

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Country-led Evaluation of Preschool Education in Timore-Leste (2012-2020)

REPORT RATING SUMMARY			
Overall Rating		72%	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			
Title of the evaluation report	Country-led Evaluation of Preschool Education in Timore-Leste (2012-2020)		
Report sequence number	Timor-Leste-60-2022-18163		
Region	EAPR		
Year of report	2022		
Office	Timor-Leste CO		
Coverage (countries)	Timor-Leste		
ToRs present	Yes		
Date of review (dd/mmm/yyyy)	February 12, 2022		
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	Yes		
Every child is protected from violence and exploitation	No		
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	Thematic area		
Evaluation type	Formative		
Evaluation strategy	Mixed methods		
Evaluation design (primary method used)	Non-experimental		
Evaluation level	Output		
Geographic scope	National		
Primary SDG(s) covered (number)	4		
EQA Summary: <i>The rater will provide top line issues for this evaluation relevant for feedback to senior management (positive and negative), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight best practice/added value elements and the level of complexity of the evaluation.</i>			
<p>The report is well structured/formatted and is written in an easy-to-understand manner. The methods used are appropriate given the evaluation's purpose, objectives, and available budget. It is evident that the evaluation team performed an in-depth desk review based on statements made, but the report sometimes uses vague language and lacks citations which would help the reader understand what is being referenced in the text itself. Providing citations would also help the reader understand how key informant viewpoints were triangulated with desk review documentation.</p>			
Recommendations for Improvement: <i>The rater will identify topline recommendations to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, resources will be cited to assist evaluation managers in overseeing future evaluations.</i>			
<p>The findings section appears to rely heavily on stakeholder viewpoints/recollections with little to no triangulation. A thorough review and revision of the findings section should have taken place in order to present the findings in a manner that clearly shows how findings were triangulated. Where triangulation was not possible, this should have also been noted so that readers understand that certain findings must be taken with caution. Further, a thorough review of the findings section and accompanying revision may have helped to remove what appear to be contradictory findings (e.g., the last finding under Section 5.4 Efficiency relates to the lack of effective communication between municipal and national levels (see Section 5.4.3, last paragraph on page 62). In the following page (Sustainability section), the report includes the following as a key finding: 'Overall, the preschool subsector has effective communication between national and municipal levels.'--see feedback below for more details). The findings section lacks citations for referenced documents (e.g., policies/directives) and these should have been included prior to report finalisation.</p> <p>The report would have benefited from more clearly laying out the recommendations (to clarify, up front, who the responsible party(ies) is/are.</p>			
SECTION RATINGS			
SECTION A: EXECUTIVE SUMMARY (weight 5%)		67%	Comments on Rating
Question 1.	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).	Yes	The report fully meets this criterion. It can serve as standalone document, and is concise.

	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 executive summary sufficiently describes the purpose of the evaluation, key findings, conclusions, and recommendations. The executive summary does not include explicit information regarding the intended audience but one of the stated objectives was to create the lessons learned and recommendations that were useable by the government and its key partners. The executive summary did not contain sufficient information regarding an overview of the preschool education in Timor-Leste. The executive summary would have benefited from a paragraph or two detailing Timor-Leste's preschool education subsector. The report would have also benefited from more details regarding the methodology (e.g., the executive summary does not note the two selected in-depth data collection municipalities (Ainaro and Lautem) and the less in-depth municipality-level data collection in Baucau, Bobonaro, Viqueque and Oecusse.
	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	As stated above, the executive summary is missing an overview of the preschool subsector in Timor-Leste. Overall, the executive summary includes all significant information required to understand the evaluation but would benefit from including information required to understand the evaluation (preschool subsector). It does not introduce any new information.
SECTION B: BACKGROUND (weight 5%)			64%	Comments on Rating
Question 2. Is the object of the evaluation clearly described?				
	i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Yes	This is a thematic evaluation of the preschool education subsector between 2012-2020. The report offers sufficient details regarding the scope of the assignment including locations and time period covered.
	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	The report does not include a description of intended rightsholders and duty bearers by type, by geographic locations, and in terms of numbers reached with disaggregation by gender, age, and disability. The report does, however, list some duty bearers (e.g., government entities, UNICEF, New Zealand Aid) and rightsholders including teachers, teacher aids, pre-school-aged children.
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 sufficiently describes recent developments in the preschool education subsector. The report would have benefited from a brief overview of the socioeconomic situation in the country (e.g., #/% of preschool aged children living below the poverty line, any disability-related statistics, gender parity, etc.) as well as cultural factors that may influence the preschool subsector.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	The report references SDG 4.2.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Partially	The report discussed the net and gross enrolment rates in country and the general benefits of and consequences to missing preschool education, but does not go into detail regarding the needs of preschool-aged children in the country.
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	No stakeholder map is presented. Relationships between stakeholders are not clearly described. However, the report does note the government's role and states that there are development partners such as UNICEF, New Zealand Aid and NGOs, but the linkages between them are not specified. The report would have benefited from a stakeholder map that clearly laid out the actors in the preschool education subsector, their roles/responsibilities, and the linkages between them.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	The report provides an overview of what actors in addition to the government contribute to the preschool education subsector. The report does not give an in-depth overview of the specific contributions and roles of key stakeholders. The report does not offer details on the evaluation's primary focus - the government's development and implementation of a comprehensive preschool education system during the 2012-2020 timeframe (which coincides with the launch of Timor-Leste's 2012 Policy Framework for Preschool Education). The report would have benefited from providing an overview of what actors are responsible for what components, etc. from 2012-2020 and how this ties into the launch of Timor-Leste's 2012 Policy Framework for Preschool Education.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)			75%	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	The purpose of the evaluation is clearly defined, but the report does not explain why the evaluation is being undertaken at this time. The evaluation's intended use and users are detailed. The report would have benefited from including some information regarding the 'why now' of the evaluation.
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 evaluation objectives are clearly stated in section 3.2 and there are no noted changes from 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 3.3 clearly lays out the thematic, geographic, and chronological scope.
Question 7.	Is the theory of change, results chain or logic well articulated?			
	i	Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Partially	While some outputs and the outcome are clear, the presented ToC contains an output and outcome that lack clear descriptions. For example, one output is titled 'system capacity' and one outcome is 'supportive contexts & home & in the community'. It is unclear what is meant by these two pieces and more detail is required.
	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).	Partially	Some links between outputs and outcomes is unclear. For example, one output is called 'system capacity' without any other qualifiers (e.g., adequate system capacity, increased system capacity, etc.) and the outcome 'access to pre-primary education programming' also lacks qualifiers and does not take into consideration that the outcome may want to be increased enrolment and/or participation in pre-primary education programming. In other words, access is important, but uptake is what the ultimate achievement (outcome) should be. The outcome 'supportive contexts at home & in the community' is vague as well. It is unclear how some of the outputs and outcomes will be measured in future assessments/evaluations or monitoring activities.
	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 created retrospectively and finalised in cooperation with the ERG.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)		77%	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 listed in Section 4.2 are appropriate for meeting the purpose and objectives of the evaluation. The corresponding evaluation criterion is noted per question. The introductory paragraph preceding the questions list state that 14 questions were finalised. However, only 12 evaluation questions are listed. The evaluation approach (Section 4.3) lists six modules (planning and budgeting, curriculum, workforce, family and community engagement, quality assurance, and enabling environment). However, the evaluation matrix, questions list, and findings are not organised by these modules. The conclusions section include some of these modules as subtitles (but not all) and the conclusions section is not presented in the same order as the six modules in Section 4.3.
	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.	Yes	Annex B contains an evaluation matrix that specifies the indicators, data sources, and data collection techniques.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?			
	i	Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	The evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives, and scope. Copies of the tools are available as an annex.
	ii	Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	The evaluation utilised appropriate data sources including document review and interviews with duty bearers at different levels of government, implementing partners, and others.

	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Yes	Section 4.4 describes the sampling strategy (purposive) and includes information regarding how the municipalities for in-depth and lighter touch were selected in coordination with the ERG. The sampling strategy for identifying individuals for interviews was simply the ERG sharing a list of who to speak with. While this alone (i.e. without the evaluation team confirming the list and adding to the list based on desk review or the snowball method) is not an ideal sampling strategy, the strategy is provided so the report fulfils this criterion.
	iv	Clear and complete description of the methods of analysis.	Yes	The report provides extensive details on how the data collected through desk review and interviews was coded and analysed.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Partially	The methodology described in the evaluation report had the potential to allow for drawing of causal connections between outputs and outcomes. However, the lack of triangulation means that some causal connections could not be drawn.
	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	The limitations section highlights some important constraints but lacks mitigation measures/how these limitations were addressed by the evaluation team. One limitation that was not highlighted was gaps in evidence regarding how specific strategies were carried out (or whether they were ever carried out). The report presents the views of stakeholders when it comes to the extent to which specific items under a specific strategy were implemented, but the lack of secondary data meant that these views were often not triangulated. The authors should consider adding a limitation regarding the availability of evidence related to the pre-school education subsector between 2012-2020.
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).	Partially	The report states that the team completed the work in line with United Nations evaluation norms, standards, and guidelines. However, explicit and contextualized reference to the UNEG obligations of evaluators or the UNEG Ethical Guidelines was not made.
	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 evaluation employed appropriate ethical considerations given the sample population and the subject matter under evaluation.
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 be 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 evaluation employed appropriate data collection (e.g., document review, interviews) and analysis techniques, but there is no evidence of innovative practice(s) used during the evaluation.
SECTION E:		EVALUATION FINDINGS (weight 25%)	70%	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	As expanded upon below in question 13.i and 13.ii, the evaluation findings lack details required to understand how/if findings were triangulated. There are many findings that appear to rely on a single stakeholder viewpoint, for example.
	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The evaluation findings are organised by OECD-DAC criteria and the findings do not make explicit use of the ToC. However, findings do respond to elements of the ToC. For example, the findings explore the inputs in the form of government leadership, policies and legislation, public demand, and financing. The findings also explore planning and budgeting, workforce development, curriculum development, family/community engagement, and a quality assurance and/or accreditation system.
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>Many findings lack sources/citations. Some findings use information from multiple informants but do not include secondary data from the document review. Findings like the one quoted below are examples of findings that should have been able to provide information from the secondary documentation and cite specific references:</p> <p>---'As noted above, stakeholders did not think Timor-Leste's targets or timelines for preschool were realistic (see Section 5.5.1). The country does not appear to have any alternative targets or plans that are more realistic. Timor-Leste also lacks policy directives with specific, integrated timelines and roles/responsibilities for preschool service provision. The country does have plans around preschool expansion and strengthening the education system as a whole (such as through improved infrastructure), but these plans do not appear to have been implemented as formal policies.'</p> <p>--- 'Regarding accreditation systems, stakeholders noted that there is a recently established system of preschool accreditation whereby organisations that want to open preschool institutions must undergo a licensing process and meet certain requirements. One MoEYS member explained that organisations are required to meet basic criteria to open preschools, including having adequate facilities and a playground or other suitable place for children to play and express their creativity. Another national-level stakeholder said that MoEYS staff collaborate with the local authorities to decide whether opening the school is logistically feasible. If so, the school can open promptly.'</p> <p>---'In addition, many preschools have play materials to align with the curriculum, due in part to the active support of ECD partners who provide</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>Findings related to question 8 are weak and do not provide sufficient information to answer the evaluation question. (The evaluation question was: '5.3.6. To what extent does the preschool education system promote equity – including reaching the most marginalised – empowerment of women and girls, respect for linguistic diversity, and inclusion of people with disabilities?') The evaluation report states that in order to answer this question, 'the evaluation assessed the extent to which Timor-Leste's preschool education subsector strategy focuses explicitly on building equity'. Assessing the extent to which the strategy focuses on building equity does not sufficiently answer the question at hand.</p> <p>Some findings appear to contradict others. For example, the last finding under Section 5.4 Efficiency relates to the lack of effective communication between municipal and national levels (see Section 5.4.3, last paragraph on page 62). In the following page (Sustainability section), the report includes the following as a key finding: 'Overall, the preschool subsector has effective communication between national and municipal levels.'</p> <p>Another finding that appears to be contradicted elsewhere is in 5.3.4. In that</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).	Not Rated	<p>This criterion does not fit well with the thematic evaluation under review. However, the report notes some causal factors leading to the (likely) non-achievement of enrolment. However, findings do not analyse the logical chain.</p>
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.	Yes	<p>The subsector did not have a monitoring and evaluation framework based on a ToC. However, the evaluation report does contain some discussions around the monitoring system sector wide: The report findings note that the country's subsector planning does not appear to address some elements that are required for a well-functioning pre-school system including 'comprehensive monitoring that captures coverage, participation and child outcomes as well as the above-mentioned quality indicators'. The report also states that Timor-Leste's education sector monitoring system considers diversity elements. However, in terms of monitoring, the report states that 'Timor-Leste does not appear to use data monitoring to track relationships between funding and resulting access, quality, effectiveness or equity in the country's preschool subsector. Municipality-level inspectors claimed that no such system exists. According to a national MoEYS leader, the Timorese government has tracked the implementation of the budget and programmes through the Dalan ba Futuru/Way to the Future system and Unit for Planning, Monitoring and Evaluation.⁸⁹ Dalan ba Futuru tracks outputs of government funding but does not track outcomes. For education specifically, the school inspection system and EMIS track school quality, although incompletely, as described in Section 5.4.3. Whilst the government does monitor the quality of the programmes they fund, the informants in the interviews did not mention using these data for any cost-effectiveness analyses that link outcomes to programme costs.'</p> <p>The report goes on to state that 'Timor-Leste took an important step with the integration of preschool data into its EMIS in 2013.⁹³ In addition, there is a set of goal-related indicators that addresses several aspects of the preschool subsector in the ESP 2020–2024.⁹⁴ However, as also discussed in Section 5.1.1, Timor-Leste does not have a unified and comprehensive preschool</p>
SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	75%	Comments on Rating
Question 15.	Do the conclusions clearly present an objective overall assessment of the intervention?		

	i	Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking (if a formative evaluation or if the implementation is expected to continue or have additional phase).	Partially	Overall, conclusions are clearly formulated; however, there are some conclusions (noted below in question 15.ii) that are not clearly formulated based on findings.
	ii	Conclusions are derived appropriately from findings, and present a picture of the strengths and limitations of the intervention that adds insight and analysis beyond the findings.	Partially	<p>Most conclusions are derived from the findings and present both strengths and weaknesses of the pre-school education subsector. However, there are some conclusions that do not clearly link back to findings. For example:</p> <p>-- However, relationships between different stakeholder organisations were viewed as existing one on one rather than as a strong, collaborative network. This view applied to both relationships between the government and outside organisations, and between non-governmental entities. (Note that the assessor cannot locate a coinciding finding for this conclusion.)</p> <p>-- Government stakeholders reported that there was good communication between the national and municipal levels. (Note that another evaluation finding appears to contradict this conclusion and it is unclear why the authors chose to only highlight government stakeholders' opinions. As stated above: Some findings appear to contradict others. For example, the last finding under Section 5.4 Efficiency relates to the lack of effective communication between municipal and national levels (see Section 5.4.3, last paragraph on page 62). In the following page (Sustainability section), the report includes the following as a key finding: 'Overall, the preschool subsector has effective communication between national and municipal levels.')</p> <p>In future, evaluation reports should reference the supporting findings (e.g., section number or EQ number) so that readers can more easily understand how evidence from the findings section has flowed into the conclusions section.</p> <p>The conclusions section summarises the findings but does not include analysis beyond the findings.</p>
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.	Yes	<p>As presented, the lessons learnt are phrased in a way that allows for wider applicability.</p> <p>The lessons learnt appear to have stemmed from the evaluation findings, however, the report would have benefited from referencing which finding(s) influenced each lesson learnt.</p>
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	Aside from lesson learnt #3 requiring more details to fully convey the lesson learnt, the section concisely presents the lessons with sufficient detail to be useful for the intended audience.
SECTION G:		RECOMMENDATIONS (weight 15%)	63%	Comments on Rating
Question 17.		Are recommendations well grounded in the evaluation?		
	i	Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Yes	The recommendations align with the evaluation purpose and are easy to understand (although some headline/bold text introductions to the recommendations do not reflect the actual recommendation - see below). Overall, the recommendations stem from findings and conclusions. The report would have been stronger if each recommendation referenced the contributing findings/conclusions.
	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	<p>The section provides useful recommendations that are often written in an actionable manner alongside specific guidance for implementation. However, the assessor has the following comments regarding the presentation of certain recommendations:</p> <p>An example of a finding that would benefit from some revision in terms of presentation is Recommendation 5 which reads 'determine how best to build upon the diversity of preschool delivery models in use' -- the recommendation is actually to undertake a national study to compare different models.</p> <p>Recommendation 7 does not have a responsible party.</p> <p>Recommendation 12's headline is 'provide preschool education to children with disabilities'. However, the recommendation within that is to carry out a situational analysis study of education for children with disabilities. The recommendation to perform a situational analysis does not necessarily lead to providing preschool education to children with disabilities.</p> <p>Further, many of the recommendations have MoEYS taking the lead on the suggested initiatives. It is unclear whether some recommendations are useful and actionable for MoEYS based on evaluation findings. For example, one finding noted that 'MoEYS currently lacks sufficient personnel with knowledge of effective preschool education standards to engage in planning on its own but receives much technical support from partners such as HANDS, UNICEF and the World Bank. One national MoEYS staff member summarised the issue well: "From the government side, especially the [MoEYS], we do not have experts, but we work with partners, because they have experts that work</p>

	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 report states that the recommendations were developed based on a combination of direct input from stakeholders in the course of interviews, and review of the evaluation data across sources to identify actionable areas of need or gaps. Based on the information presented at the beginning of the recommendations section, it appears that duty bearers in the ERG were not involved in formulating the recommendations. Rights holders were not involved in the evaluation. Earlier on in the report (Section 4.5), the authors state the following, however, this was not reflected in the recommendations introduction paragraph: 'To gain at least some parental input, the Timorese NGO Alola met with some parents in communities in March 2022 to ask for their feedback and input on the recommendations from this evaluation, particularly in relation to family engagement.'
Question 18.		Are recommendations clearly presented?		
	i	Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Partially	The recommendations section clearly justifies its prioritization of recommendations. However, some of the recommendations do not clearly identify the group(s) or duty bearer(s) responsible for action for each recommendation. The report would have been stronger if each recommendation included more actionable language by clearly referencing responsible party(ies). As written, recommendations include language like 'This is the kind of activity that UNICEF or HANDS would be well positioned to facilitate (or the ECD Working Group, if it is relaunched), but the MoEYS should be viewed as a strong leader in this activity.' A stronger way of stating this recommendation would have been to identify the responsible finding up front for easy reference and use stronger language.
SECTION H:		REPORT STRUCTURE AND PRESENTATION (weight 5%)	93%	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.	Yes	The report fully meets this criterion.
	ii	Annexes include: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework (unless included in report body), list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology, information about the evaluator(s), etc.	Yes	The report fully meets this criterion.
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).	Yes	The report is well formatted and includes clear titles/subtitles to help with report 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 report structure follows UNICEF guidelines for evaluation reports.
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.	Partially	The ToR states that the final report should not exceed 45 pages exclusive of executive summary and annexes. The report length is approximately 70 pages in length. The assessor has marked this as 'partially' given that the report is not verbose in nature and contains important information regarding the pre-school education subsector over an eight-year period.
	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 generally easy to read and free from grammar, spelling, and punctuation errors. There are some small errors (e.g., SGD instead of SDG) and these may have been caught through a thorough proofread and copyedit. Note that the report visuals are not accessible to individuals using screen readers and requiring alternate text for these visuals would increase accessibility.
	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.	Yes	The report makes use of call-out boxes, photos, tables, and other visuals to help convey key information. Tables/charts are clearly labeled and referenced in the text.
SECTION I:		EVALUATION PRINCIPLES (weight 10%)	76%	Comments on Rating
Question 22.		Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity?		

	i	Reference and use of rights-based framework, and/or CRC, CCC, CEDAW and/or other rights related benchmarks in the design of the evaluation.	Yes	The evaluation report references CEDAW, CRC, and CCC.
	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	Rights holders were not included in this evaluation. Duty bearers were involved as members of the ERG as well as participants in data collection.
	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 empowering and inclusive language and differential results are assessed wherever possible (e.g., regarding children with disabilities or linguistic minorities).
	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	While the report does not explicitly state that it analyses the extent to which the evaluand addresses child rights and LNOB, the report is centred around a child's right to education and includes discussion and analysis of marginalized groups including children with disabilities and linguistic minorities (linguistic diversity), and girls.
Question 23.		Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards with detail provided below)	4	
	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.	Partially 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? The evaluation includes assessments of what information is collected to monitor the system, but does not provide an assessment of whether sufficient information is collected related to human rights and gender equality results. Data such as enrolment rates are sex disaggregated (figure 4).</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? The evaluation does not include an objective specific to assessment of human rights and gender equality considerations. However, the subject matter of the evaluation relates to human (child) rights and there are gender equality considerations within the evaluation questions and subquestions.</p> <p>c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? No, but gender and human rights were 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? Yes (question 8). However, the findings related to this question are weak and do not provide sufficient information to answer the evaluation question. (The</p>
	ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	<p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? The report states that 'Team members paid special attention to differences amongst respondents according to gender, location and role, as well as to any gender equality and human rights issues.' However, there is no evidence that collected data was disaggregated by sex or how the collection and analysis methods integrate gender considerations.</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations? Yes.</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? Yes.</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 frame includes a diverse range of stakeholders (duty bearers only).</p> <p>e. Were ethical standards considered throughout the evaluation and</p>

iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially 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? The evaluation does not include a background section with an intersectional analysis of the specific social groups affected by the issue, but does include reference to the relevant normative instruments 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? Yes.</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? N/A.</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? No.</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?