

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Evaluation of the UNICEF Contribution to Improved Teaching and Learning

REPORT RATING SUMMARY

Overall Rating	66%	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	Evaluation of the UNICEF Contribution to Improved Teaching and Learning
Report sequence number	Evaluation Office 65-2023-19114
Region	HQ
Year of report	Mar-23
Office	UNICEF Evaluation Office
Coverage (countries)	Global
ToRs present	Yes
Date of review (dd/mmm/yyyy)	November 23, 2023
Name of review firm	IOD PARC

CLASSIFICATION OF EVALUATION REPORT

Management of evaluation (Managerial control and oversight of evaluation)	UNICEF managed
Unicef goal areas (Alignment with strategic plan priorities)	
Every child survives and thrives	Yes
Every child learns	Yes
Every child is protected from violence and exploitation	Yes
Every child lives in a safe and clean environment	Yes
Every child has an equitable chance in life	Yes
Gender equality (cross-cutting)	Yes
Humanitarian action (cross-cutting)	Yes
Evaluation object	Thematic area
Evaluation type	Summative and formative
Evaluation strategy	Mixed methods
Evaluation design (primary method used)	Participatory
Evaluation level	Output & Outcome
Geographic scope	Multi-region/Global
Primary SDG(s) covered (number)	4

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

This evaluation report is satisfactory and can be used by decision makers. The complexity of the evaluation in covering the global span of data sources is well noted. The main strengths of the report and process are as follows.

- The executive summary is clearly presented and within the page limits.
 - The description of rights holders and duty bearers includes state and non-state actors and strategic partners, and a detailed list of activities undertaken by UNICEF is provided and UNICEF's role in global partnerships.
 - The global context includes the impact of the pandemic on education, and human rights goals in SDG 4 and the relevant UNICEF strategies are referenced.
 - The purpose, objectives and scope are clearly set out as specified in the TOR.
 - The Theory of Change for the UNICEF Strategic Plan (2018-2021) Goal Area 2 was effectively modified by the evaluation for the pandemic emergency response and contains relevant inputs, outputs and outcomes.
 - The evaluation is well structured according to three themes, and the questions modified by the evaluation team and agreed with management and the reference group; the evaluation framework is well developed.
 - The methodology is strong and data collection tools and methods of analysis are clearly presented.
 - The findings and conclusions are cohesively presented and include both the negative and positive. The conclusions are written as higher levels of analysis.
 - Value is added through a rigorous process to select case study countries using specially designed selection criteria.
- The main issues are as follows.
- The recommendations are not strongly linked to the conclusions and each recommendation is not supported by sufficient details on its implementation and the responsible parties.
 - There is insufficient reference in the findings to the planned outcomes in the causal chain in the Theory of Change .
 - Overall, there is weak attention in the findings to gender equality relative to UNICEF's strong focus on the equity issues.
 - There is weak attention in the methodology and findings into the situation of girls' education relative to the magnitude of those issues.
 - Need for a clear tracking of the phases and management of the evaluation process such as the TOR and inception report, and the reference group roles and responsibilities.
 - The human rights language is scant among the background, findings, conclusions and recommendations.
 - The references to UNICEF's response under the Education in Emergencies global strategies can be more frequent, including the guidelines issued by UNICEF, INEE and others, as well as in the findings to note how well these guidelines were respected.
 - The main report lacks numbers in the sections to correspond to the Table of Contents and contents in both the main report and appendices are not hyperlinked for easier navigation.
 - The title of the report does not reflect its unique focus on education in the pandemic emergency response.
 - The report does not have a limitations and mitigation measures section.
 - The executive summary is missing the audience, the management and the main guiding principles in order to present a more complete picture to readers.

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

- Restructure the recommendations section (6) to more strongly link recommendations to the conclusions, citing the relevant conclusions by number, and support each with details on the suggested implementation measures noting who is to participate. Suggest to present the recommendations in order of priority with an approximate time frame for their achievement. Place the justification with relevant conclusions first, followed by the recommendation and then operational considerations for each sub-recommendation. Refer to evaluations with high QA scores to note how the recommendations section can be structured.
- Add an introductory paragraph to the recommendation section (6) reviewing the evaluation purpose and objectives, explaining the process to develop the recommendations and who was involved, and mentioning the next steps in the process.
- Fortify the findings sections (5) with stronger references to the planned outcomes in the causal chain in the Theory of Change, stressing the progress activities are making toward the outcomes and why data on outcomes may be missing, adding more language regarding achievements and shortfalls relative to the UNICEF and other rights based guiding principles such as for Education in Emergencies.
- Use and note the appropriate UNSWAP guidance (see guidelines below) to more clearly situate gender equity among the objectives, rights holders' groups, methodology, findings and recommendations. A stand-alone recommendation or sub-recommendation regarding gender equity is strongly advised.
- Devote more analysis to the situation of girls' education relative to the magnitude of those issues in the background (1.2) and findings sections (5), both before and after the pandemic, noting why disaggregated information is rarely presented, and if possible what impact the pandemic had.
- Bring in more frequently the evaluation components (general) as they contributed to the process, ensuring needed information is brought in from the TOR and the inception report and discussion of the role of the evaluation reference group.
- Change the title of the report to reflect its unique focus on education in the pandemic emergency response, noted in the executive summary as "UNICEF contribution to the education response to the COVID-19 pandemic".
- Strengthen the executive summary (page 5) by adding the audience, the management and the main guiding principles.
- Add a limitations and mitigation measures section to the methodology (4) revised from that in the inception report.
- Number the sections in the main report to correspond to the Table of Contents and contents and hyperlink the contents of both main report and appendices.

SECTION RATINGS

SECTION A: EXECUTIVE SUMMARY (weight 5%)		67%	Comments on Rating
Question 1.	Can the executive summary inform decision-making?		
i	Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Yes	The executive summary is clearly presented and minus the photos is approximately 5 pages. It can be used as a standalone document for decision-making. (Note: Please hyperlink the Table of Contents to the page numbers.)
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 summary includes most necessary elements including the background, overview, purpose and objectives, key findings and conclusions and key recommendations. It is missing the intended audience. Suggest that the intended audiences are included as stated under the Purpose of the evaluation, page 13, and noted as primary and secondary audiences.
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	Most information is provided but there is room to provide a few more paragraphs. The main guiding documents such as the UNICEF education strategy and UNICEF's guidance for education in emergencies plus the UNEG standards for evaluation are important to brief the readers. Optional would be a brief run down on the activities undertaken.
SECTION B: BACKGROUND (weight 5%)		79%	Comments on Rating
Question 2.	Is the object of the evaluation clearly described?		
i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Partially	A general description of the intervention appears in the Introduction (Section 1), page 11, which covers the UNICEF education response to the COVID-19 pandemic (2020-2021), with emphasis on the most disadvantaged children and young people. The investment of \$1.29 billion in 2020 includes one third for personal protective equipment (PPE) with a similar amount transferred to implementing partners, thus (assumed from this sentence) approximately \$2.58 billion; the 2021 budget is not mentioned. Suggest to clarify the timeline more precisely in terms of the two years being covered, the approximate times of school closures and reopening's, and to clarify the investment, whether the mentioned amounts were the total investment covering 2021 as well.
ii	Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cities, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Yes	The proportion of girls (147 million) of the estimated 301 million children (rights holders) reached and numbers of schools and preschools reached across 89 countries during lockdowns and school closures are noted in the Introduction. Information regarding duty bearers is found in the discussion of the evaluation audiences under the section on the Purpose of the Evaluation (2.1) and includes state and non-state actors. Information regarding children with disabilities is scattered among the findings and usually country specific.
Question 3.	Is the context of the intervention clearly described?		
i	Clear and relevant description of the context of the intervention (i.e. relevant policy, socio-economic, political, cultural, power/privilege, institutional, international factors) and how context relates to the implementation of the intervention.	Yes	The general global context is referred to in the Introduction including the impact of the pandemic on education which was the largest disruption of education in history. In this context, UNICEF partnered with numerous state and non-state actors. The objectives of the interventions, to protect against exposure to the virus, to minimize the impacts of public health measures (should be noted: on education), to address the socio-economic consequences and increase access to basic social services are presented. The section provides a brief overview of the activities supported.

	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Partially	Linkages are drawn to SDG 4 and some of the relevant strategies. The first paragraph under the Introduction mentions the <i>UNICEF Strategic Plan (2018-2021)</i> education programming for teaching and learning within Goal Area 2 with the outcome that children particularly the most marginalized and those affected by humanitarian situations are provided with inclusive and equitable education. The implementation of <i>Every Child Learns: UNICEF Education Strategy 2019-2030</i> began less than a year before the onset of the pandemic. It is important to stress the alignment of "UNICEF's COVID-19 response strategy in 2020 with the WHO's COVID-19 Strategic Preparedness and Response Plan and the UN's Global Humanitarian Response Plan led by the Office for the Coordination of Humanitarian Affairs (OCHA)". This sentence is found in the inception report and could be added here to highlight alignment with other global strategies in addition to general coordination with others that is mentioned in the evaluation report. The limited progress toward SDG 4 prior to the pandemic is briefly noted, however, it would be useful to briefly elaborate on reasons to further set the context and to tie into the evaluation findings. (Some reasons typically cited include shortcomings in quality and inclusiveness, inadequate infrastructure, low enrolment, gender disparities and digital divide.) SDG 3 is also relevant in terms of ensuring capacity for risk reduction and management of global health risks. Further, UNICEF's emergency response guidelines are applicable particularly for EIE, and use of INEE minimum standards.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Partially	The Introduction notes that the pandemic has affected nearly 1.6 billion learners in more than 190 countries and school closures affected 19 in 20 students. The last sentence in the Introduction section: "Due to the large scale disruptions ... the UNICEF education programme reached more children in 2020 than any year previously" requires clarification or more explanation as to why this is the case. Suggest to add more detail on the factors contributing to vulnerability of the most disadvantaged children and young people in the third paragraph of the Introduction as well as the most marginalized described in the third paragraph of the Evaluation Background to clarify why they require strong attention, and note the groups left behind who may include out of school children and those who are actively migrating or in conflict situations.
Question 4.		Are key stakeholders, their relationships and contributions clearly identified?		
	i	Identification of implementing agency(ies), development partners, right holders, and additional duty bearers and other stakeholders; and of linkages between them (e.g., stakeholder map) (if relevant).	Yes	Key stakeholders are mentioned in the introduction and in more detail in the Purpose and Objective section (2.1) and include government officials, children as rights holders, parents, guardians and communities, key implementing and development partners and donor agencies and institutions and strategic education partners, such as GPE, UNESCO, INEE and Education Cannot Wait.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	The roles of the key stakeholders are mentioned in Appendix A in reference to the Theory of Change including a detailed discussion of UNICEF's role in the global partnership and public goods during the COVID-19 emergency and cooperative roles with key actors including fund raising. UNICEF's role in technical guidance for Education in Emergencies (EiE) is mentioned along with key partnerships and coordinating bodies to support EIE. (Text should refer to Appendix A).
SECTION C:		EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)	67%	Comments on Rating
Question 5.		Is the purpose of the evaluation clearly described?		
	i	Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.	Partially	Section 2 covers the evaluation purpose, objectives and scope. The purpose is clearly set out as organizational learning and accountability. As to why it was needed at that point in time which the IR notes as being the two year anniversary since the end of the evaluation period is not mentioned. As to the intended use, it is to inform UNICEF strategic direction after the midpoint of the current global education strategy (2019-2030 - thus expected to be around 2025?). Summative and formative language can be brought in. The TOR refers to the evaluation as largely formative while the description in the report would suggest both summative and formative. The intended users are described as primary and secondary audiences. Noted that Table 1 in the inception report effectively covers the users and their interests which would enrich the discussion in the text or as an annex.
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 three objectives are those mentioned in the TOR and the objectives clearly state what is to be achieved.

	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	The scope of the evaluation is well set out in terms of time period, geographical coverage, thematic scope, and stakeholder coverage. This is the scope generally specified in the TOR and/or agreed in the inception phase, which can be mentioned.
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	The Evaluation background (1.2) refers to the <i>UNICEF Strategic Plan 2018-2021</i> Goal Area 2 Theory of Change (ToC) and the fact that the assumptions therein (e.g., stable economic, political and social environment) ceased to be true with the onset of COVID-19. The three key change strategies from Goal Area 2 (global partnerships and public goods; system strengthening; and service delivery) remained relevant. A clear explanation is needed as to how these change strategies were interwoven in the evaluation. It is noted that outside of the discussion in this section and the Purpose section (2.1), the ToC is not mentioned again in the report.
	ii Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Partially	The Theory of Change for Goal Area 2 which has been modified for the evaluation to be relevant to the COVID-19 response is found in graphic form in Appendix A, Figure A1 and the text describes UNICEF's activities related to the three key change strategies. The text should clearly explain that the ToC was modified for the evaluation as described in the inception report. The Evaluation background section describes the six outcomes to be evaluated and notes that the evaluation will examine to what extent UNICEF activities contributed at global, regional and country levels. The text does not explain that these are intermediate outcomes as found in the ToC and this should be noted with stronger reference to the ToC.
	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.	Partially	The Evaluation background (1.2) notes that a specific ToC was not developed for UNICEF's response. However in the inception phase (as per the inception report, page 3) the evaluators developed what is considered "a high level" ToC using outputs and activities from Goal Area 2. The graphic presented in Appendix A thus contains activities/outputs relevant to the pandemic response such as school reopenings, and distance learning platforms. The IR notes that the ToC will be under review in the data collection phases and a final version is presented in the evaluation report. The revision by the team should be clearly mentioned and confirming that the final version is presented in the report following the team review.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	86%	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	The evaluation effectively joins the evaluation objectives, themes and OECD-DAC criteria and does a good job in modifying the TOR questions as found in Appendix B. The evaluation questions are organized according to three Themes, which are stipulated in the TOR. Gender and girls are interwoven with the questions and girls are considered as part of the marginalized groups.
	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	Table C1 page 24 of Appendix C includes the Evaluation Framework which provides important information on which the team based its analysis. In addition to the questions, the framework includes judgement criteria, indicators, data sources and data collection methods.
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 robust for the purpose, objectives and scope. The approach is utilization focused and participative, and methods are described in Appendices B, C and D. The methods used are typical for capturing large data sets while using select cases for deeper analysis. Appendix B covers the approach, data collection methods and data analysis. Appendix D includes the primary data collection tools such as the key informant interview instruments, and the survey questionnaire.
	ii	Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	Data sources are described in Appendix B. The primary data sources are shown in Tables B 3-5 for interviews globally, regionally and locally. The online survey yielded some quantitative data. Table B6 indicates secondary sources which would also provide quantitative data. They are all well described.
	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Yes	Table B2 provides a summary of the sample across primary data sources. Between the global, regional and country based interviews in addition to the online survey for UNICEF Chiefs of Education, a diverse perspective was captured. Key informants other than UNICEF included partners such as funding partners (e.g., GPE) and education in emergencies partners (e.g., UNHCR) and other UN agencies. While regional interviews focused on UNICEF staff, the country level included government, NGOs and private sector in the seven case study countries. The selection of case study countries was rigorous as per the selection criteria depicted in Table B7, including main and additional criteria, and the matrix of countries selected and the criteria met shown on Table B8.
	iv	Clear and complete description of the methods of analysis.	Yes	The methods of analysis are described in Appendix B and included data coding using Dedoose, thematic analysis, iterative processes, and triangulation.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Partially	The diversity of data sources and methods of data collection and analysis allows causal connections to be defined. Reference to outputs and outcomes as per the Theory of Change or the results based language is generally missing. Suggest that some sentences are added to the main report and the annex methodology sections to indicate that elements of the ToC form part of the analytical perspective, particularly how progress toward outcomes will be assessed, and referring to the indicators on the framework matrix.
	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	Table B4 provides a summary of the planned and actual interviews at the regional level and Table B5 shows the country level. While most were achieved some respondents were not available. With regard to the seven case studies, the text notes that it could be difficult to generalize based on the small sample and numerous variables. Other forms of data supplemented the findings. A data and methods limitations section does not appear in the report while a very detailed matrix is found in section 5.4 of the inception report. This matrix should be updated with the limitations as actually experienced with the mitigation measures used by the team. It may appear as an annex or in the text.
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	Appendix B covers ethical considerations. Suggest to insert a few sentences or paragraph in the main text possibly under the approach to note that the evaluation was guided by UNEG evaluation principles and appropriate safeguards for participants have been employed with more details in Appendix B.
	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 considerations of confidentiality as well as behaviours of the team members and data collectors are mentioned in Appendix B. The interview guides start with a pledge of confidentiality and anonymity prior to the interviews as well as gaining permission to record the interviews. It can be mentioned that the rights holders were not interviewed and no children were involved in primary data collection.
Question 11.		Does the evaluation incorporate innovative practice that adds value to the evaluation process?		

	i Innovation practice is used to improve the quality of evaluation process. This could evident in several ways such as the design of the methodology (i.e. use of technology for data gathering, extensive participatory processes, systematic analysis processes such as collaborative outcomes reporting and incorporation of big data, specific strategies to address complexity such as outcome harvesting, strong child rights focus), or ways of sharing of evaluation results.	Yes	Value is added to the evaluation process using specifically designed selection criteria for the case study countries, using 'literal replication' as a method of case selection, meaning that the same type of results was predicted for each of the cases, despite the use of diverse interventions in differing contexts. The case study countries vary in terms of context, barriers, interventions, structures and geographical location. Variation along these dimensions provides the greatest coverage and the best chance of identifying patterns of similarity or difference in terms of the extent to which UNICEF strategies and interventions were effective and contributed to delivering results. The analysis focused on themes which allowed comparison across countries.
SECTION E:	EVALUATION FINDINGS (weight 25%)	67%	Comments on Rating
Question 12.	Do the findings clearly address all evaluation objectives and scope?		
	i Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Yes	The evaluation findings and conclusions are presented in Section 5. They are well organized by three themes as per the evaluation framework. First, key questions are presented along with the key conclusions followed by the numbered findings. Key findings are highlighted followed by discussion of the evidence. The evidence presented supports the key conclusions and findings statements.
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Partially	As the object Theory of Change pertains to UNICEF's long term education strategy and has been modified for the evaluation to be covid response specific, reference to the outputs, and outcomes, while possibly implicit, should be explicitly referred to as possible. Suggest to review the findings and more effectively draw in the ToC, the evaluation framework indicators, and the key humanitarian and development principles. Tie the discussion to the object planning tools, what was planned and aimed for, and what was actually accomplished. The addition of a few sentences will give additional clout to the findings and conclusions. In addition to enriching the discussion under each finding, a lead paragraph can be useful to remind of the evaluation use of the Theory of Change, what the evaluation has planned as per the framework as well as the global principles, human rights and education in emergencies, as well as humanitarian response, that underlie the strategy.
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.	Yes	The discussion of the findings effectively draws in evidence from both qualitative and quantitative data analysis. Triangulation is evident through multiple data sources and the text includes numerous supporting examples from the case study and other countries.
	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	The findings support both positive and negative conclusions. A greater drawing in of performance indicators, such as from the evaluation framework, standards and benchmarks, such as from the Theory of Change causal chain, would enrich the discussion. For example, under finding 1.1.3, the conclusion that the "response was not as fast and flexible as emergency procedures could be", the analysis could mention how well the response compared to UNICEF standards (or those of INEE, WHO or PAHO) or other emergency responses in regard to the speed and flexibility of its response to public health emergencies, and use of procedures as stipulated for Level 3 response.
	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	The findings pinpoint the causal factors as possible given that the vast number of variables among countries. A number of successes in inputs and outputs are mentioned along with organizations strengths at global, regional and country levels. The discussion of causal factors for less successful outputs are also touched upon. At the outcome level, the discussion could be strengthened. For example, remind readers that the intermediate and long term outcomes/results as per the ToC target the marginalized in particular and stress equity and inclusion and this is why the finding that UNICEF was not very successful in this regard during a widespread response is critical to future planning and why the assumptions may not be realistic.
Question 14.	Does the evaluation assess and use the intervention's Results Based Management elements?		
	i Assessment of the adequacy of the intervention's monitoring system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	Partially	At the global and regional level, evidence is presented of the effective use of the multiple indicator cluster surveys (MICS) data to plan the response and notes various activities of the UNICEF Division of Data, Analytics, Planning and Monitoring such as household surveys. The tracking of learning outcomes for children, in view of the strong collection of input and activity data, according to conclusion 5, little evidence was found regarding the effect of these successes on learning outcomes. Since learning outcomes form part of both intermediate and long term outcomes on the Theory of Change, and this finding eventually leads to a recommendation more clarification could be presented as to why the learning outcome data was not or could not be collected during the response period.

SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	50%	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	The conclusions are well formulated and strongly connected to the purpose, objectives, themes and questions. The forward looking language could be strengthened in some cases, noting what the formula is for success in future emergency responses. For example, the shortage of resources from HQ and regional stands out as a key factor in the beginning of recovery in Conclusion 9, however, it seems that country offices still did a good job although unclear how other resources would have helped. In view of recovery being a critical part of the emergency response, the falling out of higher level management resources can be given more clout as a limitation.
	ii Conclusions are derived appropriately from findings, and present a picture of the strengths and limitations of the intervention that adds insight and analysis beyond the findings.	Partially	The conclusions generally reflect a higher level of analysis rather than a summary. There is some room for improvement to lead in more strongly to the recommendations. It is noted that broad sweeping statements such as Conclusion 1: "UNICEF was a global leader..." seems to be something that was known prior to the evaluation and not a particular finding unless the aim was to validate this while connecting to whether that ranking helped the outcomes to be more effectively achieved. Suggest that a more valuable finding is the cooperation and coordination within and without used to create and share the resources which may be more forward looking and aim to be replicated.
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.	Not Rated	The TOR does not specify that a stand alone set of lessons will be prepared. The lessons and good practices gathered from data sources are incorporated in the findings discussions. The report notes that collection and documentation of lessons took place at some regional and country levels and by other stakeholders.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Not Rated	Lessons and good practices incorporated in the findings are not specifically phrased as lessons but the global nature of the evaluation suggests that they are widely applicable.
SECTION G:	RECOMMENDATIONS (weight 15%)	50%	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.	Partially	Some supporting conclusions are brought into the recommendations. To more firmly link the recommendations to the conclusions, the numbers of the relevant conclusions should appear so they can be traced back to the relevant findings and the relevant overarching evaluation questions. This will indicate the trail of logical thinking leading to the recommendations and to firmly ground them.
	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 useful and relevant to the intervention and capture the needs for reaching marginal groups and strengthening Education in Emergencies. They are currently too open ended and floating. Suggest to restructure the way the recommendations are presented to give them more clout, the justification first and then the recommendations, each followed by the operational considerations, or steps to be taken. For example, under Recommendation 1, the conclusion that information flow was top down seems not to correspond to the recommendation or note specifically how the recommendations will address this problem. For the mechanisms to support the review to consolidate the resources, will it be a global, regional or country based exercise and who will participate and even though the priority is high, within what time frame should this be accomplished? Strongly suggest that recommendation 2 devotes more text or a separate sub-recommendation to respect the SWAP guidance iiid. <i>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?</i>
	iii Process for developing the recommendations is described, and includes the Involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	Partially	The process for developing the recommendations is described in the TOR and the inception report and this information should form part of the introduction to the Recommendations section. The introduction should mention the objectives and purpose of the evaluation to set the stage and then note the process that management and the team used to develop the recommendations, as well as roles for duty bearers and rights holders, followed by an explanation for the way the recommendations are presented with relevant conclusions and suggested implementation measures. The team can refer to other evaluation reports for a format that works, this may include a matrix that organizes the components, and clarifies the priority, and who is responsible.
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	It is noted that the management response is always the responsible office/team and this may not be needed, as the introduction to the recommendation section should note the next steps in the process such as the reviews, dissemination of results and the management response. The programme group is the most important and this is noted. Consider presenting high priority recommendations first followed by the medium priority.
SECTION H:	REPORT STRUCTURE AND PRESENTATION (weight 5%)	86%	Comments on Rating
Question 19.	Does the evaluation report include all relevant information?		
	i Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes; list of acronyms/abbreviations, page numbers.	Partially	The opening pages include the evaluated object, the timeframe, the evaluating and commissioning organizations and the complete table of contents. Suggest that the title of the report alludes more specifically to the evaluation object which is the UNICEF Education response to the COVID-19 public health emergency or similar. Suggest that the word "object" is used in relation to the Introduction (section 1) description of the intervention as this is generally the practice in UNICEF reports and it is found as such in the inception report. The date of the report should appear on the front page in addition to the last page.
	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 annexes include the bibliography, the Terms of Reference, the evaluation framework matrix, details on the UNICEF education key change strategy and the reconstructed Theory of Change diagram, data collection instruments, and list of documentary evidence. Others are details on the methodology which includes the actual numbers of key informants and their locations, and the roles of the interview respondents, planned and reached at global, regional and case study country levels. The proportion of planned respondents that did respond to the staff survey are found in Figure B1.
Question 20.	Is the report logically structured?		
	i Structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles, well formatted).	Partially	The structure is well organized but the sections should be numbered as per the Table of Contents and need to be hyperlinked both in the main report and the annexes. The use of different colours in the conclusions and findings among the themes is assumed to differentiate them but this should be explained in the executive summary or else not used there.
	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 accords to the 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.	Yes	A page length is stipulated in the TOR for the inception report but not found for the evaluation report. The main report is about 50 pages leaving out the covers and the executive summary which is acceptable for a global evaluation.
	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 written in an accessible manner for the intended audiences, free from grammar, spelling and punctuation errors.
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Yes	The findings are well supported by visuals to present the results particularly of the survey. They are clearly labelled and referenced.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	50%	Comments on Rating
Question 22.	Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity?		
	i Reference and use of rights-based framework, and/or CRC, CCC, CEDAW and/or other rights related benchmarks in the design of the evaluation.	Partially	The TOR sets the stage with mention of rights based efforts (implied based on relevant human rights conventions) and include Education for All goals, child-friendly schools, among other types of interventions, research and development of rights-based standards working toward inclusion and equity. This discussion is not carried forth in the IR or evaluation report. The evaluation report can benefit from inclusion of rights-based instruments to highlight the rights based outcomes being aimed for and note UNICEF's efforts with others toward realizing the rights of the child, particularly in inclusion of the most marginalized prior to the pandemic, and augment rights based discussions and focus on the outcomes.
	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).	Partially	The methodology targeted a range of duty bearers as interviewees while the rights holders did not participate in primary data collection, however, data on them was available in secondary sources. Importantly an evaluation reference group was involved in the process which is described in the inception report (section 6.1.2) and should be mentioned. The TOR mentions the reference group under the Phase 4 activities and provides guidance for its formation and its role.

	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	The language avoids biases. The rights based language can be used more frequently to promote the right-based mind set in emergencies. The reasons that data is not disaggregated should be explained and form part of the limitations. The evaluation generally groups the marginalized including girls and the evidence on the individual groups is found under Question 02.5 and Conclusion 8. Noted that "UNICEF HQ and regional offices produced numerous reports, papers and guidance on inclusive remote learning and the return to school for children with disabilities. UNICEF country offices commonly sought to reach children excluded from distance learning, sharing such approaches with government." Examples from countries are provided in targeting ethnolinguistic minorities, remotely located children, and migrant learners. The examples generally provide numbers of each group that benefited from activities.
	iv Evaluation assesses the extent to which the implementation of the intervention addresses child rights and Leave No-one Behind (gender and other excluded and marginalized groups). It is disability inclusive.	Yes	The key goal of the ToC is realizing the rights of every child and child rights appear in the intermediate outcomes particularly (implied) right to education, right to inclusion and equity and protection from harm. The principle of Leave No-one behind is included in the theory of change in terms of inclusion and equity. Children with disabilities are included among the marginalized, as are girls. Since "no child left behind" is a rights based goal in education that UNICEF promotes, it can be mentioned and the rights based language implied throughout.
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)	3	
	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	The TOR sought gender and cultural balance among the evaluation team; Question 2.5 sought "evidence of the extent to which UNICEF was instrumental in the initiatives to identify and reach (provide for learning needs) children in particular the most marginalised groups, e.g. girls, children with disabilities, refugees, minority ethnic/language groups". Noted that the TOR, Theory of Change and Evaluation Framework offer scant mention of the need for focus on girl's education, including them among the marginalized, while UNICEF and other education in emergency proponents such as Education Cannot Wait put forth the principles of more strongly supporting girls and gender equality in the emergency response. Given the strategic focus and recognition that inclusion of girls required stronger focus in emergencies, the evaluation framework devoted minor attention to GEEW and possibly lost the opportunity to collect more robust data sets on girls' education in the COVID-19 response.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	UNICEF gender and equity staff were among the planned and actual interviewees. Data is not gender disaggregated except the total numbers of assisted. Methods and tools sought evidence of inclusion of girls among other marginalized but generally fell short of being considered strongly gender-responsive.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially integrated	The finding 2.5.2 notes: "UNICEF work focusing on the challenges and needs of girls mainly targeted the increased risk of violence against girls during the school closures and their need for psychosocial support. In early 2020, at the onset of COVID-19 outbreaks in the region, the East Asia and Pacific Regional Office collaborated with the Inter-Agency Standing Committee's Gender in Humanitarian Action Working Group to develop a joint advocacy brief calling for attention to gender inequalities in the emerging pandemic. This informed the development of the UNICEF global technical note 'Five Actions for Gender Equality in the COVID-19 Response'. " The conclusions and recommendations include girls in the discussions of marginalized but do not focus on them or differentiate them by age. Thus an important gender analysis is mainly lacking recognizing that girls are disadvantaged in education generally including those not marginalized in other ways and those experiencing society pressures to drop out of education such as adolescent girls.

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.**
- 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?
 - Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?
 - Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?
 - 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 GEEW 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 GEEW issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area?