

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Summative Evaluation of UNICEF Interventions to Create an Enabling Environment for WASH (OUTPUT-1) in Pakistan (2018-2022)

REPORT RATING SUMMARY			
Overall Rating	86%	Satisfactory	
●●●●●	Exceptional (96% - 100%)	5	
●●●●○	Highly Satisfactory (87.5% - 95.99%)	4	
●●●○●	Satisfactory (62.5% - 87.49%)	3	Meets UNICEF/UNEG standards for evaluation reports. Decision makers may use the evaluation with confidence
●●○●●	Fair (35% - 62.49%)	2	
●○○●●	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
Title of the evaluation report	Summative Evaluation of UNICEF Interventions to Create an Enabling Environment for WASH		
Report sequence number	Pakistan 64-2023-19045		
Region	SAR		
Year of report	2023		
Office	Pakistan		
Coverage (countries)	Pakistan		
ToRs present	Yes		
Date of review (dd/mmm/yyyy)	February 12, 2024		
Name of review firm	IOD Parc		
CLASSIFICATION OF EVALUATION REPORT			
Management of evaluation (Managerial control and oversight of evaluation)	UNICEF managed		
Unicef goal areas (Alignment with strategic plan priorities)			
Every child survives and thrives	Yes		
Every child learns	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	Yes		
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)	SDG 6 (particularly 6.1, 6.2) and SDG 4 (4.3)		
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>			
<p>The evaluation report of the 'Summative Evaluation of UNICEF Interventions to Create an Enabling Environment for WASH (output 1) in Pakistan' is a robust and thorough evaluation that can be used with confidence by decision-makers.</p> <p>Strengths of the evaluation include:</p> <ul style="list-style-type: none"> - the purpose, scope, objectives of the evaluation are clearly shared and any changes from the ToRs are clarified - the methodology is 'for-for purpose' using participatory, theory-based and utilisation focused approaches , and mixed methods - the context shares the general background contextual information, as well as relevant WASH issues - the key stakeholders, and their roles and responsibilities are shared well - the theory of change is described and critiqued, and causal pathways explored well and analysed within the findings - the findings are evidence-based with clarity regarding the types of evidence utilised - the use of visuals to illustrate points and engage the reader is a strength of the report - the conclusions are arranged logically by evaluation criteria and summarise the findings in a crystallised way - the lessons are rooted within the findings of the evaluation, but written in a generalised way appropriate for different contexts - the recommendations are clear with overall steers and practical recommended actions. They are also prioritised with clear timeframes. - a strength of the report is the treating of gender and equity as separate issues (unlike many evaluations), and also consideration of human rights and climate change as cross-cutting themes <p>Shortcomings include:</p> <ul style="list-style-type: none"> - the report significantly exceeds the page limit at over 100 pages - the evaluation matrix does not include an overarching evaluation question for each criteria area - although the conclusions summarise the findings well, they do not add further strategic insight to the findings as would be expected within conclusions - some of the recommendations are targeted to more than one primary stakeholder which risks a lack of clarity around which organisation is taking responsibility - whilst there is use of some benchmarks within the analysis there is scope to do more (thematic issues such as open defecation , access to clean water, expected utilisation rates) - the conclusions and recommendations cover gender in a commendably systemic way, but include very limited discussion of other cross-cutting issues (climate change, LNOB including people living with disabilities) 			
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>			

The report could be improved by:

- within the evaluation matrix, an overarching evaluation question for each criteria area could be included
- conclusions could add further strategic insight beyond being a summary of the findings (e.g. focusing strategically on the WASH policies/ governance/ system and UNICEF's role and contribution)
- ensure clarity of one lead organisation for each recommendation (and if needed sub-recommendation)
- explore other benchmarks that could be woven into the analysis to ensure that there is comparative analysis against international best practice/ global standards
- consider whether the conclusions and recommendations should include other cross-cutting issues (climate change, LNOB including people living with disabilities)
- review how to reduce the length of the report to bring it closer to guidance (with further use of the annexes)

SECTION RATINGS		
SECTION A: EXECUTIVE SUMMARY (weight 5%)	83%	Comments on Rating
Question 1. Can the executive summary inform decision-making?		
i Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Partially	The executive summary is clearly presented and flows logically. The content is reflective of the report as a whole, and it can be generally used as a standalone document. The ToRs state that the executive summary should be 4-5 pages in length but it exceeds this at 6 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.	Yes	The executive summary includes all expected components - the evaluation purpose, objectives, methodology, findings, lessons learned and recommendations.
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	All of the important information to understand the intervention and evaluation is present .
SECTION B: BACKGROUND (weight 5%)	100%	Comments on Rating
Question 2. Is the object of the evaluation clearly described?		
i Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Yes	The introduction clearly states that the evaluation is of UNICEF's Country Programme focused on WASH (Output 1 of Outcome 5). The budget is set out in 1.2.7 (US\$ 96.7 million)
ii Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cities, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Yes	The report discusses direct (government ministries, NGOs) and indirect beneficiaries. There is a stakeholder analysis which provides clarity around roles and responsibilities . Indirect beneficiaries (individuals using open defecation practices and substandard sanitation practices) including vulnerable groups, are also mentioned. Numbers are aggregated (10 million involved in open defecation, 43 million in substandard sanitation), but this seems reasonable given that it is a national programme (and rightsholders are indirect beneficiaries) .
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 context provided is that of WASH global issues, and then with a focus on Pakistan. The context is focused on access to clean water and sanitation (the scale of the issue, policy framework, governance, international commitments). It also provides a solid overview of the country in terms of socio-economic, political and institutional factors .
ii Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	There is sufficient reference to SDGs and targets (6.1, 6.2, 4.3).
iii Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	Whilst the direct beneficiaries (government, decision-makers, NGOs)' s status and needs are not discussed, it is sufficient that the rationale for the WEE work in terms of what is needed to improve the WASH EE (section 1.3) is included. The WASH landscape and numbers that it affects are discussed regarding indirect beneficiaries.
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 key stakeholders are set out clearly within Table 4 and include government departments, local government, sector partners, academia and others.
ii Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	The roles and responsibilities for each of the stakeholders is set out clearly in Table 4.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)	92%	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 overall purpose of the evaluation is set out clearly as per the ToRs 'enhance accountability through an independent assessment of attained (or unattained) results and the underlying factors that influenced these outcomes' , chiming with the summative nature of the evaluation . Section 2. 2 sets out very well the key stakeholders, their roles in the evaluation and uses. Within this it states that the evaluation will be used to inform any adjustments in the next country programme document.

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 eight objectives are listed, and the report states that there has been some rephrasing. On review, the essence remains the same.
	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 scope of the evaluation sets out the geographical coverage, timeframe, and thematic coverage, presented well in figure 6. However, it could be helpful to include the month for the start and end (2018 and 2022) to give an accurate picture of the timeframe under review.
Question 7.	Is the theory of change, results chain or logic well articulated?		
	i Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Yes	Section 2.4 shares and explains 'Output 1's theory of change at length, with discussion of the 'if.. then' theory of change (ToC) logic and description of each of the key components .
	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 between outputs and outcomes are shared well in section 2.4 both graphically (figure 7) and also within a description sharing the assumptions that are linked to the causal pathways. Annex 3 also provides more depth.
	iii For theory-based evaluations, the theory of change or results framework is assessed, and if requested in the ToR, it is reformulated/improved by the evaluators.	Yes	The theory of change is assessed in terms of its comprehensiveness and validity. The evaluation scope did not extend to reconstructing the ToC and conceptual framework.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	82%	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 criteria are shared and justified, and comprise each of the six OECD-DAC criteria as requested in the ToRs and an additional cross-cutting criteria (gender, equity, human rights, environmental sustainability). The overarching evaluation questions for each criteria are shared in a visual (figure 9) and questions/ sub questions are included in table 6 (which does not include the overarching evaluation questions). The questions respond to both the accountability and learning objectives of the evaluation, and are well aligned to the evaluation criteria. 'Cross-cutting' criteria of gender equality, equity, climate change and human rights are included within two questions (relevance, impact) and there is also a focused criteria with three further questions . It would be helpful to include the overarching evaluation question for each criteria within the evaluation table (table 6) and evaluation matrix (annex 4). Including the question for each of the criteria would also help to guide the analysis and narrative. Within the findings section, there is a lot of cross referral to other evaluation questions/ sub questions (or dealing with multiple questions together) suggesting question overlap. A consultative process was taken during the inception phase to finalise the questions.
	ii In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions and/or other processes from which the analysis can be based and conclusions drawn.	Partially	A comprehensive evaluation matrix is provided in Annex 4 , including indicators, data collection methods, data analysis, data sources. This is a useful planning tool. However, it would be useful to include the overarching evaluation question for each criteria.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?		
	i Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	The evaluation design and set of methods are fit-for-purpose and follow current good practice, using a participatory, theory-based and utilisation focused approaches (and each is explained well for the context of this evaluation). A mixed methods approach is utilised, but with quantitative data from secondary sources and an emphasis only on qualitative data within primary data collection. This was given the focus of in-depth and holistic understanding from the qualitative approach. Each of the methods is well described , with articulation of why they were selected and how.

	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 (set out in section 3.5) are appropriate. They include an initial desk review to establish comprehensive understanding of the context, and map the different initiatives. Data sources included quantitative data , notably UNICEF's results management system to gauge the impact of the WEE activities. Key informant interviews were pivotal to the evaluation to gather feedback. It clearly (and justifiably) states that community members were not included as they are not direct beneficiaries of 'Output 1'. The data sources are clearly described in terms of both why they were selected and how they were utilised.
	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Yes	The sampling strategy (section 3.5.4) used purposeful (or purposive) sampling so that information-rich sources were sought, which makes sense in the context of the evaluation. A table sets out the number of participants across different stakeholders (disaggregated by sex) which shows that appropriate stakeholders were spoken to.
	iv	Clear and complete description of the methods of analysis.	Yes	The methods of analysis are explained (3.7) and included thematic analysis and content analysis. It appears methodical, systematic and sufficiently exploratory. Different forms of triangulation (section 3.8) are also shared, with helpful explanation as to how they reduce the risk of biased conclusions .
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The methodology enabled the tracing of causal connections, as shared in (3.3.2) the description of the theory of change with articulation of the mapping of the desired changes and processes / factors to enable or hinder the changes. The data collected is also appropriate.
	vi	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Yes	The limitations and constraints are grounded in the reality of the evaluation process, and the mitigation measures taken were plausible to reduce/ minimise impact of the limitations.
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 ethical considerations (section 3.12) include reference to UNEG norms and standards and UNICEF procedures, and the principles are outlined and explained as to how they were applied within the evaluation.
	ii	Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.	Yes	Ethical safeguards were in place including informed consent, avoidance of harm , privacy for participants . As the evaluation did not interview community members, children or adolescents no specific safeguards were needed.
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.	No	The evaluation did not incorporate any innovative practice (but used tried-and-tested) methods.
SECTION E:	EVALUATION FINDINGS (weight 25%)		83%	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	Each of the evaluation questions and sub-questions is systematically responded. The analysis is evidence-based and clearly states the types of evidence that are used to formulate the findings throughout the report. However within the narrative some are combined e.g. efficiency and cross-cutting criteria, to avoid duplications, and others refer back to different evaluation questions to answer the questions which is not ideal.

	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	Section 4.4.2 provides a commentary on the ToC, setting out the strengths and weakness in detail. It is explicitly referred to , particularly around questions on effectiveness.
Question 13.		Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		
	i	Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.	Yes	The evaluation uses and refers to multiple data sources e.g. listing the main desk review sources that have led to particular findings. Interviews are also cited, with information about different stakeholder types. Triangulation is evident throughout.
	ii	Findings are clearly supported by, and respond to, the evidence presented, including both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.	Partially	The findings present strengths and limitations of the WASH interventions, including of unintended results. Some benchmarks are included (e.g. 4.1.2) but more global standards/ indicators could have been referred to and used (e.g. regarding open defecation, access to clean water, utilisation rates) .
	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 causal factors are identified , highlighting strengths and limitations of evaluation results. The enabling and constraining factors (technical, policy, expertise, legislative, governance), influencing results are also included.
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	There is a specific question (EQ10) which asks about the M&E framework , and assesses its credibility, reliability and accuracy to measure the outputs, outcomes and impact of the WASH strategy. The lack of an overall M&E framework is highlighted and the tiered monitoring structures and collaborative efforts are also noted.
SECTION F:		EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	88%	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 provide assessment of the WASH outcomes and outputs , both for accountability and learning objectives. They are organised around the evaluation criteria, responding to the evaluation objectives.
	ii	Conclusions are derived appropriately from findings, and present a picture of the strengths and limitations of the intervention that adds insight and analysis beyond the findings.	Partially	The conclusions summarise the findings well, and present a balanced synopsis of strengths and limitations. Whilst the concise text crystallises the findings, the conclusions are not more strategic to add additional strategic insight beyond the findings.
Question 16.		Are logical and informative lessons learned identified? [N/A if lessons are not presented and not requested in ToR]		
	i	Identified lessons stem logically from the findings, have wider applicability and relevance beyond the object of the evaluation.	Yes	The lessons are derived well from the findings, and are generally focused on lessons to optimise the efficacy and effectiveness of scaling up. They have wider applicability beyond the Pakistan WASH context.
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	The lessons are written to be more broadly relevant, but have enough detail to be useful for the key users of the report.
SECTION G:		RECOMMENDATIONS (weight 15%)	88%	Comments on Rating
Question 17.		Are recommendations well grounded in the evaluation?		
	i	Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Yes	The recommendations align with the evaluation purpose and are logically derived from the findings and conclusions. It would have been helpful to include a mapping of findings, conclusions and recommendations so that they could be traced.
	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 strategic and focus on the systems, policy / legislative framework and governance of the WASH sector. Recommended actions are provided for each of the recommendation areas.
	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).	Yes	The draft recommendations were presented to the ERG for their validation, and some refinements were made.
Question 18.		Are recommendations clearly presented?		

	i Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Partially	The recommendations are prioritised and timeframes provided. Primary and secondary duty bearers are identified, however some of these are multiple and it would be useful to assign one lead organisation. If this is not possible for all of the recommendations, the suggested actions could be targeted to different stakeholders.
SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)		79%	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.	Yes	The opening pages include the name of the evaluated object, timeframe, commissioning organisation, evaluation organisation and team, table of contents, glossary and list of acronyms.
	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	Annexes include terms of reference, evaluation matrix, detailed methodology, data collection instruments, list of interviewees, list of references, mapping of initiatives, theory of change, extended findings sections.
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 structure is highly navigable with clear headings and sub-headings and numbered sections.
	ii Structure accords to UNICEF guidelines for evaluation reports: context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations.	Yes	The structure adheres to UNICEF guidelines for evaluation reports, with each of the key components in the appropriate sequence.
Question 21.	Is the report well presented?		
	i Report is of reasonable length; it does not exceed number of pages that may be specified in ToR.	No	The report is extremely long at 100 pages (excluding the executive summary) and far exceeds the guidance (up to 70 pages).
	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 and engaging way. The narrative is well structured and written, with few grammar, spelling and punctuation errors.
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Partially	The visuals are notable in sharing information in an accessible and visually appealing way e.g. figure 9 of the evaluation matrix, There are 15 figures and 14 tables. Photos are also used to add interest to the report. The links from the table of figures may need checking (e.g. figure 4 does not work).
SECTION I: EVALUATION PRINCIPLES (weight 10%)		93%	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 WASH strategy is assessed against its alignment with CRC and CEDAW. Human rights is one of the cross-cutting issues brought into evaluation design, findings analysis and recommendations.
	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	There is clarity about the types and numbers of evaluation respondents. An evaluation reference group was established by UNICEF.
	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 within the report does not indicate any bias. There is differentiated analysis for different groups.
	iv Evaluation assesses the extent to which the implementation of the intervention addresses child rights and Leave No-one Behind (gender and other excluded and marginalized groups). It is disability inclusive.	Partially	The evaluation uses cross-cutting criteria (gender equality, equity including disability, climate change) within evaluation design and findings. The impact on different marginalised groups is discussed broadly rather than tested. The conclusions discuss the LNOB principles / systems that have been brought in to interventions; the recommendations focus more on gender rather than other marginalised groups.
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)	9	

	<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>	Fully integrated	<p>The report is notable in that the evaluation 'lens' encompasses gender equality, equity, human rights, and environmental sustainability as distinct and integral areas to assess (within a 'cross-cutting' criterion) within the evaluation matrix. Leaving no one behind including GEWE is also considered within two other evaluation criteria (relevance, impact).</p>
	<p>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</p>	Fully integrated	<p>A gender responsive methodology is included within the methodology, methods/ tools, planning (female team members and female only interviews for female respondents) and analysis. Table 8 helpfully articulates the different evaluation stages and how GEWE/ human rights measure were incorporated.</p>
	<p>iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</p>	Fully integrated	<p>The findings , conclusions and recommendations reflect a gender analysis</p>

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>