GEROS Evaluation Quality Assurance Tool

Version: September 2021

Joint Evaluation of Improving Access to Water, Sanitation, And Hygiene (WASH) in Rural Fishing Communities Programme in Sierra Leone (2019-2022)

REPORT RATING SUMMARY

Overall Rating		73%	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	Joint Evaluation of Improving Access to Water, Sanitation, And Hygiene (WASH) in Rural Fishing Communities Programme in Sierra Leone (2019-2022)
Report sequence number	Sierra Leone/61/2024/19196
Region	WCAR
Year of report	2024
Office	UNICEF Sierra Leone
Coverage (countries)	Sierra Leone
ToRs present	Yes
Date of review (dd/mmm/yyyy)	October 2, 2024
Name of review firm	IOD PARC

CLASSIFICATION OF EVALUATION REPORT

Management of evaluation *(Managerial control and oversight of evaluation decisions)*	Not clear from report
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	No
Every child lives in a safe and clean environment	Yes
Every child has an equitable chance in life	No
Gender equality (cross-cutting)	Yes
Humanitarian action (cross-cutting)	Yes
Evaluation object	Programme
Evaluation type	Summative
Evaluation strategy	Mixed methods
Evaluation design *(primary method used)*	Participatory
Evaluation level	Output & Outcome
Geographic scope	Other
Primary SDG(s) covered *(number)*	1, 2, 3, 4, 5, 6, 11, 14, 16

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

Strengths:
--The report is logically structured and written in a manner that is accessible to a wide audience.
--The evaluation includes a detailed context section that details the national context and highlights specific needs of certain populations, including persons with disabilities.
--The evaluation report clearly states the evaluation purpose, scope, objectives, ‘why now’, intended use and intended users.
--The findings section is organised by evaluation question.
--The report offers valuable lessons learned and actionable recommendations.

Weaknesses:
--The executive summary does not contain sufficient information to be a standalone document useful for decision making.
--The report's length (over 100 pages exclusive of executive summary and annexes) is not aligned with the ToR which stated: ‘max 60 pages with the rest to be placed in annexes and an executive summary of a maximum of 5 pages’. The length of the report may impede uptake of information and reduce accessibility to relevant audiences.
--The conclusions section provides summaries of findings and does not analyse and interpret the significance of those findings (i.e., does not add insight and analysis beyond the findings which is required of conclusions).

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

The report would benefit from a more detailed executive summary. As written, the executive summary does not contain sufficient detail to be a standalone document used to inform decision making.

The report would benefit from some revision to cut down on its length. It is currently over 110 pages, exclusive of the executive summary and annexes. This is likely a barrier to uptake/consumption of the evaluation output and is not aligned with the ToR which stated a 60-page limit with all other content annexed.

The report would benefit from a thorough review of the findings section to ensure that sources of evidence are either noted in the body of the text or provided as footnotes. As written, it is unclear what the underlying data is for many findings.

The conclusions section should include interpretation of the significance of evaluation findings, adding insights and analysis beyond the findings themselves. As written, the conclusions section is simply summaries of findings.

The recommendations section should have used the same scale for all recommendations. As written, the timeline column uses two scales (e.g., uses ‘medium’, ‘high’, and ‘short’) and the report should have been revised to ensure that the same scale was used for each recommendation (either priority level – high, medium, low – or timeline – short-term, medium-term, long-term).

SECTION RATINGS

SECTION A:	**EXECUTIVE SUMMARY (weight 5%)**	50%	Comments on Rating
Question 1.	Can the executive summary inform decision-making?		

i	Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Partially	The executive summary is five pages in length. It is clearly structured and provides a good summary of the detailed findings as well as the purpose, scope and objectives of the evaluation. The executive summary provides a brief overview of the project itself in the introduction but lacks details on the project activities and the national context which is required to understand the project being evaluated. The lessons learned and recommendations do not provide enough detail. In order for the document to serve as a standalone document useful for informing decision making, the executive summary should include: 1) a short description of the national context as it relates to WASH, 2) details on the project's activities (1-3 sentences/bulleted list), 3) more detailed lessons learned, and 4) more detailed recommendations.
ii	Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions on findings, lessons learned if requested, key recommendations) as per the ToR.	Partially	The executive summary includes a brief overview of the intervention but lacks crucial details including key project activities. The executive summary lists the larger purpose of the evaluation (accountability and learning) but does not note the objectives of the evaluation or the intended audience. The executive summary does not present key conclusions (note that the conclusions section of the report does not contain conclusions. Rather, the section contains summaries of the findings per evaluation criterion). The executive summary requires more details regarding the listed lessons learned and recommendations, as noted above. The executive summary contains sufficient information on the evaluation methodology and findings.
iii	Includes all significant information needed to understand the intervention and the evaluation AND does not introduce new information from what is presented in the rest of the report.	Partially	The executive summary does not include all significant information needed to understand the intervention (i.e., it is lacking project activities, as noted above) and the evaluation (i.e., it is lacking details on the objectives, lessons learned, and recommendations, as noted above). The executive summary does not introduce new information from what is presented in the rest of the report.
SECTION B:	**BACKGROUND (weight 5%)**	100%	Comments on Rating
Question 2.	Is the object of the evaluation clearly described?		
i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Yes	Table 3 on pages 7-8 fully meets this criterion. It includes location (Sierra Leone), budget, duration, overall objective of the project, expected outcomes and outputs.
ii	Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals…), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cites, sub-regions…) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Yes	Table 3 describes the intended rightsholders and the partners (duty bearers). The NGO partners are disaggregated by geographic location. The government partners are for all areas of the project. Annex 8 provides information on numbers reached for Goderich and Konacrydee together and then for Tombo separately. This does not include disaggregation by gender, age, disability or other populations. However, this is not the fault of the evaluation team as they have provided the data made available to them. As such, the criterion is considered to be fully met.
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 report provides a detailed description of the context including socio-economic and policy-related/institutional environments, and it is clear how these contexts relate to the implementation of the intervention. Given the length of the report (over 100 pages), the context section could have been distilled into 2 pages to shorten the overall report length. The context section is currently 6.5 pages.
ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	The final report clearly states the relevant SDGs and their interventions/targets within the context section. These include SDG 1-6 and 16.
iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	The context section of the final report notes findings from several studies including the Sierra Leone integrated household survey (2018), WASH NORM (National Outcome Routing Mapping) 2022, and the WHO/UNICEF Joint Monitoring Programme 2020. This helps to paint a picture of WASH needs in the country. The report links WASH to issues of malnutrition and stunting, but does not provide data on the current statistics. While this was a missed opportunity, the report provides sufficient details to frame the object of the evaluation.
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	Table 3 includes the implementing agencies and partners (duty bearers) as well as rightsholders. Section 2.3 provides an overview of the roles and responsibilities of duty bearers. Further, findings from evaluation question coherence 2.2 provides additional details on how duty bearers worked together in the execution of the projects.

ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	The report is clear that UNICEF and the Government of Iceland provided funding (Table 3). Per Section 2.3, technical oversight of the project was provided by the MoFMR with support from the MoHS and MoWR. At the subnational level, UNICEF and its implementing partners worked with district representations of the MoFMR, MoWR and MoHS as well as with the Western Area Rural and Port-Loko district councils, and district education offices from the MBSSE to carry out the project.
SECTION C:	**EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)**	100%	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.	Yes	Sections 3.1 and 3.2 clearly convey this information, including why it is needed at this point in time (i.e., to inform the next phase). Table 4 includes a detailed description of the users and uses of the evaluation findings.
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	Section 4 clearly states the objectives per the ToR. There are no changes from what was requested in the ToR.
ii	Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Yes	Section 5 clearly states what will be covered thematically (project's social and constructional components), chronologically (8 Feb 2019-31 Dec 2021 in Tombo and 24 Jan 2020-October 2023 in Goderich and Konacrydee) and geographically (all project activities in Tombo, Goderich and Konacrydee through a representative sample from the project sites).
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.	Yes	As part of the inception process, the evaluation team constructed a ToC based on the results frameworks for each of the two projects (1) Tombo and 2) Goderich and Konacrydee). This is presented in Figure 4 and narrative form as output and outcome level indicators in section 2.2.
ii	Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Yes	The ToC created by the evaluation in Figure 4 presents the causal relationship between outputs and outcomes in graphic form.
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 projects did not have a ToC. The evaluation team assessed the results frameworks for Tombo and Goderich and Konacrydee and created a ToC.
SECTION D:	**EVALUATION DESIGN AND METHODOLOGY (weight 20%)**	77%	Comments on Rating
Question 8.	Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? *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	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	Taken as a whole, the evaluation questions and sub questions are appropriate for meeting the objectives and purpose of the evaluation. However, some questions would have benefited from edits for clarity and others could have been removed or transferred to more appropriate criteria. The evaluation matrix and findings section align the evaluation questions to their appropriate evaluation criteria (i.e., relevance, coherence, effectiveness, efficiency, sustainability, and gender equality, human rights, equity and the environment). However, there are some questions that do not fully fit within the OECD DAC criteria. For example, under efficiency, it is unclear how questions such as 'EFFI 1.3 What were the strengths and weaknesses of the programme's supplies and delivery to the communities?' and 'EFFI 1.4 What monitoring and other evidence generation activities did the programme entail to ensure results-based management?' relate to the OECD DAC efficiency criterion. Similarly, a question under sustainability does not fully fit the criterion: 'S1.3 To what extent the capacities of community structures are strengthened to effectively manage, operate and maintain installed WASH facilities and to create demand for sanitation through Community Led Total Sanitation (CLTS)?'

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	The evaluation matrix (Annex 4) provides the main evaluation questions, sub questions, indicators, data collection methods/tools, data sources and data analysis techniques.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?		
i	Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	The evaluation utilised mixed-methods approach including secondary data and literature review, household survey, school survey, KIIs, FGDs and O&M checklist/observations. These methods were adequately robust for the evaluation's purpose, objectives and scope and are fully and clearly described (including sampling techniques scored below).
ii	Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	Data sources are appropriate and include both qualitative and quantitative sources. All data sources (secondary data and literature review, household survey, school survey, KIIs, FGDs and O&M checklist/observations) are described in section 7.2 (but note that the subsections within 7.2 are labeled 7.3 in error).
iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Partially	The multi-stage sampling strategy employed by the evaluation team is described in section 7.3. These included purposive, convenience and random + quota sampling techniques. Considerations of representativeness of sample as well as gender balance and other equity issues (e.g., inclusion of persons with disabilities) were also noted. Reasons for not meeting gender balance were noted. The report would have been strengthened by including language on whether or how children with disabilities were considered in the school survey (e.g., tailored tools), in addition to the sample being balanced across target grades and by gender.
iv	Clear and complete description of the methods of analysis.	Yes	The report details the data analysis methods in section 7.5.1. These included quantitative analysis (descriptive statistics, crosstabs) as well as qualitative analysis (narrative and content analysis).
v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The methodology used for the evaluation allows the team to draw connections between outputs and outcomes. The analysis included triangulation of multiple sources of information per evaluation question (see Annex 4).
vi	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Yes	Section 7.7 details the limitations/constraints faced by the evaluation as well as mitigation measures taken by the evaluation team.
Question 10.	Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
i	Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Yes	Section 7.6 and annex 14 (as well as information presented in the inception report and referenced in section 7.6 of the final report) references the UNEG Ethical Guidelines for Evaluations. The authors confirmed that the following ethical guidelines/principles were followed: utility and usefulness, independence, impartiality, credibility, conflicts of interest, honesty and integrity, respect for dignity and diversity, human rights, equity and equality, confidentiality, avoidance of harm, omissions and wrong-doing, beneficence, and informed consent.
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.	Partially	The evaluation received formal IRB approval prior to data collection. The ethical principles/guidelines are detailed in Section 7.6 and include issues of confidentiality/informed consent, avoidance of harm, etc. Further, the informed consent process is available in Annex 14. One item that is missing is a form of assent for children under 18 which should have been included. The inception report states that all enumerators were to undergo child protection training and sign a statement of commitment to Montrose's safeguarding policy as confirmation that they will abide to while in the field. The ethical considerations section (Section 7.6) refers back to the inception report. The report would have been strengthened if it included a summary of the ethical considerations actually carried out (as opposed to planned at inception).
Question 11.	Does the evaluation incorporate innovative practice that adds value to the evaluation process?		

i	Innovation practice is used to improve the quality of evaluation process. This could evident in several ways such as the design of the methodology (i.e. use of technology for data gathering, extensive participatory processes, systematic analysis processes such as collaborative outcomes reporting and incorporation of big data, specific strategies to address complexity such as outcome harvesting, strong child rights focus), or ways of sharing of evaluation results.	No	The evaluation utilised a mixed-methods approach and gathered much of the data via tablets, but there was no evidence of an innovative method used during the evaluation.
SECTION E:	**EVALUATION FINDINGS (weight 25%)**	75%	Comments on Rating
Question 12.	Do the findings clearly address all evaluation objectives and scope?		
i	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Partially	The evaluation findings section is divided by evaluation criteria, then evaluation question and then evaluation sub-question. There are findings presented for each question and, overall, the findings marshal sufficient levels of evidence. However, there are some instances where the evaluation authors detail when the evaluation team received conflicting information and were forthright when clear assessments could not be made with available data. Additionally, there are instances (as noted below) where the report presents information without citing specific data sources (e.g., documentation, quantitative data, FGDs, KIIs), making it difficult to assess whether enough evidence is provided to answer the question accurately. A thorough review of the report to ensure that all findings are clearly linked to their data sources, either within the text or in footnotes, would enhance the reader's ability to understand the evidence behind each finding and judge whether there was enough evidence to answer each question.
ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The evaluation findings explicitly reference the reconstructed ToC and accompanying results framework. This is especially clear under the effectiveness criterion where the authors divide findings per output and outcome, but is also present throughout the findings.
Question 13.	Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		
i	Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.	Partially	The evaluation utilised credible forms of qualitative and quantitative data from diverse sources/stakeholders (see Annex 9). The report notes both output and outcome level information. The findings section would benefit from stronger triangulation, as this was not consistently evident throughout. (Although it is important to note that many findings were clearly triangulated, using language such as 'As reported by UNICEF in its activity reports and confirmed during the KIIs, the project faced...'). That said, the report also presents many statements without supporting evidence, making it difficult for the reader to identify the source of the findings. Additionally, without more context, the reader cannot determine whether these findings are based on a single data point or multiple sources (i.e., whether triangulation was used). There are numerous instances of this issue throughout the section. For example, the following statement illustrates this concern: 'The MoWR wished to have been more involved in designing the project and selecting the project locations. The project was implemented mainly at the community level, with few activities at the district level even though the district level had been involved and informed during the planning stages' (pg. 38). The report would have benefited from a thorough review of findings to ensure that data sources are either cited within the body of the report or as footnotes.
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	Where relevant, the evaluation report discusses progress towards indicators with full details provided as annexes. The evaluation highlights both positive and negative findings. As noted above, there are many instances where findings are clearly supported by, and respond to, the evidence presented. Other times, the report includes findings that do not clearly link to specific evidence sources (e.g., reference to desk review documentation, quantitative data, KIIs/FGDs, etc.). The report would have benefited from a thorough review to ensure that evidence for each finding was clearly stated either within the body of the text or within footnotes.

iii	Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyse the logical chain (progression -or not- from implementation to results).	Yes	The findings of the evaluation do a good job at noting the causal factors that either led to achievement or non-achievement of specific results, where data was available. For example, under Effectiveness, the report clearly lays out the factors that contributed to project successes and hindrances (e.g., internal factors, external factors, social factors, economic factors, political factors, COVID-19). In efficiency, the report investigates how the human resource capacity and financial capacity-related causal factors impacted achievement or non-achievement of results. Throughout other evaluation criteria, causal factors are also referenced and are used to support the development of lessons learned, conclusions and recommendations.
Question 14.	Does the evaluation assess and use the intervention's Results Based Management elements?		
i	Assessment of the adequacy of the intervention's monitoring system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	Yes	The evaluation includes an assessment of the adequacy of the results frameworks from the project including reconstructing a ToC and including a revised results framework to better reflect vertical and horizontal logic. The evaluation also provides some analysis and reflections on the monitoring efforts carried out under the proramme (e.g., spot checks, regular project monitoring visits and joint monitoring visits). The evaluation team highlighted that internal factors that contributed to project successes were related to monitoring and information sharing, among other aspects. The evaluation team concluded that the various monitoring tools used by the project helped to ensure results-based management and monitoring, and helped UNICEF detect bottlenecks during project implementation. In answering the efficiency sub-question 1.4, the evaluation stated that the team was 'given access to the database and was able to verify the functionality of the monitoring tools using with the data collected.' The evaluation also assesses the adequacy of the intervention's financial monitoring system, noting several shortfalls in section 8.4.1. Shortfalls of other aspects of the monitoring system include that the information management system for monitoring the WASH activities was built while the project was running. As a result, the evaluation team noted that baseline data for Tombo may be incomplete. Further, the evaluation noted that the lack of readily available assessment reports as highlights some issues related to the knowledge management of the project (see relevance section). Annexes 7 and 8 display the expected outputs and planned targets alongside performance at the time of the evaluation per location. Data was available for all outputs/targets which supported the evaluation's findings that the monitoring systems in place were adequate, although recommended improvements were provided by the evaluation team.
SECTION F:	**EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)**	63%	Comments on Rating
Question 15.	Do the conclusions clearly present an objective overall assessment of the intervention?		
i	Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking (if a formative evaluation or if the implementation is expected to continue or have additional phase).	No	The conclusions section does not provide conclusions but instead provides summaries of the findings per evaluation criterion. As such, it does not fulfil this criterion. Detailed feedback on the text presented in the conclusions section follows under questions 15.ii.
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	*Note that the conclusions section does not present conclusions. Rather, the conclusions section presents summaries of the findings per evaluation criterion. Since the section presents a picture of the strengths and limitations of the intervention, but does not add insight and analysis beyond the findings, the reviewer has marked this criterion as 'partially' addressed.* The conclusions section presents specific evidence from the evaluation and focuses on factual information and evidence gathered during the evaluation without making broad or speculative statements beyond what the findings demonstrate (i.e., conclusions). Rather than just listing brief summaries of the findings under different headings (Relevance, Coherence, Effectiveness, etc.), the authors should weave them into a cohesive narrative that highlights how they relate to one another and contribute to the overall assessment of the project's success. Instead of simply summarising the findings, the authors should have analysed and interpreted the significance of those findings, going beyond describing the issues and start interpreting the significance of the findings.
Question 16.	Are logical and informative lessons learned identified? [N/A if lessons are not presented and not requested in ToR]		
i	Identified lessons stem logically from the findings, have wider applicability and relevance beyond the object of the evaluation.	Yes	The four lessons learned stem logically from the findings and refer back to key information from the evaluation. These lessons learned have wider applicability and relevance beyond the object of the evaluation and would be useful to other similar WASH-related programming.
ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	The lessons are concisely presented (one paragraph each) yet have sufficient detail to be useful for the audiences of the evaluation and beyond. The lessons refer back to examples from the findings which readers could go back to and investigate further if interested. The lessons are also supported by additional details via the recommendations.
SECTION G:	**RECOMMENDATIONS (weight 15%)**	63%	Comments on Rating

Question 17.	Are recommendations well grounded in the evaluation?		
i	Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Yes	The recommendations align with the evaluation's purpose to help the GoI, UNICEF Sierra Leone and its partner to guide the design and implementation of the next phase. All of the recommendations are clearly formulated and logically stem from the findings, but there are some issues highlighted in the findings related to environmental elements of the project structures that seemed to warrant a recommendation. However, this was not included in the recommendations section: --Findings related to environmental elements of the structures: 1) under Efficiency questions 3.1 and 3.2: 'Similarly, the roofs of some fish platforms in Tombo had to be replaced because they were not designed to withstand high winds. This should have been considered at the design stage, as it would have been less expensive to build the roofs adequately than to refit them retrospectively. However, UNICEF quickly learned from this situation and decided to change the roof design for the fish platforms in Konacrydee and Goderich before construction began.' 2) finding under Sustainability 2.2. and 2.3: 'The schools in Goderich are vulnerable to the salty winds from the sea, which have quickly damaged the roofs and locks of the school latrines. The SMC requested that the roofs of the latrines should be built in concrete rather zinc.')
ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	Overall, the recommendations are useful and address critical issues such as sustainability, inclusivity (for people with disabilities), gender equity, and governance, which are essential for the success of future programming. Recommendations are actionable and provide a clear direction for UNICEF WASH and the GoSL to follow, particularly the recommendations on addressing immediate operational challenges. That said, the recommendations could have been made even stronger through the inclusion of detailed implementation steps for each recommendation. For example, as worded, strategic recommendation 5 is not actionable for the primary intended users without additional information. If this recommendation is specifically for UNICEF-supported WASH projects, then it may be actionable for UNICEF. If this is meant for all WASH projects, then it is less appropriate and actionable. Guidance is also given for implementation regarding timeline for completion, but the recommended timelines do not always logically align with the recommendation itself. For example, strategic recommendation #2 states 'Incorporate realistic project assumptions into the logical framework and ToC such as internal and external factors identified in this report as well as the sustainability constraints found during the evaluation.' The timeline associated with that recommendation is 'medium'. It is unclear why this is recommended for the medium term when realistic project assumptions should be incorporated into the new/revised ToC at the outset of any new iteration of the project. In other words, it is unclear why this recommendation is not marked as 'short term' or 'high priority'. Further, the timeline uses two scales (e.g., uses 'medium', 'high', and 'short') and the report should have been revised to ensure that the same scale was used for each recommendation (either priority level – high, medium, low – or timeline – short-term, medium-term, long-term).
iii	Process for developing the recommendations is described, and includes the Involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	Partially	The report describes the process for developing the recommendations. The evaluation reference group (ERG), co-chaired by UNICEF, the Government of Sierra Leone and the Government of Iceland, reviewed and gave feedback on the recommendations. The evaluation team then refined the recommendations based on the feedback received from the ERG members. The development of recommendations did not include rights holders. It is unclear whether this was because it was not feasible. The report would have been strengthened if it included an explanation for why rights holders were not involved.
Question 18.	Are recommendations clearly presented?		
i	Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Partially	The report clearly identifies groups or duty-bearers responsible for each recommendation. The report also prioritises each recommendation, but the system used is unclear. The title of the prioitisation column is 'timeline' but the text used for each recommendation ranges from short, medium to high. It is unclear what high means (i.e., did the authors mean to recommend that it be carried out quickly/in the short-term). The authors should have ensured that the same prioritisation scale was used for each recommendation. Further, the evaluation report does not provide general definitions for what is meant by short, medium and high.
SECTION H:	**REPORT STRUCTURE AND PRESENTATION (weight 5%)**	71%	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 name of the evaluated object, date of the report, location of the evaluated object, names of the evaluators, table of contents and list of acronyms/abbreviations. The opening pages are missing the following information which should have been presented: timeframe of the evaluation and name of organization commissioning the evaluation.

ii	Annexes include: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework (unless included in report body), list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology, information about the evaluator(s), etc.	Yes	The report includes all of the required annexes noted in this criterion. In addition, the final report provides the following annexes: Annex 2 – Risk register, Annex 8 – Final summary of key results against final targets , Annex 10 – CLTS raw data –UNICEF M&E system, Annex 11 – JMP standards WASH Assessment, Annex 13 – Quantitative tables for HH and school surveys, Annex 14 – Consent Form for the KII and FGD, Annex 15 – Reconstituted multi stakeholder engagement, Annex 16 – Additional needs expressed by the respondents, Annex 18 – Authorization letter for field data collection, and Annex 19 – Recommendations from the ERG meeting.
Question 20.	Is the report logically structured?		
i	Structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles, well formatted).	Yes	The report uses a clear structure to organise information. Sections and subsections are titled clearly and are numbered for easy reference. The table of contents contains hyperlinks for easy navigation. The document is well formatted with varying font/text size for sections, subsections and table/figure titles.
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 of the report accords to UNICEF guidelines for evaluation reports. The context, purpose and methodology precede the evaluation findings which are followed by lessons learned, conclusions and recommendations.
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.	No	The report does not fulfil this criterion. The report is over 110 pages. This exceeds the limit of 60 pages in the ToR. The length of the report may present a major barrier to uptake and accessibility. Many sections of the report would have benefited from thorough editing for succinct/concise background information and findings. Additional, more detailed information could have been provided as annexes. This would have aligned with the requirements set out in the ToR: 'Final Evaluation report (max 60 pages with the rest to be placed in annexes and an executive summary of a maximum of 5 pages)'.
ii	Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Yes	Overall, the report is clear and accessible to its intended audience, with minimal grammar, spelling, and punctuation errors. While a final copyedit or proofread could have addressed these minor issues, they do not affect the reader's ability to understand the content.
iii	Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labeled, and referenced in text.	Partially	The report makes frequent use of visual aids to help convey key information. These visuals are clearly presented and labelled. Regarding references in the text, the report would have benefited from including consistent reference to the specific table/figure numbers for clarity and easy reference. As written, many tables/figures are not referenced specifically, but the data contained therein is discussed in the body of the report, with some exceptions. Figures that require reference/explanation in the text or require additional information in the text include: Table 15: Access to basic water supply by district Graph 1: Used time to fetch water before and after the project Table 18: Water facilities visited (O&M audit) Table 19: Proportion of households that considered water was unsafe (before and after the project) Figure 9. JMP standards in school Table 25: Access to basic sanitation facilities Graph 10: Percentage of HHs that pay for water per quantile and location Table 34: Summary findings on water points (water facilities) from O&M audit Table 35: Strengths of sanitation facilities against risks Table 36: Summary findings O&M audit Table 37: Summary findings of the O&M audit of latrines (sanitation facilities) Figure 18 Disability inclusion in Sanitation facilities at school Additionally, Figure 5. Map of project locations in Sierra Leone on page 11 and Table 11: Limitations and mitigation measures on page 23 would have benefited from an introductory sentence or two. There were two references to visuals are mistitled or the reference source is not found. These include:
SECTION I:	**EVALUATION PRINCIPLES (weight 10%)**	74%	Comments on Rating
Question 22.	Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity?		
i	Reference and use of rights-based framework, and/or CRC, CCC, CEDAW and/or other rights related benchmarks in the design of the evaluation.	Yes	The evaluation report notes that the use of a rights-based approach was employed in the design of the evaluation. This was done through adhering to UNEG and UNICEF ethical standards, ensuring representation in data collection, and exploring questions around inclusion and equity within the evaluation framework. The report references the SDGs and the CRPD.

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 report clearly indicates the level of participation of key rights holders and duty bearers as participants in data collection (see Section 7 and Annex 9). The evaluation report notes the presence of an evaluation reference group (ERG), co-chaired by UNICEF, the Government of Sierra Leone and the Government of Iceland, but it does not specify all stakeholders who were a part of the ERG. Further, the report does not specify the role/responsibilities of the ERG beyond stating that the ERG reviewed and gave feedback on the recommendations. More detail regarding the members of the ERG and its role/responsibility should have been included.
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 report often uses language that is empowering and inclusive (e.g., rights holders, persons with disabilities instead of disabled persons). However, the report also uses the term beneficiary/ies in some areas, including within the reconstructed ToC. The authors should have ensured that all references to project participants used a more rights-based term other than beneficiary, such as 'participants' or 'rightsholders' to reflect the empowerment of individuals and communities in claiming their rights. The report indicates some challenges encountered due to a lack of disaggregated data which prevents some differential results from being assessed. The data collected by the team includes disaggregation by gender as well as disability (either households with persons with disabilities or interviews with persons with disabilities). O&M audits also included accessibility components for persons with disabilities. The evaluation places a strong emphasis on inclusion and equity. Aside from the occasional use of the term beneficiary/ies, the report makes a clear effort to amplify the voices of marginalized groups, uses inclusive and empowering language that avoids gender, heteronormative, age, cultural, and religious biases, and provides disaggregated data whenever possible. Based on these efforts, the reviewer concludes that the report fully meets this criterion.
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	To the extent that data allowed, the evaluation team assessed the extent to which the implementation of the projects addressed child rights and Leave No One Behind. The evaluation contained some disability-inclusive aspects including ensuring that persons with disabilities as representatives from organisations of persons with disabilities were included as key informants. The report does not note whether the team made an effort to included children with disabilities within the samples from schools. Information around efforts to include children with disabilities in data collection should have been noted.
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)	8	
i	GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Fully integrated	**SWAP Rating Guidance:** GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected. a. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results? **YES.** The evaluation reflects on whether sufficient evidence was collected during the project on indicators related to human rights and gender equality. b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives? **YES.** The evaluation mainstreamed human rights and gender equality throughout its four objectives. c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? **YES.** The evaluation included a criterion on gender, human rights, equity and the environment. d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? **YES.** The evaluation includes questions and sub-questions related to GEEW.

ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Fully integrated	**SWAP Rating Guidance:** 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? **YES.** The evaluation included details regarding how gender issues were addressed in the methodology and disaggregated collected data by sex. b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations? **YES.** The evaluation employed a mixed-methods approach. c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? **YES.** The evaluation included a diverse range of data sources including persons with disabilities, women, and children. d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? **YES.** The evaluation methods and sampling frame sought to capture the diversity of stakeholders including women, children and persons with disabilities. e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? **YES.** The evaluation followed UNEG and UNICEF ethical guidelines and standards and received formal IRB approval prior to data collection.
iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Satisfactorily integrated	**SWAP Rating Guidance:** 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? **YES.** The Context (background) section of the report explores the needs of specific social groups (e.g., persons with disabilities) and includes reference to the SDGs and CRPD. b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable? **YES.** The findings include diverse voices, including key informants from organisations of persons with disabilities who are persons with disabilities themselves. c. Are unanticipated effects of the intervention on human rights and gender equality described? **PARTIALLY.** Some of the unintended positive outcomes described on page 63 include issues related to human rights. There were no unintended positive outcomes related specifically to gender equality. d. Does the evaluation report provide specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area? **YES.** Strategic recommendation #4 addresses GEWE issues specifically.

SWAP Rating Guidance

i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.
a. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results?
b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?
c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?
d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation?

ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.
a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex?
b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?
c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?
d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?
e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?

iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.
a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality?
b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable?
c. Are unanticipated effects of the intervention on human rights and gender equality described?
d. Does the evaluation report provide specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area?