

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Evaluation of WASH in Schools Programme in the Karamoja region

REPORT RATING SUMMARY			
Overall Rating	70%	Satisfactory	
●●●●●	Exceptional (96% - 100%)	5	
●●●●●	Highly Satisfactory (87.5% - 95.99%)	4	
●●●●●	Satisfactory (62.5% - 87.49%)	3	Meets UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with confidence
●●●●●	Fair (35% - 62.49%)	2	
●●●●●	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
Title of the evaluation report	Evaluation of WASH in Schools Programme in the Karamoja region		
Report sequence number	Uganda/63/2024/19556		
Region	ESAR		
Year of report	2024		
Office	Uganda Office		
Coverage (countries)	Uganda		
ToRs present	Yes		
Date of review (dd/mmm/yyyy)	January 17, 2025		
Name of review firm	IOD PARC		
CLASSIFICATION OF EVALUATION REPORT			
Management of evaluation (Managerial control and oversight of evaluation)	UNICEF managed		
Unicef goal areas (Alignment with strategic plan priorities)			
Every child survives and thrives	No		
Every child learns	Yes		
Every child is protected from violence and exploitation	No		
Every child lives in a safe and clean environment	Yes		
Every child has an equitable chance in life	No		
Gender equality (cross-cutting)	Yes		
Humanitarian action (cross-cutting)	No		
Evaluation object	Programme		
Evaluation type	Summative		
Evaluation strategy	Mixed methods		
Evaluation design (primary method used)	Theory-based		
Evaluation level	Output & Outcome		
Geographic scope	National		
Primary SDG(s) covered (number)	4, 6		
EQA Summary: <i>The rater will provide top line issues for this evaluation relevant for feedback to senior management (positive and negative), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight best practice/added value elements and the level of complexity of the evaluation.</i>			
Overall, this is a satisfactory report relating to the UNICEF-KOICA project 'Improved access to WASH to enhance adolescent-friendly school environments programme in Karamoja'. The topline issues are presented below.			
<ul style="list-style-type: none"> •The executive summary offers a coherent standalone document, that is helpful to decision making. Although it would have been helpful if, as in the main report, the methodology section said that there was a comparison of the before and after intervention – as this is an important methodological approach. Also, as in the main report, the conclusion section is not a conclusion, reflecting on the purpose of the evaluation, but instead it offers a summary of the findings. •The intervention is suitably described, presenting a graphical and narrative description of the intervention, along with a strong description of the rights holders and the social, political, environmental context affecting the rights holders and intervention. A clearly labelled section stating that a section describes the rights holders, and, in general, presenting a more coherent narrative would have improved the readability of the intervention and context. The findings, assessing the theory of change, are presented in the introduction section. Although this appears to have been suggested by the Evaluation Reference Group, it does not seem appropriate. •Further explanation of the approach would have been helpful. For example an explanation of a "cross sectional design" and explicitly stating this was a theory-based evaluation approach. Data sources and sampling seem suitable for this evaluation, though it would be useful to consider whether the total population source was excessive, and not the best use of resources. The evaluation drew on an innovative methodology, the results of which were not in the report, but seem very helpful. •The findings draw on strong mix of qualitative and quantitative data. Greater referencing of the sourcing of data, against the findings may have provided more confidence in the findings. Contextual factors seem appropriately considered. •Rather than offering a reflection on the purpose, the goal and objectives of the evaluation, the conclusions offer a brief, well formulated, summary of the findings of each evaluation question. The evaluation team may have considered a different approach to developing the conclusions. •Lessons stem from the findings, and have wider applicability. Some of the lessons are quite specific, whilst some are quite generic. However, they all seem useful, to different degrees. •The recommendations are highly aligned with supporting long-term planning for what comes after the program. They clearly come from the findings of the evaluation. With a couple of exceptions, the recommendations are well formulated. •The report has a clear structure that follows a logic sequence in terms of sections. There are some issues with inconsistent heading numbers and formatting and some inconsistency with spelling. Generally, though, the report is well written, with only a small number of sentences that may have benefited from revisiting. 			
Recommendations for Improvement: <i>The rater will identify topline recommendations to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, resources will be cited to assist evaluation managers in overseeing future evaluations.</i>			

There are four topline recommendations suggested to improve the evaluation.

- The evaluation report may have benefited from more attention to the overall coherence of paragraphs within sections and sentence structure. Writing a report that presents information in a coherent way, so that all related information is presented in one place greatly improves readability. Similarly, more attention to the sentence structure also improves readability, but generally offers the reader more confidence in the overall quality of the evaluation.
- Even though the Evaluation Reference Group may have suggested presenting the findings from the evaluation question on the theory of change in the introduction section, it does not help provide a logical flow to the report. It is recommended that the findings on the theory of change are presented with the rest of the findings, along with any suggested revisions to the theory of change.
- Referencing evidence, when presenting the findings, is helpful for the reader to understand where the findings have come from, and the strength of confidence in the findings. It therefore helps with overall confidence in the evaluation.
- The conclusions presented in the evaluation report are a summary of the findings. This is a common issue for evaluation teams when writing reports. However, the conclusions need to be more than this. It is helpful, when writing the conclusion section, to reflect on the goal and purpose of the evaluation, and to see how and why the evaluation has met these.

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 serves as a coherent standalone document, summarised into 5 pages. The most useful points from the evaluation are presented, therefore being helpful to decision making. As in the main body of the report, all the acronyms in the executive summary have not been spelled out it would be useful if they had been spelled out the first time that they are used. In the executive summary introduction, the summary describes the two main impact pathways, which is very helpful to understand the intervention. However, the executive summary introduction then provides the findings supporting the hypothesis and assumptions of the theory of change. As in the main report, this is presented before a description of the methodology and not in one coherent place, with the rest of the findings. Therefore, presenting the findings against the ToC in the introduction of the executive summary seems out of place. Most of the paragraphs in findings section in the summary start in the same way "To a large extent"; the team may want to consider using other terms and phrases to avoid repetitiveness.
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 all the necessary elements. However, as in the main report, the conclusions section is not a conclusion reflecting on the purpose of the evaluation, but instead it offers a summary of the findings. As a summary of the findings has just been presented in the executive summary, this section is an unnecessary repetition. As in the main report, for the conclusions section, the evaluation team perhaps should have reflected on whether the report has contributed to the purpose of the evaluation.
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 states the number of evaluation sites and describes the evaluation subjects. This is all well described, though it may have been helpful to note that the number of evaluation sites is the total number of implementation sites. The executive summary states that "quantitative data was analyzed using SPSS to produce the required descriptive statistics". However, in the report, the methodology section says that there was a comparison of the before and after intervention, using all outcome indicators in the results framework. Furthermore, the statistical tests carried out appear to have been used as inferential analysis rather than just descriptive. The findings section in the executive summary provides a coherent, mainly accurate, reflection of the main report findings – though some of the picture has not been presented. For example, when describing Program Coherence findings, the executive summary states that "No program was undermining the program interventions"; however, in the main report, the evaluation team report some issues of this nature.
SECTION B: BACKGROUND (weight 5%)		71%	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	The intervention is suitably presented in the introduction, at the start of the main body of report. The report provides sufficient information to understand the subject of the evaluation; with the location, dates and extension of the intervention, details of the funder, the budget, as well as the goal, objectives and outputs and implementation status of the intervention. All this information is concisely presented. Some sequencing of sentences could have been improved – such as fourth paragraph which introduces the date – could have preceded the third paragraph which presents the end date. Further details of the intervention are then presented in section 2 following a description of the context of the intervention. There are some sentence structure issues that impedes comprehension (e.g. "All facilities provided were aimed at enabling the schools to meet the recommended national standard of 1:40 pupil stance ratio" pg 7). Structure of this section could be improved, for example paragraph 2 of page 7 talks about the evaluation, whilst the preceding and subsequent paragraph discuss the intervention. Readability may be improved by presenting a more coherent narrative. When describing the intervention, there is a clear section on the program theory of change, both a narrative and graphic representation.

	ii	Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cities, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability... (as appropriate to the purpose of the evaluation).	Partially	There is a strong, relevant description of the ethnic groups of the rights holders, a description of the WASH context, and a description of social context of school learners. There is a clear section describing gender inequalities and girls with disabilities who are the main groups of rights holders. Providing a clearly labelled section that states that these are the rights holders may have strengthened this section. When describing the object of the evaluation (the intervention) the report appropriately states how many schools were involved, how many learners were reached, which is disaggregated by gender and disabilities. However, the number of learners reached is not aligned with what is stated in the Terms of reference (the programme has been able to reach 42,665 learners with basic drinking water, sanitation and hygiene services and 18,243 adolescent girls with menstrual hygiene products). The report provides, in Table 3, a very clear summary of stakeholders with duty bearing responsibilities and the roles and responsibilities they have. The table provides accessible reading and easy comprehension.
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 strong description of the context of the intervention, including a description of the ethnic groups – and on page 4, a strong section describing the how the context affects gender inequalities and girls with disabilities. There is a strong description of the context, in terms of needs and environmental challenges to address needs and how these relate to the intervention. And similarly, there is a good description of the economic and political contexts and how they relate to the implementation. This is a strong section of the report, which could be strengthened by adding some additional contextual description which is presented in the ToR. For example, the ToR notes about the challenge of declining water sector funding.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	Within the policy context section, there is a clear subsection on Sustainable Development Goals. This section outlines the goals and indicators that are relevant to the intervention objectives.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	Rights holders are described in several places within the context section. When describing the rightsholders, there is an appropriate description of their needs, as related to the aim of the intervention. This is particularly strong when describing gender inequalities and girls with disabilities and associates needs around menstrual hygiene.
Question 4.		Are key stakeholders, their relationships and contributions clearly identified?		
	i	Identification of implementing agency(ies), development partners, right holders, and additional duty bearers and other stakeholders; and of linkages between them (e.g., stakeholder map) (if relevant).	Partially	As described above, the report presents a table (table 3) presenting the stakeholders of the intervention. Both rights holders and duty bearers are presented, and their role in the project is briefly described. The description of UNICEF's role also describes the link / relationship between UNICEF and other stakeholders. Ideally, a bit of further detail describing the relationships between stakeholders could have accompanied the description of all stakeholders. Presenting this in a map may have been interesting addition, as this would have started to illustrate the WASH in school system in Uganda, which might have prompted further thought and understanding of the intervention.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	As described above, Table 3 describes the stakeholders and their role in the intervention. As suggested above, providing the links between stakeholders and how roles relate may have provided greater understanding of the context and the WASH 'system'.
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 3.1 presents the Evaluation Purpose. Within this subsection, the three purposes of the evaluation are described, although the clarity of the description could be improved. For example, the forward-looking purpose and the utilisation purpose appear to be the same thing – however, it is recognised that this is how the purpose is described in the ToR. The need for the evaluation is outlined as a 'a learning moment for long-term planning' beyond the intervention (pg.12). The users of the evaluation are described in the same subsection. This could be improved by being more specific about some of the users, such as describing which line ministries are the main users, and whether any specific staff within UNICEF are main users.
Question 6.		Are the objectives and scope of the evaluation clear and realistic?		
	i	Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Yes	The evaluation objectives are clearly presented in subsection 3.2. They are presented and described exactly as presented in the TOR. This is seen as sufficient, as long as no other objective were identified or deepened during the evaluation.

	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).	Partially	The description of the scope of the evaluation is copied, more or less from the Terms of Reference. One main difference is that the Terms of Reference asks that during the inception period the communities in target areas will be sampled. The scope in the Evaluation Report does not discuss this. Instead it appears that communities did not participate in the evaluation, only community leaders. Describing this variation in scope could have strengthened this section.
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	Under Section 2.3 "Object of the Evaluation" and subsection "Program Theory of Change" (also labelled 2.3), the results of the intervention being tested by the evaluation are described. The description is presented in several places in these sections, rather than in one place as a coherent description. The evaluation team may have considered stating clearly that the evaluation will assess each part of the results chain / impact pathway; the outputs, outcomes and impacts.
	ii	Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Yes	The causal relationship of the intervention is presented in a summarised narrative and graphical theory of change taken from the Terms of Reference.
	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	It is not explicitly stated that the evaluation is a theory-based evaluation. However, the approach uses the theory of change (ToC) as an assessment framework. The evaluation report states that the Evaluation Reference Group recommended that the ToC is assessed in subsection 2.3, under the Program Theory of Change. Therefore, in this section the evaluation team have described the evidence to validate the ToC. Evidence is presented along the impact pathways of the theory of change, and an overall assessment of the assumptions is presented. Even though presenting the assessment of the ToC in this section may have been recommended by the Evaluation Reference Group, it seems a strange place to present the evidence and findings. An assessment of the ToC may be better placed under the findings section. Despite this, positively, the section introduces two positive unintended consequences of the intervention. It would, therefore, have been beneficial to have an updated graphical ToC showing these additional casual links. There does not appear to be an assessment of any unintended negative impacts of the intervention.
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 criteria and questions are presented in an appropriate place, after the evaluation purpose, objectives and scope
	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 is suitably presented in Appendix 2, with reference to its location in the main report. The matrix includes a column which presents several indicators against each evaluation question. The indicators are suitable for identifying suitable evidence sources and providing an analysis point against the evaluation question.
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.	Partially	The report states that it adopted a cross-sectional design – but does not provide a description of what it means. Providing a simple definition of this design may have helped with comprehension for a broader set of non-research educated readers. The report explains clearly the use of a 'before and after comparison' of all of the outcome indicators. This is relevant and adequately robust for purpose of the evaluation. The report also draws on a theory-based approach, illustrated by the use of the programme theory of change and examination of the causal chain. However, the report does not explain that this was a complementary method to the before and after comparison. Explaining this may have provided a more complete description of the evaluation design.

	ii	Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	The data sources are fully listed in the evaluation matrix in Appendix 2, where they are presented against the indicators for each of the evaluation questions. Additionally, in subsection 5.2 the report presents the evaluation population, from which both qualitative and quantitative data is sourced. Further, the data collection methods are also presented, later in section 5.7, where both qualitative and quantitative methods are clearly described.
	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 method for calculating the sample size for learners is presented in section 5.4. However, the name of the formula is not provided, which may make it challenging for readers to understand whether this formula is appropriate. Although the report does not explicitly say how different perspectives are captured, the narrative provided demonstrates that different perspectives were included. For example, subsection 5.2 describes the evaluation population, describing different groups of beneficiaries. In the section (5.3), the report describes, very well, the sampling strategy for the inclusion of girls and learners with disabilities. And in subsection 5.5, there is a description of how learners were purposely sampled for interview. This strategy illustrates diverse selection of learner perspectives was captured. The rest of the evaluation population appears to be fully included as data sources with no sampling necessary, therefore capturing a range of perspectives involved in the intervention. In total, this is a large population source. There is no discussion of the implications of this on resources, or whether the evaluation team felt the data collection methods achieved saturation at an earlier point.
	iv	Clear and complete description of the methods of analysis.	Yes	There is a clear and complete section on data management, cleaning and analysis of both qualitative and quantitative data. The methods of quantitative and qualitative data analysis are appropriate. One minor suggestion to improve this section would be to describe why the quantitative tests are appropriate and to describe whether coding structures were used and how they were developed from the evaluation questions for qualitative analysis. The report states that the evaluation team pretested (piloted) the data collection tools, which is very good practice.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The report states that the evaluation team drew on Cross-tabulations, bivariate and multivariate regression as quantitative analysis method. These can all be inferential methods which can test hypothesis and causal relationships, e.g. between outputs and expected outcomes.
	vi	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Partially	The report has a separate subsection (5.11) describing the limitations of the evaluation. Four limitations are described well, with suitable mitigation methods also described. For example, the report states, quite rightly, that regression analysis can help mitigate the effects of timing and disentangle the influence of when an intervention was completed from its impact on observed changes. The report does not explain how this was achieved (it can be achieved by incorporating time as a variable in the regression model). The report also states that one limitation was that baseline outcome data was not disaggregated by school. The report states that "To mitigate this limitation, administrative data for before and after intervention was captured". This mitigation method would benefit from further elaboration.
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	There is explicit and full reference to conducting the evaluation according to UNEG standards.
	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	There is explicit reference to following the UNICEF procedures for Ethical Research Involving Children. Earlier in the report, in section 5.3, it is stated that a human-rights based approach was incorporated, with consent obtained, privacy and confidentiality ensured. As an illustration of good practice, the ethical protocols were reviewed by an appropriate university review board, as well as the Program Evaluation Review Committee. This section, then also further elaborates on the process of gaining and following informed consent and maintaining confidentiality. In total, this is a strong section.
Question 11.		Does the evaluation incorporate innovative practice that adds value to the evaluation process?		

	i Innovation practice is used to improve the quality of evaluation process. This could be evident in several ways such as the design of the methodology (i.e. use of technology for data gathering, extensive participatory processes, systematic analysis processes such as collaborative outcomes reporting and incorporation of big data, specific strategies to address complexity such as outcome harvesting, strong child rights focus), or ways of sharing of evaluation results.	Yes	The reported evaluation process seems high quality, in the data collection methods used, in the extent of intervention population participating as data sources, and in the ethical procedures reported. There appears to have been good triangulation of data from suitable qualitative and quantitative methods. Perhaps the most 'standout innovative' practice is the capturing of human-interest stories from girls and learners with disabilities. Three methods were used to capture a range of stories: 3-minute videos, audio recording of stories, and a photo booklet. The capturing of these stories of change was an opportunity to capture the richness of an intervention implemented in complex environments – that the other methods, particularly the quantitative methods would not capture.
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.	Yes	The findings present an adequate level of evidence against each of the evaluation questions. The evidence is, mostly, well described, allowing an evaluative judgement to be made on each evaluation question. In the report, when answering the effectiveness and efficiency questions, the findings are nicely supplemented by tables of data. One suggestion which could have strengthened the findings section is indicating more consistently where the evidence came from; i.e. which source(s). This can often be done well through the use of footnotes.
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	To evaluate program effectiveness the evaluation team adopted a systematic approach assessing outputs and then objectives – therefore suitably following the program Theory of Change.
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 findings section uses a strong mix of qualitative and quantitative data. However, as reported above, it is often unclear the source of the data, whether it came from the survey, monitoring reports, interviews or focus groups. Additionally, occasionally, the report offers quantitative data that has been sourced from qualitative interviews, for example, "The findings from interviews with girls show that 87% had used the washrooms during menstruation". Whilst this is entirely valid, when further quantitative findings are presented without the source referenced, it leaves a question over the strength of the finding.
	ii Findings are clearly supported by, and respond to, the evidence presented, including both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.	Yes	For each evaluation question, evidence and findings are first presented, with a "Overall Assessment" summarising the evidence. This is the conclusion finding, against the evaluation question. These findings sections do clearly draw from the preceding evidence, and present both positive and negative findings. These findings are aligned with the performance indicators presented in the Evaluation Matrix in Appendix 2.
	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 often consider contextual factors, particularly for the non-achievement of results. For example, the report explains that the target quantity of latrine blocks was not met because of the COVID-19 pandemic and the war in Ukraine; and water systems were not working due to the unavailability of equipment in local markets. There is a specific evaluation question that asks about the factors influencing achievement or non-achievement of results that discusses contextual and managerial issues contributing to non-achievement of results. The evaluation question about sustainability introduces and discusses the role of contextual, organisational and managerial issues, such as quality of construction, availability of spare parts in local markets etc. As discussed above, the findings related to the effectiveness evaluation question analyse and are presented according to the progression of the programme Theory of Change.
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.	No	There does not appear to be an explicit assessment of the programme monitoring system. Although there is not a direct question related to the monitoring system, the evaluation team could have considered the adequacy and role of the monitoring system in program effectiveness and efficiency.
SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	25%	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 are presented and structured according to the DAC criteria and evaluation questions. Rather than offering a reflection on the purpose, the goal and objectives, of the evaluation, the conclusions offer a brief, well formulated, summary of the findings of each evaluation question. Reflecting on the purpose of the evaluation would have been a good approach to developing the conclusions section.

	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.	No	As the conclusions are direct summaries of the findings against each of the evaluation questions, they are appropriately derived. However, as discussed above, the evaluation team may have considered a different approach to developing the conclusions.
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.	Partially	Lessons are structured according to the DAC criteria used in the evaluation. Lessons do stem logically from the findings, and have wider applicability. Some of the lessons are quite specific, whilst still being very helpful, for example, with reference to Karamojong social context. Many of the lessons, whilst helpful, are quite generic. For example, the lesson for the Relevancy criteria suggests, "Stakeholder consultations and engagements are critical in program design... beneficiaries, local government leaders, and the line ministry are some of the key stakeholders who must be consulted and engaged".
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Partially	Lessons are presented as one or two short sentences, mainly in bullet form. Although short and concise, the lessons do offer enough detail to be useful. However, the sentences are not always clearly structured and would benefit from being revisited. For example, "Health benefits of a safe water supply can be maximized through a reduction in the disease burden and improved nutrition through micro-irrigation of school gardens. These gardens can produce vegetables and food."
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.	Partially	The recommendations are highly aligned with supporting long-term planning for what comes after the program. They clearly come from the findings of the evaluation. With a couple of exceptions, the recommendations are well formulated. The exceptions are where findings have been presented along with the recommendation, which is not necessary. For example, recommendation 2 starts as, "WSSBs are unable to raise adequate funds for major repairs like replacing a water pump. The District Water Offices (DWOs) do not have sufficient budget for O&M to handle additional requirements. The majority of the beneficiary communities did not contribute to the O&M of the WSS".
	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 appear actionable for the intended user. The intended users, however, may find additional detail useful for guiding implementation of the recommendation.
	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 evaluation report describes the process by which the recommendations were developed. The process involved "key stakeholders" and members of the "Evaluation Reference Group". However, there is no further detail on describing the composition of the Reference Group or which stakeholders were involved. Therefore, the reader of the report is unsure of whether this includes duty-bearers, as well as rights holders.
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.	Yes	The recommendations are presented in a clear table, with columns for the recommendation, for the intended user of the recommendation, and the suggested timeline for implementation, prioritising recommendations. This is well presented and appropriate.
SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)			64%	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 present all the relevant and necessary information very clearly. There are some minor errors in the numbering in the table of contents (and therefore with the heading numbering within the report). And some minor formatting with List of tables (the first "table" should have an upper case "T"). A list of acronyms has been provided, but it is noted that not all acronyms are in the list. A map of the country and the intervention region has been provided, which is great, but it is not referenced in the report.
	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	An extensive annex is presented, with all required annexes and also appropriately referenced in the table of contents.
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 report has a clear structure that follows a logical sequence in terms of sections. There is a useful one-page introduction, that lays out the report sequence and the main details of the report, so the reader can get an early understanding of the subject of the report. The findings section is well structured, following and answering each of the evaluation questions in turn. And at the end of each of the findings to the evaluation questions, there is a helpful overall assessment. The report, on the whole, has used headings appropriately. However, there are some issues with inconsistent heading numbers and formats, (for example Section 2.3 is a main heading on page 6 and then 2.3 is also a subheading on page 8; heading 6.3 is not the same size as heading 6.1 or 6.2). If this formatting was more consistent, it would help readers navigate the report more easily.
	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.	Partially	On the whole, the report follows UNICEF guidelines for structure of evaluation reports. The only divergence appears to be that lessons are presented before the conclusions. The report states that the Evaluation Reference Group recommended answering the evaluation question about the relevance of the program Theory of Change in an earlier introduction section (section 2.3). However, this is a bit strange, as it means findings are presented in the introduction section, including tables of statistics and results. The assessment of the Theory of Change may be better suited under the appropriate question in the findings section.
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	The main body of the report is 42 pages, which is reasonable and not exceeding the number of pages specified in the ToR (45 pages).
	ii) Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	Generally, the report is well written, though with some sentences that may have been revised (e.g. first sentence of third para on page 1). Some of the sequencing of the paragraphs may have been improved (e.g. the fourth paragraph, which introduces the dates of the intervention, could have preceded the third paragraph which presents the end date of the intervention). There are some minor issues with spelling consistency (for example, program and programme), and some acronyms are not spelt out when first used, both in the executive summary (e.g. NDP page x; IEC material page x; O&M page xi; MFPED page xi) and in the main body (GOU page 1). There is inconsistency with the spacing, fonts, and sizing of the footnotes. If corrected, this would have improved presentation.
	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	There is a good use of a variety of visual aids, including tables, infographics, maps, tables, figures and photos. They have been appropriately labelled and referenced. However, many of the figures of graphs are photos, and are blurry, and in some instances are covering text (particularly in the Section 6.3). Furthermore, some of the tables (e.g. Table 5, and Table 6) are presenting statistical analysis, which would be better placed in an Appendix, with appropriate narrative description in the body of the report.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	90%	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 report references and appears to follow the principles and standards of human rights into the process of evaluation - CRC, CCC, CEDAW have not been referred to.
	ii) Clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation (for example, a reference group is established, stakeholders are involved as informants or in data gathering).	Yes	The evaluation approach was participatory, engaging both rights holders and duty holders as evaluation informants through the use of several different standard data collection tools, such as survey and focus group discussions. These tools were pretested in some of the targeted schools to assess suitability. Furthermore, the evaluation findings, lessons learned, and recommendations were presented to the key stakeholders during the regional validation workshop and the draft evaluation report was submitted to an Evaluation Reference Group (ERG). To increase the level of participation from both rights holders and duty holders, the evaluation may have considered designing the tools with representatives from each group – this can be particularly useful when designing tools for children. However, engaging children in tool design can be good practice it requires resources that may not have been available to the team for this work.
	iii) Language is empowering and inclusive, avoiding gender, heterosexual, age, cultural and religious bias, among others; use terminology of rights holders and duty bearers; data is disaggregated by marginalized group; differential results are assessed (distribution of results across different groups).	Yes	The language used throughout the report is appropriate, using terminology such as rights holders and duty bearers, but also "learners with disabilities". Considering the topic of the evaluation, there is a strong level of data disaggregation, between genders, but also between areas with different ethnic groups. There is also a consistent and appropriate level of focus on disabilities throughout the evaluation report.

	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	As mentioned above, there is appropriate level of focus on disabilities throughout the evaluation report, indicating an inclusive evaluation process – both in terms of gender, disabilities, and ethnic groups. This appears to have been facilitated by the topic of the evaluation, and the evaluation questions, which for example have specific questions on cross-cutting issues, such as gender, disability and human rights. However, the evaluation appears to have carried this out well.
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)	7	
	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.	Satisfactorily integrated	The evaluation does not explicitly assess whether sufficient information was collected during the implementation period on result indicators to measure progress on human rights and gender equality results. However, both the goal and the objective of the evaluation have specific focus on gender-sensitivity, disability, and adolescent sensitivity. And although there was a dedicated evaluation question on gender equality and human rights.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	The evaluation report specifically states that gender has been mainstreamed throughout the evaluation, in sampling, data collection, analysis and report writing. The report also states that female researcher assistants interviewed girl learners. Although gender inclusion appears to have been appropriately considered, the evaluation team may have also considered how and whether power dynamics (as well as inclusion) were addressed by the intervention. The report appropriately states that ethical standards were considered, and as mentioned above, the evaluation team sought appropriate ethical clearance.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	The evaluation has a specific background section that includes an intersectional analysis of inequalities within and between boy and girl learners and girls with disabilities in the education system, as well as how social context affects women and girls in the intervention location. Findings report evidence from different sources reflecting the voices of the different social groups involved in the intervention, such as interviews and focus groups with girls, survey results that involved participation of girl and boy learners, and senior woman teachers. Unanticipated results affecting gender equality are reported – however, all unanticipated results are positive. For example, unexpected results from the WASH intervention included “time saved by learners from walking long distances to collect drinking water” and “Teachers have been attracted to these schools which has also reduced teachers’ absenteeism from classes”. The evaluation offers two recommendations specifically addressing gender equality issues, targeted to the appropriate stakeholders.

SWAP Rating Guidance

<p>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.</p> <p>a. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results?</p> <p>b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?</p> <p>c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?</p> <p>d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation?</p>
<p>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</p> <p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex?</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?</p> <p>d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?</p>
<p>iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</p> <p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality?</p> <p>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable?</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described?</p> <p>d. Does the evaluation report provide specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area?</p>