

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Summative evaluation of the ECHO-supported WASH programme in Afghanistan

REPORT RATING SUMMARY			
Overall Rating		87%	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 the ECHO-supported WASH programme in Afghanistan	
Report sequence number		Afghanistan/64/2024/19993	
Region		SAR	
Year of report		2024	
Office		UNICEF Afghanistan Country Office	
Coverage (countries)		Afghanistan	
ToRs present		Yes	
Date of review (dd/mmm/yyyy)		December 8, 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		No	
Gender equality (cross-cutting)		Yes	
Humanitarian action (cross-cutting)		Yes	
Evaluation object		Programme	
Evaluation type		Summative	
Evaluation strategy		Mixed methods	
Evaluation design (primary method used)		Theory-based	
Evaluation level		Output & Outcome	
Geographic scope		National	
Primary SDG(s) covered (number)		6.1, 6.2	
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.			
The information presented in the report is comprehensive and it is clear that this has been a large evaluation that has managed to provide a lot of useful information. The report has a number of strengths: •The Executive Summary is clear, tidy and succinct and it serves well as a standalone document useful for informing decision making. It is of relevant conciseness and depth for key users, presenting all main findings, conclusions and summary recommendations in a clear and tidy manner. •Clear linkages are drawn to the SDGs and relevant targets and indicators for the areas being evaluated. •The evaluation's purpose and objectives are clearly presented and the evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. •The evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope. They are fully and clearly described. The methodology allows for drawing causal connections between outputs and expected outcomes. •The findings are presented with appropriate evidence to back them up. •The conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. •The recommendations align with the evaluation purpose. They are clearly formulated and they are useful and actionable for the primary intended users and uses. •Clear linkages are drawn to the SDGs and the relevant targets and indicators that are being evaluated (i.e. SDG 6.1 and 6.2). It is stated that Afghanistan's progress towards the Sustainable Development Goals (SDGs) is stagnating or decreasing for almost all goals, and that it is not on track to achieving SDGs 6.1 and 6.2. The report had some weaknesses, however: •There is no reference to- and use of relevant rights-based frameworks such as CRC, CCC, or CEDAW, although CRC appears in the list of acronyms.			
Recommendations for Improvement: The rater will identify topline recommendations to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, resources will be cited to assist evaluation managers in overseeing future evaluations.			
The following are recommendations for improvement: •The report would benefit from a clearer reference to and use of relevant rights-based frameworks such as CRC, CCC or CEDAW.			
SECTION RATINGS			
SECTION A:	EXECUTIVE SUMMARY (weight 5%)	100%	Comments on Rating
Question 1.	Can the executive summary inform decision-making?		
i	Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Yes	The Executive Summary is clearly presented, serves as a standalone document useful for informing decision making. It is of relevant conciseness and depth for key uses and does not exceed 5 pages.
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 necessary elements, including an overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions and findings, lessons learned and key recommendations, as per 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 includes all significant information needed to understand the intervention and the evaluation and it does not introduce new information that was not presented in the main part of the report.
SECTION B: BACKGROUND (weight 5%)			93%	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	There is a clear and relevant description of the intervention, including locations, timelines, overall cost and implementation status.
	ii	Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cites, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Yes	There is a clear and relevant description of the intended rightsholders and duty bearers by type, geographic locations in terms of numbers reached, and where possible, disaggregated by gender, age and disability. A stakeholder analysis is presented in Table 4 on page 6.
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	A thorough, clear and relevant description of the context of the intervention is presented in the Background section of the report. It includes relevant policy, socioeconomic, political, cultural, power/privilege, institutional and international factors, and it also provides a clear outline of how the intervention relates to the context. For example, the highly volatile political situation with the de facto Government is described, including the restraints on women and girls and their rights, including the WASH-related constraints in the country, justifying the need for this intervention.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	Clear linkages are drawn to the SDGs and the relevant targets and indicators that are being evaluation (i.e. SDG 6.1 and 6.2). It is stated that Afghanistan's progress towards the Sustainable Development Goals (SDGs) is stagnating or decreasing for almost all goals, and that it is not on track to achieving SDGs 6.1 and 6.2.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	There is a clear and relevant description of the status and needs of the rightsholders of the intervention. For example, The latest data from the 2022-23 Multiple Indicator Cluster Survey (MICS) survey shows that 69% of people have access to at least basic drinking water with a clear urban rural disparity (93% in urban areas; 61% in rural areas). It furthermore indicates that 20% of Afghans practice open defecation (4.2 million people) with high disparities between rural (26%), and urban areas (1%) and for the semi-nomadic "Kuchi" population (4.5% of the Afghan population) with high open defecation practices (71%). Handwashing is limited with nearly half (42%) lacking access to basic facilities. 74% of health facilities have available sanitation facilities. However, there is considerable regional variation with much lower availability in facilities in western (53%) and northern (59%) regions. Access to WASH facilities in schools is limited, there is only 1 female toilet for every 200 male toilets in schools.
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	There is clear identification of the implementing agencies, development partners, rights holders and additional duty bearers and other stakeholders, including the linkages between them in the stakeholder map presented in Table 4 on page 6.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	The report identifies the specific contributions and roles of key stakeholders, including UNICEF, in both financial and non-financial terms. However, it does not provide a clear overview of the financial contributions made by each donor.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)			83%	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 clearly defined, including why it was needed at that point in time, its intended use and its key intended users. The evaluation has a dual purpose: Firstly, to assess the ECHO-supported WASH Programme's overall design, strategies, interventions, and results to guide WASH programme planning and aid decision-making at the country office level to deliver high-quality WASH programmes in advancing children's rights, especially the most vulnerable populations in Afghanistan. Secondly, to promote and foster organizational learning and improvement by identifying key strengths, gaps, bottlenecks, and satisfaction levels among affected populations, thereby facilitating targeted interventions and programme enhancements.
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	There is a 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. The evaluation objectives as agreed during the inception phase are clearly laid out in section 3.1 on page 7.

	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	There is a description of the scope of the evaluation, thematic, temporal and geographic in table 5 on page 8. What will not be covered is not specified, neither are the reasons for the scope.
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	There is a clear description of the intervention's intended results that are being tested by the evaluation. A graphic presentation of the ToR can be found on page 5.
	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 between outputs and outcomes are well presented in graphic form in the Theory of Change on page 5.
	iii	For theory-based evaluations, the theory of change or results framework is assessed, and if requested in the ToR, it is reformulated/improved by the evaluators.	Partially	The Theory of Change was constructed by the evaluation team retrospectively, as none existed prior. The ToC was as such not assessed by the evaluators.
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.	Yes	The evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The evaluation questions are organised under evaluation criteria such as Relevance, Coherence, Effectiveness, Efficiency, Sustainability, Coverage, Coordination and Cross-cutting patterns, as outlined in Table 6 on pages 8-9.
	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 (Appendix 6, pages 17-36) includes the evaluation questions and sub-questions, areas of inquiry/potential indicators, data collection methods, data sources and analysis type.
Question 9. Does the report specify adequate methods for data collection, analysis, and sampling?				
	i	Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	The evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described. The evaluation used hybrid approaches with mixed methods, combining theory-based design, contribution analysis and post-intervention single group research. The Evaluation Team constructed the ToC using the descriptive logic of interventions from the programme documents. Using the Theory of Change (ToC), the evaluation team (ET) mapped pathways of change and analyzed programme inputs and activities. By comparing post-intervention data with pre-intervention conditions, the evaluation assessed changes in WASH services.
	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 thoroughly described, both in narrative and in a table in the main report and in more detail in Appendixes 11, 12 and 13. They include both qualitative and quantitative sources including Document Reviews, Key Informant Interviews, Focal Group Discussions, Household Surveys and Observations.
	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 is provided. It includes a description of how diverse perspectives are captured. For quantitative data, due to the complexity of the WASH interventions geographically and programmatically, a complex multi-stage cluster sampling strategy was utilised. In the absence of a reliable and comprehensive national level household listing or a complete programme beneficiary list, a cluster sampling approach was considered one of the most effective methods to draw a statistically representative sample while taking the logistical and budgetary limitation in consideration. For qualitative data, a purposive or 'judgmental' sampling for qualitative data collection was applied, to 1) ensure selection of relevant categories of stakeholders at all levels; 2) include a diverse range of locations and communities; and 3) reach communities that received the maximum possible benefits from the programme to maximize representation of programme participants.

	iv Clear and complete description of the methods of analysis.	Yes	There is a clear and complete description of the methods of analysis. For the qualitative data, the evaluation team undertook a systematic content and thematic analysis of secondary data, FGDs, KIs and field notes/recordings using Excel-based tools. The process involved reviewing and summarizing unstructured textual content, followed by identifying and coding relevant data into common themes. Summarized data was organized into matrices to extract patterns on programmatic aspects and participants' perspectives. For quantitative data collected through household surveys and WASH Checklists (for schools, health facilities, and water supply scheme) underwent a thorough analysis using analysis plans formulated during the data collection phase, guided the overall analysis, including descriptive statistics (averages, means, frequencies), and cross-tabulations for gender analysis, equity, disability, human rights, and climate considerations related variables where feasible and applicable.
	v Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The methodology chosen, which is theory based, allows well for drawing causal connections between the outputs and expected outcomes. These are clearly drawn up in the Theory of Change and they are also used when presenting the evaluation findings.
	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	There is a clear and complete description of the evaluation's limitations faced by the evaluation and the corresponding mitigation measures taken by the evaluation in table 9 on pages 14-15.
Question 10.	Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
	i Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Partially	There is explicit reference to the UNEG obligations of evaluators, but these are not specified in the report (i.e., independence, impartiality, credibility, conflicts of interest and accountability). The evaluation design was approved through the institutional review board (IRB) process prior to implementation. The evaluation design and execution adhered to UNICEF evaluation policy, UNICEF Standard on Information Security (2018), UNICEF Policy on Personal Data Protection (2020), UNICEF procedure on ethical standards in research, evaluation, data collection and analysis (2021), and UNEG and UNICEF prescribed evaluation norms, and standards among several other guidance documents on evaluation ethical guidelines (also outlined in Appendix 14).
	ii Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.	Partially	The report mentions that data collection at the community-level was conducted with same-gender enumerators i.e. men interviewed men and women interviewed women. Special attention was given to respecting the sensitivities of topics (menstrual hygiene management (MHM) and personal hygiene), especially during discussions with women as FGD participants. However, no further description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they were applied is provided (i.e. regarding respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm).
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	Innovation practice was not used in this evaluation.
SECTION E:	EVALUATION FINDINGS (weight 25%)	100%	Comments on Rating
Question 12.	Do the findings clearly address all evaluation objectives and scope?		
	i Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Yes	The evaluation findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria. Evidence sources are clearly presented and quoted (household survey results, interview content, etc) to back up findings and conclusions.

	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	There is explicit use of the intervention's Theory of Change in the formulation of the findings. For example, it is evident in the summary part of the findings presented under Effectiveness, where the links between activities, outputs and programme outcomes become clear: The programme demonstrated substantial progress in achieving its stated objectives by improving access to clean and safe drinking water for vulnerable populations in disaster prone areas and some newly accessible districts. Through the construction and rehabilitation of water systems, the programme increased communities' reliable access to improved water sources. Providing water systems at both household and community levels reduced reliance on unimproved water sources, such as surface water and water trucking. Furthermore, the programme integrated its interventions with health and nutrition sectors, addressing root causes of issues like diarrhoea and malnutrition, while also strengthening groundwater monitoring systems to address drought and natural disasters effects in the long run. The programme was mostly successful in reaching marginalized and previously inaccessible regions, with nearly half of activities implemented in newly accessible districts (35 of 83). Expanding service to these areas provided essential WASH services to some of the most underserved populations. This achievement highlights the programme's success in extending its reach to the most vulnerable groups, in alignment with its objectives.
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 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. This can be seen, for example, in paragraph 106 on page 26 where findings are presented based on diverse data sources such as the household survey and interviews.
	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	The findings are clearly supported by, and respond to, the evidence presented, including both positive and negative. Findings are based on clear performance indicators or other means of comparison as relevant for each question. For example, it is stated that "the hygiene promotion interventions, aimed to foster behavioural change and improve hygiene practices in the targeted provinces across Afghanistan. The programme educated communities on key practices such as handwashing and proper waste disposal to prevent waterborne diseases, and the distribution of hygiene kits during emergencies further supported these practices. Stakeholders identified gaps, including partial integration of HP within all projects, selective targeting, inconsistencies of content and dissemination of messages by different entities, and weak technical capacities of the hygiene promoters on the ground, which affected the effectiveness and outcomes of these efforts."
	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	Causal factors leading to achievement or non-achievement of results are clearly identified, and the findings also analyse the logical chain. For example, it is stated that " These gaps (the gaps described in 91D above) were mainly related to the fact that different local IPs and INGOs were implementing different types of interventions (6 different types of water supply projects, water trucking, and sanitation activities, etc.) and hence every implementer adopted different approaches to integrating hygiene promotion within these activities."
Question 14.	Does the evaluation assess and use the intervention's Results Based Management elements?		
	i Assessment of the adequacy of the intervention's monitoring system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	Yes	The evaluation included an assessment of the adequacy of the intervention's monitoring system to support decision making. It assessed the quality and overall coherence of internal monitoring and evaluation (M&E) and knowledge management systems utilized by the programme.
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 are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking. For example, the conclusion for Effectiveness states that the programme made notable progress in improving water access by constructing and rehabilitating systems. Key achievements include transitioning from water trucking to more sustainable sources and delivering water to vulnerable communities, including those in previously inaccessible areas. The programme's emergency response during crises, such as the Paktika and Herat earthquakes, strengthened from largely effective to fully effective respectively. Efforts also focused on groundwater monitoring and promoting integrated water resource management. The programme integrated WASH with health and nutrition sectors, thereby contributing to efforts in addressing root causes of polio, AWD and malnutrition. The Programme's success was driven by key enabling factors including flexible funding allocation, effective crisis modifier utilization, and strategic prepositioning of supplies. Strong partnerships, coordination with the DfA, and community involvement, coupled with effective coordination mechanisms and adaptability, ensured that WASH services were delivered effectively despite evolving challenges. The programme's effectiveness and implementation were also challenged by factors like policy and donor misalignment, bureaucratic delays, and financial management issues. These challenges, along with logistical, security, and environmental constraints, caused delays, increased costs, and limited the programme's overall reach and effectiveness.
	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 seem to be derived appropriately from the findings, but there is no reference to which findings or evaluation questions each of the conclusions is based on. The conclusions present a picture of the strengths and limitations of the intervention that adds insights and analysis beyond the findings. For example, under Efficiency, it is stated that the programme demonstrated significant strengths in terms of human resources, funding, and time management, though challenges in each area impacted overall efficiency. And furthermore that UNICEF's WASH section maintained adequate human resource capacity, but gaps in DfA sectoral departments and high turnover among IP and TPM entities affected project monitoring and consistency.
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 presented on pages 52-53 stem logically from the findings (although do not refer directly to the findings they stem from), and have wider applicability and relevance beyond the object of the evaluation.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	The lessons are clearly and concisely presented and they have sufficient detail to be useful for the intended audience. For example, it is stated that coordinating with the De facto Authorities, despite restrictions on financial transactions, is essential for areas such as situation analysis. Active engagement with DfA entities was important in facilitating their contribution to programming efforts within the ECHO-supported WASH Programme. This lesson highlights the importance of engagement with DfA entities, even in contexts of international sanctions, and can be applied to other programmes in similar contexts.
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.	Partially	The recommendations seem to align generally with the evaluation purpose, but there are no references to which evaluation questions or evaluation criteria each recommendation relates to, and hence it is unclear to what degree the recommendations cover all relevant findings or not. The recommendations are clearly formulated and are to some extent logically derived from the findings and conclusions.
	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 for the primary intended users and uses and appropriate guidance is given for implementation. Each recommendation has several sub-points further elaborating on necessary action.
	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 process for developing the recommendations is described, and it includes the involvement of both duty bearers and rights holders. The recommendations were drawn from triangulating key suggestions provided by respondents during primary data collection, considering operational and contextual challenges in Afghanistan, and corroborating with programme documents. Dedicated follow-up discussions on recommendations were held with evaluation reference group members, senior officials from UNICEF, implementing partners, and Third Party Monitoring representatives to refine the 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.	Yes	There is a clear identification of the groups of duty bearers responsible for action for each recommendation, and a clear prioritization (timewise) of the recommendations to support use.
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.	Yes	The opening pages include the name of the evaluated object, timeframe of the evaluation, date of the report, location of the evaluated object, names and organizations of the evaluators, and name of the organization commissioning the evaluation, in addition to table of contents, including tables and figures (but not annexes; they come in a separate volume), list of acronyms and glossary, and with appropriate page numbering.
	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.	Partially	The annexes include the terms of reference, an evaluation matrix, list of key informants, data collection instruments, list of documentary evidence - in addition to other relevant annexes (26 in total). No list of site visits is included in 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 structure is easy to identify and navigate with numbered sections, clear titles and sub-titles. The report is well formatted.
	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	The report structure partially accords to UNICEF guidelines for evaluation reports, the context, purpose and methodology precede the findings, which are followed by lessons learned, conclusions and recommendations. Normally, the lessons learned would follow the conclusions and not the other way around.
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 of a reasonable length with its 62 pages and does not exceed the page number limit.
	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 easy to understand, written in an accessible way for the intended audience and is generally free from grammar, spelling and punctuation errors.
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Yes	There is a frequent use of visual aids to convey key information, with 16 tables and 10 figures.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	55%	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	There is no reference to rights-based frameworks and no mention of CRC, CCC or CEDAW or other rights related benchmarks in the design of 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	There is a clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation. A reference group was established.
	iii Language is empowering and inclusive, avoiding gender, heterosexual, age, cultural and religious bias, among others; use terminology of rights holders and duty bearers; data is disaggregated by marginalized group; differential results are assessed (distribution of results across different groups).	Partially	The language in the report is generally empowering and inclusive, avoiding gender, heterosexual, age, cultural and religious bias. There is very limited use of the terminology of rights holders (only three times throughout the report) and no mention of duty bearers. Differential results are assessed whenever possible. It is reported that due to the political situation and the lack of statistics, it was challenging to disaggregate some data for gender and disability.
	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.	No	The evaluation does not refer to child rights (except from once in the formulation of one of the evaluation questions), and does not assess the extent to which the implementation Leaves No-one Behind. It has the intention of being disability inclusive, but is hampered by the limited access to both disability statistics and to persons with disabilities in Afghanistan.
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.	Fully integrated	The evaluation assesses whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results. The evaluation included an objective specific to assessment of human rights and gender equality considerations. Gender and human rights were mainstreamed into the various evaluation criteria. There was a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation.

ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	The evaluation specified how gender issues were addressed in the methodology, including: how data collection and analysis methods integrated gender considerations and ensured data collected is disaggregated by sex. The evaluation methodology employed a mixed-methods approach, appropriate to evaluating GEWE considerations. A diverse range of data sources and processes were employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility. The evaluation methods and sampling frame addressed the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate. The report is somewhat vague regarding what ethical standards were applied to this evaluation.
iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Satisfactorily integrated	The evaluation has a background section that includes an intersectional analysis of the specific social groups affected by the issue and it outlines the relevant normative instruments and policies related to human rights and gender equality in Afghanistan. The findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and disaggregates quantitative data, where applicable. Unanticipated effects of the intervention on human rights and gender equality are not described. The evaluation report provides specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area.

SWAP Rating Guidance

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