

# GEROS Evaluation Quality Assurance Tool

Version: September 2021

## Evaluation of the modeling of Alternative Early Childhood Education

REPORT RATING SUMMARY			
<b>Overall Rating</b>		71%	Satisfactory
●●●●●	Exceptional (96% - 100%)	5	
●●●●○	Highly Satisfactory (87.5% - 95.99%)	4	
●●●○●	Satisfactory (62.5% - 87.49%)	3	Meets UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with confidence
●●○●●	Fair (35% - 62.49%)	2	
●○○●●	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
<b>Title of the evaluation report</b>	Evaluation of the modeling of Alternative Early Childhood Education		
<b>Report sequence number</b>	Rep of Uzbekistan 66/2024/19437		
<b>Region</b>	ECAR		
<b>Year of report</b>	2023		
<b>Office</b>	UNICEF Uzbekistan CO		
<b>Coverage (countries)</b>	Uzbekistan		
<b>ToRs present</b>	Yes		
<b>Date of review (dd/mmm/yyyy)</b>	June 26, 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	Yes		
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	Yes		
Gender equality (cross-cutting)	Yes		
Humanitarian action (cross-cutting)	No		
<b>Evaluation object</b>	Project		
<b>Evaluation type</b>	Summative and formative		
<b>Evaluation strategy</b>	Mixed methods		
<b>Evaluation design (primary method used)</b>	Non-experimental		
<b>Evaluation level</b>	Output & Outcome		
<b>Geographic scope</b>	National		
<b>Primary SDG(s) covered (number)</b>	4		
<b>EQA Summary:</b>	<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>Overall, this is a reasonably solid report. It is well laid-out, contains all of the elements required in a good-practice evaluation, is clear and mostly well-written. In short, it achieves well what it set out to cover as specified in the evaluation design.</p> <p>The following are the top-line findings:</p> <p><b>Strengths</b></p> <ul style="list-style-type: none"> <li>•Well structured: All elements necessary are present, including useful annexes.</li> <li>•Useful and concise context and background</li> <li>•Effective Use of Theory of Change (ToC): The evaluation makes explicit and effective use of the intervention's results framework/ToC, particularly in assessing project outputs and outcomes under effectiveness criteria, which helps clarify the linkage between activities and intended impacts.</li> <li>•Robust Methodology: Data collection, analysis, and sampling methods are well-suited to the evaluation's needs. The description and application of these methods are detailed, enhancing the reliability of findings.</li> <li>•Multiple Data Sources: The evaluation utilizes various data sources, including the voices of rights-holders, enhancing the richness and comprehensiveness of insights gathered, which contributes to a robust triangulation of data points.</li> <li>•Strong Ethical Approaches: a very comprehensive approach to ensuring adherence to ethical evaluation principles.</li> <li>•Systematic approach to the analysis and presentation of findings - with one minor exception, there is full adherence to the evaluation analytical framework.</li> </ul> <p><b>Weaknesses</b></p> <ul style="list-style-type: none"> <li>•Some clearer description of the differences between the alternative ECE models and the 'standard' models would be useful</li> <li>•In places, it is somewhat general and superficial in its approach and analysis, and thus only partially achieves the evaluation's intended purpose</li> <li>•It is clear that the ECE project has achieved positive outcomes for children, but the aspect of added value beyond more broad-based investment in ECE overall (in line with what the Government of Uzbekistan has already been doing) is not fully clear.</li> <li>•A more 'experimental' approach, which sought to explore some counterfactual (i.e. whether simply rolling out existing models in areas currently lacking ECE provision) or controlled for different features of the project (notably the fee-free aspect), could have provided more useful insights for UNICEF and the Government.</li> </ul>			
<b>Recommendations for Improvement:</b> <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>			

With respect to the most substantive criticism of the evaluation - that it is insufficiently focused on the added-value of the alternative models, this is an aspect that should have been addressed at design stage. The evaluation itself followed its proposed data collection and analytical approach reasonably well. This is the role of the Evaluation Reference Group and/or Evaluation Manager. Feedback could also have been provided to the evaluators at drafting stage, to ask them to provide some additional insights or explore/unpack the nuances of specific findings that would have mitigated this criticism somewhat.

Some other, more specific areas for improvement are as follows:

- Depth and Nuance of Analysis: Certain areas within the report are only briefly examined, such as the potentially confounding effect of free ECE on the project success. A more nuanced analysis could provide deeper insights into underlying issues and their implications.
- Consistency in Measurement: There are uncertainties in how some metrics, such as "quality" of financial inputs, are measured, which may affect the accuracy and usefulness of the evaluation.
- Actionability of Recommendations: While the recommendations are relevant, they could be enhanced by adding specific, actionable steps for each stakeholder to aid in implementation.
- Sequencing of Content: The numbering of evaluation questions and sub questions is confusing, restarting under each criterion. The final sub question is missed in the analysis.
- Diversity in Disaggregation/GEEW/Vulnerability: The evaluation is quite superficial in its treatment of gender, disability and misses ethnicity dimensions, and does not meet its ambition for a comprehensive equity analysis.
- Clarity and Readability: The report exhibits minor inconsistencies in language use (US vs. UK English) and contains a range of language errors and some (minor) use of jargon.
- Representation of Management and Quality Assurance: The exact membership, role and contributions of the Evaluation Management Group (EMG) in the quality assurance and final sign-off processes are inadequately described.

SECTION RATINGS			
<b>SECTION A: EXECUTIVE SUMMARY (weight 5%)</b>		100%	Comments on Rating
<b>Question 1.</b> Can the executive summary inform decision-making?			
i	Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Yes	The executive summary, at five pages, meets the standard for length. It has sufficient detail to represent the main content of the evaluation report well without being overly in-depth. It can serve as a standalone document useful for informing decision making.
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	All of the specified elements are present and correct in the summary and therefore represents an accurate synopsis of the full report.
iii	Includes all significant information needed to understand the intervention and the evaluation AND does not introduce new information from what is presented in the rest of the report.	Yes	The summary provides an accurate representation of all of the main elements of the report. There is no additional analysis or information that is not present in the rest of the report.
<b>SECTION B: BACKGROUND (weight 5%)</b>		79%	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	Following on from the description of the context and the rights-holders intended to benefit from ECE in Uzbekistan, section 1.4 provides good detail on the specifics of UNICEF's contribution to ECE via the programme in question. The section describes the overall conceptual framework of the programme (the different ECE models being implemented), practical implementation considerations (roles and responsibilities of different stakeholders) and details the timeline of the intervention conceptualisation and development, through implementation and includes budget details (from UNICEF, but notes that the value of Govt contributions (many of which were in-kind) could not be determined. The status of the objects of the evaluation is not completely clear, but it can be inferred that the specific ECE centres were operational at the time of research, given that the model was being expanded in 2023. What is less well described - and this is a flaw that emerges across the report - is what makes the pilot projects different to regular ECE models. The 'alternative' models are discussed in depth, and many differences between these and 'standard' models are alluded to throughout the analysis, but there is no clear description (ideally a table format that presents the existing and alternative model(s) of what makes the models 'alternative'. This is important to be able to establish the added value of the new models. Table 1 presents information on two of the three models, but one is absent, and there is no easily-accessible comparison data using the parameters in the table. An expanded table would be most useful.
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).	Partially	The introductory section (1.1-1.4) provides ample detail on the status and needs of children suitable for ECE across Uzbekistan, and also more specifically (in section 1.4) related to the project under evaluation. The difference between ECE status/enrolment in rural vs. urban areas is noted from the outset in the background and context, as it is a key element of the initiative. The specifics of the number of children to be reached via the pilot ECE centres are provided, both in terms of absolute numbers and the inclusion criteria for the children. The institutional partners (duty-bearers) are noted in section 1.4, i.e. UNICEF, central Govt/MOPSE, local government. This is clearly described on page 10. Gender considerations are noted also (gender parity was implemented) as well as economic considerations, although disability has not been noted as an inclusion criterion and thus disaggregated data on disability is rudimentary (some quantitative data is included in section 1.3.1 and again in the findings e.g. under finding 1).
<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	Sections 1.1 through 1.3 provide a very comprehensive background to the overall context of ECE in Uzbekistan, with a full description of the legislative/policy environment since 2017 through to 2023 (the time of evaluation). Some useful comparisons are also provided. For example, in section 1.3, there is a comparison of Uzbekistan performance on ECE enrolment in 2017 (i.e. pre-programme implementation) with other countries in the region. The different dynamics that affect enrolment in ECE are noted, including some of the socio-economic drivers of ECE access, geographical and demographic disparities and investments over time. Most (though not all) data and statements are referenced, although some more detail on these would be welcome (this is discussed further below).

	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	Section 1.3.2 discusses the linkages of ECE in the context of Uzbek educational policy with the relevant SDG (#4) and how these are explicitly articulated in Uzbekistan's national strategy 2030. The sub-SDG relevant to ECE (4.2) is also noted with the appropriate indicator/target stated.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	The introduction, background and context sections comprehensively describe the situation of early childhood education in Uzbekistan, with section 1.2 describing the demographic and socio-economic conditions of children in the country. This complements the following section which describes the early educational status in Uzbekistan from 2017 onwards and notes the outcomes of some of the investments in ECE on children's early education indicators (notably enrolment) as well as existing disparities.
<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).	Yes	While a graphical stakeholder map has not been provided, there are a limited number of stakeholders involved in the project and their roles and responsibilities are clearly and concisely described in section 1.4 (page 10) - the description of the responsibilities/linkages between stakeholders are described in narrative format, but this is adequate. No further information is required here.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	As above, the roles and responsibilities subsection in section 1.4 notes the individual contributions of all stakeholders. While the UNICEF contribution (in USD) is described, that of national stakeholders is noted as not lending itself to quantification (or at least, such data was not made available to the evaluators). Instead, a qualitative description of their relative contributions are noted.
<b>SECTION C:</b>		<b>EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)</b>	83%	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.	Yes	Section 2.1.1 defines the purpose of the evaluation, which is clear, if quite detailed, and has a clear rationale at the point in time. The evolution of ECE in Uzbekistan is clearly explained in the previous sections and therefore the need for an evaluation of this project is justified. Key users/audiences for the evaluation are also included in this section, with details of the primary and secondary audiences provided.
<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 six evaluation objectives are clearly stated in section 2.1.2. These are largely unchanged from the TOR, although no reference to the TOR and any changes (or lack thereof) is noted.
	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	This is clearly and concisely covered in section 2.1.3, with the scope divided into programmatic (evaluand), geographic and temporal dimensions. The scope is well-defined with a (minor) limitation noted, i.e. the visiting of eight of the ten supported ECE groups (due to time constraints). No other limitations on the programme extent or data gaps were noted.
<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.	Partially	The theory of change for the original project was present in the TORs for the evaluation. The evaluators are expected to reconstruct/revise this (per the TOR) to "reflect the current operational realities". The TOC has indeed been revised and is presented in section 2.2, both in narrative and graphical forms. This said, both forms of the TOC as presented (the original in the TORs and the revised in the Evaluation Report) are somewhat generic TOCs relating to children's early education. There is little in the way of specific detail relating to the anticipated changes that the project aspires to bring about, i.e. what is the additional value of the different ECE models being piloted via the project. Further, the nature of "current operational realities" that were intended to drive the changes to the original TOC are not clearly specified. This may have been a gap that should have been addressed at project conception stage.
	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	The report contains both a theory of change and an evaluation matrix - with the TOC presented in narrative and graphic forms. These are comprehensive, although, as noted above, the "added-value" component of the ECE centres under evaluation does not seem fully fleshed-out. Further, the definition of impact of the project is debatable - as per the reconstructed TOC, the ultimate impact of the project is "equal opportunities for young children to develop and learn" - it could be argued that this is a project outcome rather than impact, and contributes to equitable and inclusive school enrolment. As noted above, this may have been better addressed at evaluation inception and, in any case, as impact-level measurement is noted in the report as not being feasible for an ongoing project, the issue is somewhat moot.
	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 assessment of the theory of change forms part of the research plan. This is noted clearly in the evaluation matrix (in Annex C) where validation of the TOC elements is covered under the Relevance criterion. As discussed above, the project TOC has been revised by the evaluators for the purposes of the evaluation.
<b>SECTION D:</b>		<b>EVALUATION DESIGN AND METHODOLOGY (weight 20%)</b>	77%	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 are clear, concise, are fully and explicitly laid out in line with the evaluation criteria chosen (section 2.3) and are well suited to addressing the evaluation purpose - although bearing in mind the previous feedback in 7i above that the evaluation could go more deeply into the relative added value (compared to other models) of the pilot centres. As also noted above, this issue could/should have been noted in the original concept for the evaluation/TORs, so is of peripheral relevance here.
	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 annexes (annex C) include an evaluation matrix that is similarly clear and concise. The specific evaluation questions are all represented (although there are some very slight omissions, and the impact/sustainability sections are arranged differently between the questions in the main report and the matrix). The matrix has a range of sub-questions under each criterion, and each has one or more indicators associated with it. Some of the indicators may be somewhat challenging to measure, e.g. "quality" of financial inputs (under the efficiency criterion) - measuring "quality" of financial inputs is unclear. Also, disability is not well represented in the indicators in relation to equity - most disaggregation is by age/gender with some ethnicity disaggregation (though not everywhere e.g. the equity indicators under effectiveness Q4 is only disaggregated by age and gender of children, if at all). Finally, evaluation questions/sub questions are not numbered sequentially (numbering restarts under each criterion) - this should be addressed for clarity.
<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 methods chosen were well-suited to the evaluation and the description of the application of the methods is excellent - clear and adequately detailed to provide the reader with a full understanding of the approach. The use of multiple methods and multiple stakeholder groups is perfectly suited to this type of evaluation.
	ii Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	The evaluators obtained and analysed data from qualitative (primarily) and quantitative sources for the evaluation, with a robust mix of primary and secondary data. Essentially all available sources of data were utilised, from duty-bearers to rights-holders, without going to the level of children themselves, which would have imposed a considerable more complex level of oversight on the evaluators. As it is, the sources are well-pitched for the purpose of the evaluation.
	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 sample and strategy for each of the primary data collection tools is provided under the respective description. The general strategy was purposive - a non-probability approach that seeks out those knowledgeable about the initiative. This is most suitable for the project. As noted, the sampling strategy involved the perspectives of both duty-bearers and rights-holders (i.e. parents of children in ECE centres). A limitation of the sampling was the preponderance of female voices (among both rights-holders and duty-bearers) - this is a reflection of the dominance of women amongst early-learning teaching and in caregiving in Uzbekistan, a feature adequately explained in section 2. No other perspectives were sought per the design. While voices from a sample of non-participating teachers/parents (i.e. a counterfactual) might have delivered some useful insights, this was not part of the evaluation design/scope from the outset.
	iv Clear and complete description of the methods of analysis.	Yes	Sections 2.4.6 and 2.4.7 describe the analysis of the qualitative and quantitative data, respectively. This is brief, but all the required information is present (although a footnote noting that NVivo is a qualitative analysis software package might be included, but this is a minor suggestion only).
	v Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The methodology for the evaluation is comprehensive, insofar as it captures the perspectives of diverse stakeholders at different levels. It also engages comprehensively with the evaluation targets, i.e. the ten pilot ECE centres are all included in the research. The evaluation questions, sub question and most of the indicators should lead to good evidence to link causal factors between outputs and outcomes, although, as noted above, there is limited scope in the evaluation to compare the pilot ECE centres with existing centres systematically. The data collection tools (included in the annexes) are comprehensive (as well as tailored to individual stakeholder groups - a positive feature) and should lead to strong findings.
	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	There are two limitations noted in the relevant section (2.4.8), which seems few. For example, the challenge of measuring impact - noted in the introductory section - is not explored in the limitations. Nor is the challenge of identifying markers of vulnerability (alluded to in other places in the report) such as ethnicity or disability. Further, there is no description of any mitigating actions. Some more consideration of the different limitations should be given.
<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	There is explicit reference to UNEG and UNICEF ethical guidelines in section 2.4.9, with footnotes linking to the relevant standards. The section provides ample explanation of the ethical safeguards and provisions in both the design and implementation of 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	These ethical safeguards are well-described in section 2.4.9, with all tools annexed including the preamble text regarding consent from participants. The evaluation did not involve data collection directly from children - parents were involved in FGDs as rights-holders. Annex G further elaborates on these ethical procedures - both in terms of directly collecting primary data and the management of data safely and securely.
<b>Question 11.</b>		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	There is no particular usage of any innovative practices in this evaluation. This is potentially a missed opportunity as the particulars of the project (a small, compact pilot education initiative, limited geographical range, easy stakeholder access) lend themselves well to some use of innovative data collection techniques. For example, outcome harvesting could have been used among parents (perhaps via the post-FGD questionnaire) to collect data on ways that the ECE centres participating in the pilot improve (or otherwise) on traditional ECE models.
<b>SECTION E: EVALUATION FINDINGS (weight 25%)</b>			<b>58%</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 analysis and findings section almost completely systematically replicates the structure of the evaluation questions/sub questions as laid out in table 2, except for the very last sub question (under sustainability, Q3) which is absent from the report.
	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The analysis does explicitly assess the project TOC outputs and outcomes under the effectiveness criterion (Q1 and Q2). The level of achievement of the outputs and outcomes and the evidence supporting this is discussed extensively across these sections.
<b>Question 13.</b>		Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		
	i	Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.	Partially	Overall, the report relies well on the different sources of evidence gathered, with data from rights-holders and duty bearers presented across responses. The analysis is somewhat thin in places, however, with, for example, positive assertions by institutional stakeholders as to the level of community ownership of ECE centres not supported by data from the community members themselves that could have bolstered the finding. In other cases, the evaluators make good use of available data (e.g. enrolment data) to demonstrate the outcomes of the initiative, but there are instances where some triangulation with secondary data sources would have changed the tenor of a finding. For example, under Effectiveness Q4, the evaluators note gender disparities amongst teachers being an issue (96% of ECE teachers are female) and discuss factors. However, this is a feature of pre-school teachers worldwide, with the Uzbekistan gender imbalance exactly the OECD average. Thus, there is more to the issue than the analysis might portray. There are other examples of where some unpacking of analysis is warranted highlighted in the text.
	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	While there is generally good linking of findings with evidence, in many cases the analysis is quite brief, with evidence or factors that determine results alluded to overly briefly. For example, under the Effectiveness criterion (Q1), high staff turnover is noted as a limitation on the effectiveness of the training of teachers and hence skills and hence teaching quality. This is the first time this limitation is mentioned and is not explored in terms of the scale of the issue, the reasons/factors involved, whether it is just the pilots where this is experienced etc. Some unpacking does come later in the report. There are quite a few instances of key areas of analysis like this that are lacking - more are highlighted via in-line commentary in the report. A review of some of the depth/nuance of the analysis would have benefited the report.
	iii	Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyse the logical chain (progression -or not- from implementation to results).	Partially	There is plenty of discussion and identification of different causal factors related to the results of the project. The evaluators make good reference to the collected evidence from different stakeholders and explore to a certain extent the different dimensions of issues noted. In places, however (as discussed above), the analysis can be somewhat superficial and would benefit from some additional unpacking and/or reference to additional explanatory data. For example, the cost of ECE (both private and public) is noted as an inhibitory factor to low-income families. The monthly costs are presented on pp 34/35, but the proportion of household income that this represents is not provided, so the reader has no frame of reference. Further, the fee-free nature of the ECE pilot doubtlessly influenced popularity and therefore enrolment. This potentially confounding factor is not explored to any significant extent. There is also a question as to whether some children enrolled might have been taken from other (fee-paying) ECE centres to avail of the free services.
<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.	Partially	The projects monitoring system is briefly explored under the Effectiveness criterion (Q1) - perception of teachers on the child progress monitoring cards (a monitoring tool) is discussed in one paragraph. While child monitoring is mentioned in other places, this is the extent of the analysis of the project monitoring. Curiously, the recommendations section includes one recommendation on monitoring of the project (#7), which is not linked to any specific section in the analysis portion of the report. This gap should be addressed.
<b>SECTION F:</b>	<b>EVALUATION CONCLUSIONS &amp; LESSONS LEARNED (weight 10%)</b>	<b>88%</b>	<b>Comments on Rating</b>
<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).	Yes	The conclusions are clear - they are well written, concise and directly reflect the purpose and objectives of the evaluation as articulated in the analytical plan (although equally, reflect the research plan's weaknesses - more on this below). They contain both summative and formative/forward-looking elements, as prescribed by the purpose and analytical approach, and integrate well into the lessons-learned and recommendations sections that follow.
	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	In as much as the conclusions are a reflection of the analysis and findings, they meet the standard well - they neatly and clearly encapsulate the overall analysis and findings. They do not, however, go beyond this to add any additional insights, and reflect some of the weaknesses of the main report. For example, the added value of the alternative ECE centres and models over existing models could be made clearer - i.e. what is it exactly that makes the alternatives 'better' than the existing? As it stands, the conclusions read more like an endorsement of ECE in general. Secondly, a key characteristic of the alternative models - that they are free of charge - could be explored as a significant factor in their popularity/success, and could usefully be discussed here (as could the global trend in offering free ECE) as this is a significant determinant of sustainability. Other potentially important factors that may have influenced the reported success of the project (such as the access to buses, or a more nuanced analysis of gender/disability responsiveness) could also be discussed here, as avenues of potential future research.
<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	There is a short, but useful lessons-learned section in section 4.2. The lessons are all clearly derived from the analysis and findings, and have wider relevance. There are probably some additional lessons that could be gleaned (e.g. the added-value of the technical assistance costs (40% of the UNICEF contribution) which included an external expert and a "study visit" to Poland by government officials), but overall it is a strong section.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	The section is clear and concise, with adequate detail to be useful.
<b>SECTION G:</b>	<b>RECOMMENDATIONS (weight 15%)</b>	<b>63%</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.	Partially	As was noted with respect to the conclusions section (15ii above), the recommendations are well aligned with the evaluation purpose, the analysis, findings and conclusions - although they are not systematically linked to numbered findings or conclusions, which is something that could be addressed. They also reflect some of the weaknesses of the analysis (per the above) and could be strengthened if these weaknesses are addressed. For example, further research on the likely added-value and feasibility of fee-free ECE (and its impact on the existing provision of fee-paying ECE services) is most likely warranted before any policy decision is made in this regard (as is noted under recommendation 1).
	ii Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	The recommendations are indeed useful and relevant to the intervention. They could be made more actionable, however. As they stand, each recommendation is a single paragraph of narrative text which could be accompanied by bulleted action points for individual stakeholders to assist in implementation.
	iii Process for developing the recommendations is described, and includes the involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	Partially	The preamble to the recommendations section notes a robust process of recommendation development, consisting of consultations with UNICEF, MOPSE and the evaluation reference group (the membership of which is not fully specified in either the evaluation report or inception report). There is no discussion of involvement of rights-holders in the development of the recommendations.
<b>Question 18.</b>	Are recommendations clearly presented?		
	i Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Yes	Notwithstanding the suggestion to itemise action points within each recommendation made in the previous comment, each recommendation has an assigned priority, relevant stakeholder(s) and recommended timeline for implementation. This is clear and useful.
<b>SECTION H:</b>	<b>REPORT STRUCTURE AND PRESENTATION (weight 5%)</b>	<b>79%</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	All of the specified elements are present and correct in the opening pages of the report. No errors or absences were noted.
	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	There is a full suite of annexes presented in a second volume. This covers all of the specified mandatory elements plus a list of the qualifications/experience of the evaluation team. The annexes are useful and relevant to the report. There is also a useful additional section (after the recommendations but before the annexes) on cost-modelling of a scaled-up project presented in section 4.4. While useful, this might be better placed as an annex, rather than an addendum to the recommendations.
<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).	Partially	Overall the structure is solid, with clear titling and heading numbering. However, the numbering of the findings sections is somewhat confusing, with each evaluation criterion numbered (3.1, 3.2 etc.) but the individual subheadings under each (corresponding to an evaluation question) starting at 1 for each subsection. Thus, referencing a specific evaluation question by number is difficult (as there are multiple EQ1s, EQ2s etc.). Each evaluation questions/subheading should be given a unique number for easier reference. Further many of the footnoted references are not complete - the specific documentary source should be noted, not the organisation and year. References should be accessible to the reader.
	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 structure is in line with UNICEF and good practice guidelines.
<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 report is 65 pages (excluding the opening pages and executive summary, and the annexes), which is very reasonable and in line with the TOR (60 pages). As noted above, some consideration could be given to moving section 4.4 to an annex, to bring the report length down to 62 pages.
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	Although largely strongly written, there are some issues: The text uses US spelling in places, UK in others - the authors should be consistent and use the UN standard. There are a range of small language errors that could be improved (many of these have been noted via in-line commentary) and some use of jargon (e.g. "makhalla" - the term is used several times without explanation until page 29 - note the explanation at the first instance only). There are several instances of a footnote bookmark error (noted via comments). Otherwise the report is quite easy to understand, accessible and well-written.
	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	By and large, the report makes good use of tables, infographics and charts. While the content of some tables (e.g. Table 1) could be expanded, this is an analysis issue discussed above. Some small tweaks could be made to enhance the impact of the charts - specifically, Figure 1 would be enhanced if the percentage of the population of children could be included, rather than numbers, which have no objective reference; Table 1 as noted above, Table 2 has a number of editing issues (see in-line commentary) Table 3 has an unclear term ("methodologist", inconsistent spelling in Figure 3).
<b>SECTION I:</b>	<b>EVALUATION PRINCIPLES (weight 10%)</b>	<b>55%</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.	Partially	The design of the evaluation, following on from the direction of the TOR, incorporates definite language around human and child rights in the purpose of the evaluation, and one of the evaluation questions specifically purports to assess how "equity, gender and human rights considerations" were integrated into the design and planning of the ECE centres. The focus on 'child rights' appears to have been watered down in this context, and indeed in the indicators associated with this question in the evaluation matrix, there is no mention of child (or human) rights. While there is mention of Uzbekistan's ratification of the CRC and CRPD in the context discussion (box 1, pg. 6), there is little further use of any specific standards related to these in either the design or analysis.
	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 are present in the evaluation in the form of parents of the children attending the ECE centres. There is good engagement with this group via primary data collection and their voices come across in the analysis. There is no representation of rights-holders on the ERG (although the actual composition of the ERG is not fully spelled out, the institutional representation on it is described).

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).	Partially	There is no apparent bias in the language used in the report - while "rights-holder" and "duty-bearer" terms are not used, neither are "beneficiaries" - instead the report speaks of the children who attend the schools. Disaggregation is to a degree - the report clearly notes the almost parity of enrolment across the genders in the ECE centres, and does discuss the gender imbalance of staffing (although this may be somewhat misplaced, as discussed in 13i above). However, there is little exploration of other dimensions of vulnerability, such as quantitative data on enrolment of children with disabilities, and the report explicitly notes that data on ethnic background was not captured by the evaluation (via primary or secondary data), so it is limited to the above elements.
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.	Partially	While a core aspect of the project and the evaluation is the promotion/assessment of equity on the basis of rights, gender, socio-economic deprivation and (sometimes) disability, it is somewhat superficial in assessing these dimensions, as well as being not systematic in analysis of disability in particular. For example, under Effectiveness Q4, disability is discussed, but there is little detail beyond a positive rating of the "overall inclusiveness" of the centres and the approach. But it lacks both data and sophistication - there is no data on the % of enrolment of children with disabilities (compared to the population). It notes that some facilities need to be more accessible - but does not explain what is it about the facilities that makes them inaccessible. There is no discussion of any proactive recruitment measures for children with disabilities, nor of how differential needs (physical/cognitive etc.) may be met. Overall, similarly to the ECE centres (from what can be determined), the report is more "gender/disability blind" (i.e. not deficient, but not adequately proactive) than "gender/disability responsive". It would benefit from a more robust and critical analysis in this regard.
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	GEEW dimensions are clearly laid out in the scope of the evaluation, with a specific question (relevance, Q3) devoted to "equity, gender and human rights considerations". However, the analytical framework (specifically the indicators related to the question) has a quite minimal interpretation of these dimensions, seeking to explore the inclusion of quite general "cultural, economic, political, social factors" in the project, as well as perceptions among policy makers and implementers, and the gender breakdown of teachers, which is not of significant relevance.
ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	Despite the rather limited approach to GEEW in the analytical approach, the data collection tools do operationalise these elements better, with specific question sets for different stakeholders on the inclusion of vulnerable children (gender, disability, ethnicity) in project targeting, uptake/access of services and implementation.
iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially integrated	As noted above, while the design of the report integrates gender and other vulnerability criteria, it is somewhat superficial in assessing these dimensions. For example, under Relevance EQ3 (pg. 28), the discussion on equity notes measures to ensure this. However, the "absence of discrimination" or emphasis on "equal opportunities" (noted in the text) is the equivalent of being "gender blind" or "equity blind" - many who suffer discrimination find it more challenging to avail of the so-called equal opportunities on offer and hence need more proactive measures to ensure true equity. The evaluation does not appear to move much further beyond the rather superficial description of some design or implementation measures to highlight potential issues.

### 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>