

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Independent Evaluation of the Accelerating Sanitation and Water for All Programme (ASWA II)

REPORT RATING SUMMARY

Overall Rating	72%	Satisfactory	
●●●●●	Exceptional (96% - 100%)	5	
●●●●	Highly Satisfactory (87.5% - 95.99%)	4	
●●●●	Satisfactory (62.5% - 87.49%)	3	Meets UNICEF/UNEG standards for evaluation reports 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	Independent Evaluation of the Accelerating Sanitation and Water for All Programme (ASWA II)
Report sequence number	Evaluation Office/65/2024/19611
Region	HQ
Year of report	2024
Office	Evaluation Office
Coverage (countries)	Burkina Faso, Cote d'Ivoire, Mali, Mozambique, Niger, Nigeria, Somalia, Republic of Tanzania
ToRs present	Yes
Date of review (dd/mmm/yyyy)	January 12, 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	No
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)	No
Humanitarian action (cross-cutting)	No
Evaluation object	Programme
Evaluation type	Summative and formative
Evaluation strategy	Mixed methods
Evaluation design (primary method used)	Theory-based
Evaluation level	Output & Outcome
Geographic scope	Multi-country
Primary SDG(s) covered (number)	3, 4, 5, 6, 10, 11, 13

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.

It is clear from reading the preface that this evaluation faced challenges, which meant that the evaluation manager had to step up and closely manage the evaluation and its outputs.

- Against this background, this is still a robust evaluation with sound findings. However, the evaluation does not communicate a detailed picture of the DGIS-funded ASWA programme, its achievements and areas for improvement. This may partially be related to the fact that the ToC was not revisited during inception and due to a long list of evaluation questions, sub-questions and judgement criteria to work through.
- Whilst the evaluation included country summaries and thematic case studies, there is no clear overview of how the programme fared in each country and of thematic learnings.
- The evaluation design and methods are sound, ethical standards were followed, and the findings speak to the questions and sub-questions but the way the data was collected and / or analysed did possibly not maximise the value-add of this large and complex evaluation.
- Whilst gender, equity and inclusion sensitive, the evaluation did not fully engage with this topic nor with different stakeholder groups; this comes across in the short description of stakeholders in the report, in particular of vulnerable groups and their needs.
- This notwithstanding, the evaluation report is clearly structured and conclusions, lessons and recommendations convey clear messages, albeit at a high level.

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.

Recommendations for improvement include:

- For future evaluations, it may be helpful to make a revisit of the ToC mandatory in the ToR and to ensure that the chosen evaluation personnel is fully competent in designing a theory-based evaluation and analysing a complex set of data from a wide range of data sources and able to capture the intervention across different levels of implementation (HQ, regional, country).
- It may also be helpful to aim for a shorter evaluation matrix with fewer, clearly prioritised main evaluation questions, supported by a small set of sub-questions and a manageable set of judgement criteria in line with the space provided in UNICEF evaluation reports for presenting and substantiating findings.

SECTION RATINGS

SECTION A: EXECUTIVE SUMMARY (weight 5%)	83%	Comments on Rating
Question 1. Can the executive summary inform decision-making?		
Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Yes	The executive summary is clearly structured, provides a good summary of the key findings and a useful document to inform future decision making about the programme. The executive summary is 5 pages long.

	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 an overview of DGIS-funded ASWA II programme, the purposes and objectives of the evaluation, the intended audiences, and the evaluation methodology (including sampling); summary findings are provided by EQ, followed by a single conclusion statement, which is appropriate. Key recommendations are listed. However, lessons learned are not included although they are part of the evaluation objectives in the ToR.
	iii	Includes all significant information needed to understand the intervention and the evaluation AND does not introduce new information from what is presented in the rest of the report.	Yes	The executive summary provides a comprehensive picture of the evaluation and does not introduce new information.
SECTION B: BACKGROUND (weight 5%)			64%	Comments on Rating
Question 2.		Is the object of the evaluation clearly described?		
	i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Partially	The programme is well summarised, and location, implementation timelines, budget and implementation status are given. Relevant UNICEF-wide strategies and guidance is referenced. However, no information is given on any ASWA II-specific programming approaches. The ToC logic is copied from the original ToR for the assignment. No reference is made to 2 countries having been part of phase I of ASWA and what this might mean in terms of programme implementation. Covid-interventions, which were added after the programme started, are not described.
	ii	Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cites, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Partially	The evaluation audiences are explained and a distinction is made between types of organisations, primary and secondary users and uses. However, the language of rights holders and duty bearers is not used. Rightsholders are mentioned in passing as 'the communities serviced by ASWA II'. Implementation areas are specified in the background section. The description of duty bearers and rights holders is very short.
Question 3.		Is the context of the intervention clearly described?		
	i	Clear and relevant description of the context of the intervention (i.e. relevant policy, socio-economic, political, cultural, power/privilege, institutional, international factors) and how context relates to the implementation of the intervention.	Partially	The policy and socio-economic context is mentioned very briefly in the background section; reference is made to WASH status, population size and economic status as well as fragility by country. However, the description is minimal and no clear reference is made to how the context relates to project implementation.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	The background section describes key targets and guidance for UNICEF WASH programming, including SDG6, the UNICEF WASH strategy and strategic plan.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Partially	The rightsholders of the intervention are mentioned in passing, no detailed information on their needs in the eight ASWA countries is provided.
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	The main project stakeholders are briefly described and their roles are sketched out.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	Whilst DGIS is identified as the donor, the role of UNICEF is not described in detail, e.g. there is no clear description of the different roles of UNICEF HQ, regional offices and country offices despite significant funding allocated to them as part of the programme. The role of implementing / supporting partners is not described in any detail and it is not clear which partners operated in which countries.
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.	Yes	The purpose of the evaluation is defined and a rationale is provided for why it is needed (to support planning for a phase III of the programme), evaluation users and uses are spelled out.
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 clear and no changes were made from the original 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).	Partially	The main thematic, geographic scope and timeframe of the evaluation are briefly outlined. The ToR also refers to advising UNICEF HQ and country offices on how the programme is positioned for a rigorous endline evaluation, including advice on planned household services. However, this aspect of the ToR is not referenced in the evaluation scope. Whilst the thematic scope covers all themes, it does not go into any detail on sub-themes of the programme such as WASH and nutrition, WASH in conflict-affected / fragile areas, knowledge management, finance and costing and climate resilience.
Question 7.	Is the theory of change, results chain or logic well articulated?		
	i Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Partially	The ToC of the programme is summarised and a copy of the programme ToC is in the narrative. However, the summary of the ToC is a copy and paste from the ToR and there is no indication in the report that the evaluation has critically examined the programme ToC.
	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 relationships are clearly presented in the programme ToC.
	iii For theory-based evaluations, the theory of change or results framework is assessed, and if requested in the ToR, it is reformulated/improved by the evaluators.	No	The ToC is not assessed or critically discussed. Whilst the ToR did not explicitly require the evaluation to revisit the ToC, some discussion on the extent to which the ToC logic held true, or did not, would be expected here. There are also some issues with the output - outcome logic of the current version of the ToC, which are problematic for assessing contributions. For example, outputs as in the ToC could be considered a level higher, at outcome level (e.g. targeted populations gain access to sustainable services, regular sustainability checks inform local and national sector planning). One of the outcome assumptions assessed in SEQ 4.1 (the extent partnerships with the private sector were mobilised as intended) is not represented in the ToC, hence creating a mismatch between the ToC and the evaluation findings - limiting the extent to which the evaluation clearly explains its 'theory-based' approach. Further, the programme adapted in response to COVID-19, spending a significant amount of money on this - however, none of the related programming elements and causal chain are reflected in the ToC and, presumably, in the results framework and the evaluation did not explicitly assess the effectiveness of this programming element.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	77%	Comments on Rating
Question 8.	Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? <i>UNICEF evaluation standards refer to the OECD/DAC criteria - Relevance; Coherence; Effectiveness; Efficiency; Sustainability; Impact (not all are necessarily relevant for all evaluations). Evaluations should also consider equity and leaving no-one behind, gender and human rights based approach (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.</i>		
	i Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Partially	The evaluation questions and sub-questions meet the purpose of the evaluation. Criteria for answering evaluation questions are developed in the form of indicators in the evaluation matrix. However, the evaluation matrix contains a large number of questions (13) some of the sub-evaluation questions contain more than one question (e.g. 3.3, 3.4) and some contain many judgement criteria (up to 14). Some sub-questions and judgement criteria distract from the main purpose of the evaluation e.g. SEQ 3.4 on the development of policies and strategies related to improved urban WASH services for a programme that is mostly aimed at rural populations. In other cases, there are simply too many criteria for each SEQ, given the overall space in the evaluation report for presenting findings. Also, the SEQs used to assess assumptions between outputs and outcomes are not aligned with the assumptions stated in the ToC (e.g. SEQ 4.1 refers to private sector partnerships whereas the assumptions stated in the ToC do not mention this aspect)
	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 includes indicators for answering each evaluation question and sub-question and specifies data sources and analysis methods.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?		

	i Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Partially	The evaluation design is based on a set of standard, robust data collection and analysis methods appropriate for the purpose and objectives of the evaluation. The methodology section summarises the data collection, sampling and data analysis methods used. However, the case study sampling and data collection approaches are not fully described in the evaluation report or annexes. For example, it is not clear from the evaluation report that thematic case studies had separate questions and separate FGDs were conducted on these topics in some countries, as described in the inception report.
	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	Yes, data sources include a range of qualitative and quantitative data; quantitative data is based on secondary data, including a separately conducted impact evaluation.
	iii Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Partially	A sampling strategy is provided and diverse perspectives are covered (children, water users, community leaders, vulnerable households, heads of schools, water operators, UNICEF at different levels, implementing partners, donor, duty bearers at different administrative levels). Reference is made to a balanced gender representation but beyond this, the sampling strategy section does not provide a detailed description of how these diverse perspectives have been included. Table 7 leaves out the following respondent groups at the village level: beneficiaries / users, community leaders and water user associations.
	iv Clear and complete description of the methods of analysis.	Yes	Quantitative and qualitative data analysis approaches are described in the report. The term 'contribution analysis' is used in the design and analysis sections, however a full contribution analysis was not conducted (setting out fully contribution stories related to specific outcomes). Rather, the evaluation assessed data in the spirit of contribution analysis, by assessing data related to outputs, outcomes and impact and making judgements on whether the programme had contributed to these in the 8 countries.
	v Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	Yes, the combination of qualitative and quantitative data analysis methods used in the evaluation allows for drawing causal connections between programme outputs and 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 evaluation report lists the risks and limitations identified during the inception phase and how these were mitigated in the evaluation process. However, the findings section identifies several data gaps related to programme monitoring data, which are not referenced in the limitations section.
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	The UNEG norms and standards are referenced and project-specific ethical considerations are described for independence, impartiality and credibility. Other standards not referred to explicitly in 10 (i) but described by the evaluation report are: disclosure, confidentiality, usefulness, informed consent and do no harm. These touch on ethical safeguarding related to confidentiality and informed consent and avoidance of harm.
	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	UNICEF Ethics review board approval was sought and granted, and the confirmation is provided in Annex 8. The evaluation involved interviewing children and the related data collection methods were part of the ethical review process mentioned above; this confirms UNICEF procedures were followed. However, no explicit reference was made to safeguarding procedures for children in the evaluation report.
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	A participatory method (evaluation H) involving interviewing school children was used to obtain their feedback on the design and use of sanitation facilities in schools.
SECTION E:	EVALUATION FINDINGS (weight 25%)	58%	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 findings provide sufficient levels of evidence to answer all questions and sub-questions. However, evidence is not provided for all criteria in the evaluation framework. Below is an illustration of this: - under 1.3 (prioritisation of equity and the needs of the most vulnerable) the findings mostly speak to criterion 1.3.1 the type of vulnerable groups considered and not to criteria 1.3.2 (the nature and type of strategy developed by ASWA II to prioritise vulnerable groups in each component) and 1.3.3, (the degree to which identified vulnerable groups are considered in WASH sector policies related to each component); - under SEQ 1.4 (response to emerging needs), criterion 1.4.3 (factors facilitating or hindering adaptations) is not assessed; - under SEQ 3.2 (WASH services in institutions), the evaluation framework specifies 14 judgement criteria but the findings do not cover all; - under SEQ 3.3 (systems strengthening, i.e. monitoring and climate resilience), the 8 judgement criteria are not strictly followed and the findings discuss climate resilient programming for service delivery rather than related to systems strengthening; - SEQ 3.4 findings report on programming approaches to reach urban poor rather than to the development of policies / strategies as specified in the SEQ; - on SEQ 3.5 (WASH and nutrition), findings do not explicitly refer to the 5 judgement criteria e.g. criterion 3.5.2 the number of countries using intersectoral coordination mechanisms on WASH and nutrition; - for SEQ 7.2 (Country offices adopting VfM strategies) the extent of VfM achieved compared to initial estimates is reported (criteria 7.2.4-6) but not findings related to criteria 7.2.1 and 7.2.2. - for SEQ 12.1 (how and to what extent did ASWA II integration of nutrition benefit children in communities and schools?) criterion 12.1.2 is not assessed (number of communities / populations reached by ASWA II nutrition activities by year and country).
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Partially	The ToC is used explicitly in the findings section. EQ 4 refers to the extent to which assumptions in the TOC between outputs and outcomes are upheld. However, in some parts, the evaluation framework moves away from the ToC in that it identifies questions that are relevant to the programme implementation and results but not reflected in the ToC. For example, under EQ 9 (capacity of WASH sector to implement sustainable WASH services), SEQ 9.2 asks to what extent the programme contributed to the development of a WASH market. However, the WASH market is not presented as an element in the ToC. See also response to Questions 8(i) in this assessment for further comments on the use of the ToC in the evaluation.
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	Yes, the evaluation uses credible evidence and some level of triangulation is made explicit. However, the evaluation does not explicitly refer to the four thematic case studies that were conducted as part of the evaluation as a source of evidence.
	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	Yes, findings are generally supported by evidence, often triangulated from different sources. However, not all findings are firmly grounded in all the judgement criteria of the evaluation framework, partially because there are sometimes too many criteria to assess them all (see response to assessment question 12 (i) for more information).
	iii Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyse the logical chain (progression -or not- from implementation to results).	Partially	Yes, causal factors are regularly identified in the findings section and both positive and negative factors are referenced. Several sub questions refer to aspects that speak to the progression from implementation to results. However, the logical chain is not systematically assessed for the overarching programme ToC.
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	Yes, the evaluation findings explicitly examine ASWA II baseline data, outcome monitoring and real-time monitoring, and the use of these tools for decision making via EQ5 and also across other findings.
SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	100%	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).	Yes	The conclusions reflect the objectives of the evaluation - which were organised along the OECD-DAC of relevance, effectiveness, efficiency, coherence, sustainability and impact. They provide a good overview and summary, which helps identifying key gaps and focus areas for the next phase of the programme.

	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.	Yes	The conclusions reflect the findings and appropriately summarise the key points, both positive and negative, of the intervention.
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 lessons build on the findings drawing out themes that have emerged across several findings. The lessons are applicable more widely in the sector (e.g. lesson on systems strengthening) and to development programming more widely (innovation, planning, monitoring and evaluation).
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	The lessons are short and concise, yet provide enough detail to be useful.
SECTION G:	RECOMMENDATIONS (weight 15%)	75%	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 purpose of using evaluation insights for shaping the next phase; they build on the findings and conclusions and are clearly formulated.
	ii Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Yes	The recommendations are useful and actionable; they are also specific to individual stakeholder groups and applicable to the next programming phase of ASWA.
	iii Process for developing the recommendations is described, and includes the Involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	Partially	The process for arriving at recommendations is briefly outlined. Duty-bearers (i.e. ERG) had an opportunity to shape the recommendations, whilst rights holders were not directly involved. However, no explanation is provided for why rights holders were not involved in shaping recommendations.
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	Yes, the recommendations are specific to individual stakeholder groups. However, they are not prioritised and no timeframe is given for actioning recommendations.
SECTION H:	REPORT STRUCTURE AND PRESENTATION (weight 5%)	86%	Comments on Rating
Question 19.	Does the evaluation report include all relevant information?		
	i Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes; list of acronyms/abbreviations, page numbers.	Partially	The introduction includes the name of the intervention that is evaluated, and the implementation countries; the evaluation commissioner and evaluation companies are described in the preface, there a table contents, the title page includes the date of the report; there is a list of acronyms; the report contains page numbers. However, the report does not have a list of tables, graphs and figures.
	ii Annexes include: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework (unless included in report body), list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology, information about the evaluator(s), etc.	Yes	The annexes include all the necessary items listed here. It would have been useful to have summaries of case studies or country findings also included as part of the annexes.
Question 20.	Is the report logically structured?		
	i Structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles, well formatted).	Yes	The report is clearly structured in line with the quality assurance guidance from UNICEF. It includes an introduction, a background section, the purpose, objectives and scope of the evaluation, the design and methodology followed by well structured findings, lessons, conclusions and recommendations.
	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	Yes, UNICEF guidelines for evaluation reports are followed, as described above in response to question 20(i).
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 report is just over 60 pages long, which is in line with other UNICEF evaluation reports. No page-length was specified in the ToR.

	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Yes	The report is written in an accessible language and is generally free from grammatical and spelling mistakes. However, the report language uses phrases like 'appears to be' in several places, questioning the confidence in the findings when this language is used (e.g. EQ4 summary p34 ' the programme appears to have been able to adapt effectively').
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Partially	The report uses visual aids, mostly in the form of tables and photos. All tables are clearly numbered. However, sources are not always provided for tables. In several places, the use of bar charts would have been appropriate to visually convey the information instead. The photos in the report do not illustrate specific findings but are give general impressions from the programme overall.
SECTION I: EVALUATION PRINCIPLES (weight 10%)		60%	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.	No	No, none of the rights-based concepts or frameworks are referenced in the evaluation.
	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	Yes, the evaluation refers to the set up of an evaluation reference group, and the documentation shows that the reference group commented on the inception report and fed back on draft recommendations. Different stakeholder groups, including rights holders, were involved as informants of FGDs, and Evaluation H (school children).
	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	Yes, the language avoids bias but does not use the terminology of rights holders and duty bearers. While views from marginalised / vulnerable groups were sought via FGDs, their perspectives do not come through strongly in the findings section, neither do the perspectives of other stakeholder groups. Concerning data disaggregation, the evaluation observes that the programme did not disaggregate data by gender, age or disability or other marginalised groups and was therefore not in a position to assess this aspect.
	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	Yes, the evaluation assesses access to WASH services and both water and sanitation are considered human (and child) rights anchored in the CCCs. The evaluation also assesses the extent the ASWA programme addressed gender (SEQ 13.2), the rights of vulnerable people (by age, female-headed HHs, economic status, persons living with disabilities) (1.3, 3.1, 3.2), and of marginalised communities (SEQs 1.3, 4.2).
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)	5	
	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	Yes, the evaluation assesses to which extent gender is included in the programme's results framework and monitoring instruments. No, the evaluation does not include an objective specifically on gender equality and human rights. However, these aspects were also not included in the purpose and objectives of the ToR. Yes, attention to gender, disability and vulnerable groups was mainstreamed in the evaluation framework via integration into several SEQs (see response to question 22 iv).
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	The evaluation methodology explains how gender and equity issues are addressed in data collection; this included separate focus group discussions with vulnerable households, a thematic case study on 'gender and inclusion' and a participatory method (Evaluation H) to consult school children. However, no reference is made to how data analysis considered gender. Yes, the mixed method approach was appropriate to assess the extent to which gender was mainstreamed in the programme but secondary data to assess the effects were poor. It is not entirely clear to which extent a diverse set of data was used to assess the accuracy and credibility of the gender-specific analysis as the findings are highly summarised and do not extensively reference triangulated evidence sources. Yes, the sampling frame included vulnerable HHs and school children as two separate stakeholder groups for data collection. Secondary data was consulted for gender-disaggregated data but this data was not available. Yes, according to the ethical protocol used, ethical standards were upheld.

iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially integrated	No, the background section does not include a detailed intersectional analysis of different social groups or references human rights frameworks. No, the findings do not explicitly triangulate the voices of different social groups and data disaggregation was not possible due to data constraints. No, unanticipated effects on human rights and gender equality are not spelled out in the findings Yes, the report recommendations speak to gender equality and women's empowerment.
-----	---	----------------------	--

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>