

GEROS Evaluation Quality Assurance Tool

Version: November, 2020

Summative Evaluation of the Enhanced and Accelerated WASH Programmes in Ghana - ACCELERATED Programme

REPORT RATING SUMMARY			
Overall Rating		81%	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		Summative Evaluation of the Enhanced and Accelerated WASH Programmes in Ghana - ACCELERATED Programme	
Report sequence number		Ghana/61/2020/16877	
Region		WCAR	
Year of report		2020	
Office		UNICEF Ghana Country Office	
Coverage (countries)		Ghana	
ToRs present		Yes	
Date of review (dd/mmm/yyyy)		April 16, 2021	
Name of review firm		DeftEdge	
CLASSIFICATION OF EVALUATION REPORT			
Management of evaluation (<i>Managerial control and oversight of evaluation</i>)		UNICEF managed	
Unicef goal areas (<i>Alignment with strategic plan priorities</i>)			
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)		Yes	
Humanitarian action (cross-cutting)		No	
Evaluation object		Programme	
Evaluation type		Summative	
Evaluation strategy		Mixed methods	
Evaluation design (<i>primary method used</i>)		Theory-based	
Evaluation level		Output & Outcome	
Geographic scope		National	
Primary SDG(s) covered (<i>number</i>)		5, 6	
coverage (countries)			
The report generally meets the requirements of UNICEF evaluations. It is a national office evaluation of a UNICEF WASH programme in Ghana. The intervention was intended to improve the health and well-being of children in schools and of women and men in communities across the Northern region of Ghana where access to WASH-related services and health outcomes are the lowest in the country. The evaluation objective was to establish whether the programme contributed to improved sanitation and hygiene with a view to promoting learning and accountability. As a theory-based evaluation, the evaluators used a mixed method approach that combined qualitative and quantitative data with a focus on reviewing and testing the Theory of Change. The assessment of the three OECD / DAC criteria (relevance, effectiveness, and efficiency) was a strength, as was the stakeholder mapping presentation. However, at 85 pages, the detailed section of findings was overly long, making it difficult for the reader to maintain focus, and contributing to the report being much longer than the requested 40 pages. The use of preliminary and final conclusions in separate sections of the report also presented challenges in understanding the overall strengths and weaknesses of the programme. The evaluation is notable for its excellent gender analysis, which included a good discussion in the context section on the gendered aspects of WASH.			
Recommendations for Improvement:			
Although this was a thorough and detailed evaluation in most respects, there are several recommendations for improvement. During the design, the evaluation could have limited the number of Effectiveness EQs to make that section more concise and sought to ensure the endline survey data collected to compare against baseline characteristics was fit-for-purpose (more aligned). With respect to the report, it would be helpful to have more detail on the scale and complexity of the intervention in the background section, to have more used more graphs and figures to more efficiently/concise present EQ findings, and to have consolidated the preliminary and final conclusions in one cohesive discussion. The evaluation could also have considered the new DAC Network on Development Evaluation criterion of Coherence that is intended to focus attention on the synergies (or lack thereof) between policy areas and on cross-government co-ordination.			
SECTION RATINGS			
SECTION A:	EXECUTIVE SUMMARY (weight 5%)	67%	Comments on Rating (include explanations for any criterion not rated)
Question 1.	Can the executive summary inform decision-making?		
	i Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Partially	The Executive Summary serves as a useful stand alone document though it is over the 5 page maximum.
	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	All of the necessary elements are present.

	iii	Includes all significant information needed to understand the intervention and the evaluation AND does not introduce new information from what is presented in the rest of the report.	Partially	Most of the information needed to understand the intervention and the evaluation is provided, though the executive summary could have included more background on the context of the programme, the duty bearers and the rights holders. Economies could have been made in the 3 pages summary of findings and 2 pages of lessons learned and recommendations.
SECTION B: BACKGROUND (weight 5%)			93%	Comments on Rating (include explanations for any criterion not rated)
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	There is a description of the intervention, though the total amount of UNICEF's contribution is not provided.
	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	There is a clear and relevant description of the intended rights holders (residents and school children in small towns in the Northern region) and duty bearers (the School Health Education Programme (SHEP) of the Ghana Education Service; the Environmental Health and Sanitation Directorate (EHSD); the Department of Community Development; Community Water and Sanitation Agency (CWSA); and, the Metropolitan, Municipal and District Assemblies (MMDAs).
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	Relevant context was carefully described in Annex 6.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	Linkages are drawn to the MDGs and SDGs under Relevance. Additional information is provided in Annex 6.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	The description of the status and needs of rightsholders is described.
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 implementing agencies identified included: as the School Health Education Programme (SHEP) of the Ghana Education Service; the Environmental Health and Sanitation Directorate (EHSD); the Department of Community Development; Community Water and Sanitation Agency (CWSA); and, the Metropolitan, Municipal and District Assemblies (MMDAs).
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	The specific contributions and roles of key stakeholders were outlined in Annex 6. This contained a detailed mapping of key sector stakeholders and the wide range of programme stakeholders.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)			92%	Comments on Rating (include explanations for any criterion not rated)
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 evaluation is clearly defined, including why it was needed, its intended use (accountability, organisational learning, targeted learning), and the key intended users.
Question 6.		Are the objectives and scope of the evaluation clear and realistic?		
	i	Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Yes	The description of the evaluation objectives is clear and complete, as presented in the ToR.
	ii	Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Partially	The description of the evaluation scope does not completely align with the ToR. The ToR required a focus on the 6 components of the program, whereas the report states it considers results at the program level as a whole.
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	The description of the intended results is clear. The intended results were long-lasting improvements in the health and wellbeing of children in schools, and women and men in rural communities.
	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 is presented in the graphic form in the Revised Theory of Change.
	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 ToC was revised to better reflect program delivery. The assessment found the indicators were output focused. The evaluators reported they went beyond these indicators to assess results (outcomes and impacts).
SECTION D: EVALUATION DESIGN AND METHODOLOGY (weight 20%)			95%	Comments on Rating (include explanations for any criterion not rated)

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 was guided by three OECD / DAC evaluation criteria, including: relevance, effectiveness, and efficiency. The evaluation team explored whether impact and sustainability criteria could be covered (not required under the ToR) and considered it was too soon to find evidence of long-term outcomes and impacts set out in the ToC. Findings related to sustainability were identified and discussed in the section on Effectiveness. In terms of meeting the objectives and purpose of the evaluation, there was an excessive number of questions (22 under effectiveness) which made it a challenge to present cohesive findings. There also seemed to be overlap of Efficiency and Effectiveness questions as much of the Efficiency findings refer back to Effectiveness.
	ii In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions and/or other processes from which the analysis can be based and conclusions drawn.	Yes	The evaluation matrix is in Appendix 2. It includes the key evaluation questions, and indicators that could be measured and advises what data sources would be used to answer each question and draw conclusions.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?		
	i Evaluation design and set of methods is 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 clearly described. It drew on a range of data sources to ensure the reliability of results, promote impartiality, reduce bias, and ensure the evaluation was based on the most comprehensive and relevant information possible. Data sources included: internal document review, secondary data review, household survey and observation, interviews, and focus groups.
	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 sources of data were appropriate and included both quantitative and qualitative sources. Quantitative data sources were used to assess programme achievements, while qualitative data sources were used to triangulate quantitative data and explore factors that contributed to achievement / non achievement of results. One limitation reported concerned the comparison of baseline indicators and endline figures aligned with global indicators. The evaluation had no directly comparable baseline. To allow some level of comparability, the evaluators reported they designed questions in the endline survey to reflect the baseline.
	iii Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or if not, provide reasons for this).	Yes	A sampling strategy was provided for the household survey, key informant interviews, and to ensure gender and 'urbanity' issues were captured. There was discussion of how diverse perspectives were captured.
	iv Clear and complete description of the methods of analysis.	Yes	The description of data analysis is clear and discussed in section 1.7.7.
	v Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The methodology allows for drawing causal connections between outputs and expected outcomes, as evidenced by both analysis of the ToC and various questions to be answered under the effectiveness criterion.
	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 description of challenges / limitations and mitigation measures the evaluation team implemented.
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 obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability).	Yes	There were explicit and contextualised references to each obligation of evaluators provided in Annex 9 of the report.
	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	The description of ethical safeguards and how they were applied is carefully outlined at Annex 9. Explicit mention was made of UNICEF procedures for Ethical Research Involving Children. The Ghana Health Service Ethics Review Committee approved the study on 9th July 2019.
SECTION E:	EVALUATION FINDINGS (weight 25%)	75%	Comments on Rating (include explanations for any criterion not rated)
Question 11.	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 indicate that a high level of evidence was marshalled to address each of the questions by criteria, though there is some vagueness in citing sources, "Unicef data" and "many stakeholders". A populated Evaluation Matrix could have resolved this issue.
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Partially	Although done to some degree, the process to develop the Effectiveness findings could have more explicitly used the ToC.
Question 12.	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 used credible forms of qualitative and quantitative data and triangulation was evident. The main findings are supported by more than one data source.
	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 carefully prepared and correspond to the evidence. They are positive and negative, and the use of metrics and benchmarks are evident.
	iii Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings analyse the logical chain (progression -or not- from implementation to results).	Yes	The causal factors and reasons leading to achievement or non-achievement of results are clearly identified by EQ.
Question 13.	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 if required in ToR) to support decision-making.	Partially	The design issue of adequacy of the intervention's monitoring system was addressed on page 16, noting most of the targets set in the Key Performance Metrics (KPM) database were at the 'output' level. The evaluators noted they went beyond this data and assessed the results in terms of outcomes and impacts. There was no assessment of how effective project monitoring was, or usefulness of that data to inform programme implementation and refinement.
SECTION F: EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)		63%	Comments on Rating (include explanations for any criterion not rated)
Question 14.	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).	Partially	The final conclusions are formulated by each of the three OECD/DAC criteria and touch on many deep findings related to key elements of the overall programme. However, the conclusions would have been more cohesive if consolidated in one section, rather than also have conclusions within the findings sections.
	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	As a set, the preliminary and final conclusions are derived from the findings and there is a picture of the strengths and limitations. However, the report would have benefitted from more perspective on the strengths and limitations of the program in accord with the 6 themes identified in the ToR. For example, this may have led to learning on why health has not considerably improved in the region and hence, how specifically to improve the benefits of the program and enhance future programming.
Question 15.	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 learned stem from the findings and are applicable.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Partially	There are clear lessons that contain sufficient detail to inform future programme development, though several introductions read as findings and more could have been done to craft the lesson and present them concisely. For example, LL6 might be, it is critical to measure results if managers want to demonstrate where the program has achieved its objectives.
SECTION G: RECOMMENDATIONS (weight 15%)		88%	Comments on Rating (include explanations for any criterion not rated)
Question 16.	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 evaluation purpose and are 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.	Partially	The recommendations are mostly useful and actionable, however some of the specifics seem to be left for the reader to determine. For example, as regards Recommendation 1, how could central government provide a more supportive policy environment? What capacity building elements are more important to continue and sustain?
	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, noting duty-bearers and the donor were unable to attend on the day of the workshop.
Question 17.	Are recommendations clearly presented?		

	i Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Yes	The evaluation report identifies groups responsible for action of each recommendation and assigns a priority to the action.
SECTION H:	REPORT STRUCTURE AND PRESENTATION (weight 5%)	64%	Comments on Rating (include explanations for any criterion not rated)
They	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	Most of the required pages are included, though readability would have been enhanced if the Annexes were listed in the Table of Contents. This was more important due to reliance on use of annexes for relevant content on context and ethical considerations.
	ii Annexes include, if not in report body: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework, 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, results chain, information about the evaluator(s).	Partially	Most of the required Annexes are included, though a list of the 38 persons interviewed and sites visited could have been provided for transparency purposes.
Question 19.	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 main structure of the report is easy to identify and navigate.
	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 accords with UNICEF guidelines.
Question 20.	Is the report well presented?		
	i Report is of reasonable length; it does not exceed number of pages that may be specified in ToR.	Partially	At 85 pages (excluding the Executive Summary and annexes), the length of this report significantly exceeded the ToR requirement for a 40 page report.
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	The language was clear, however accessibility is somewhat impeded by the lengthiness of the findings and the conclusions being in different places.
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labeled, and referenced in text.	Partially	The evaluators made only limited use of visual aids to convey key information.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	86%	Comments on Rating (include explanations for any criterion not rated)
Question 21.	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, and/or CCC, and/or CEDAW and/or other rights related benchmarks in the design of the evaluation.	Partially	Although rights were fore fronted - a rights-based approach started from the premise that all people had the right to participate in and benefit from all development actions - there was not specific mention of a framework or convention.
	ii Clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation (for example, a reference group is established, stakeholders are involved as informants or in data gathering).	Partially	While there is a clear description of the level of participation of rightsholders in the evaluation, and the number of stakeholders involved as informants, there is no discussion of the evaluation reference group referred to in the ToR. There was mention of a technical review meeting in Accra, though this was only attended by UNICEF and the evaluation team.
	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 is empowering and inclusive, used rightsholder and duty bearer terminology, and it avoided bias.
	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	The evaluation assessed the programme targets achieved, with discussion on access of disabled groups to WASH. The report notes that during the Inception Report the evaluators sought to understand the context of WASH-related inequities in Ghana, including gendered division of household chores such as water collection, the specific hygiene and sanitation needs of women and girls, and inequities in access to WASH services faced by marginalised groups.
Question 22.	Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards)	9	
	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	Gender equality and human rights were fully integrated in the scope of analysis. The evaluators reported mainstreaming gender and incorporating a rights-based approach through the evaluation questions and selection of appropriate indicators.

ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Fully integrated	Gender-responsive methodology, methods, tools and data analysis techniques were applied. The evaluators reported adopting a mixed-methods approach to ensure both quantitative and qualitative evidence of equity considerations was sought and to ensure the voices of all stakeholder groups were heard. Inclusion of women and girls in the evaluation process was ensured through setting 50/50 male/female targets for the survey sample, and by holding gender-segregated focus group discussions for students. As an example, the survey data was disaggregated by key variable, gender and urbanity, to assess for any statistical differences between gender and settlement types.
iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	The findings, conclusions and recommendations reflect a gender analysis. There was a discussion of the gender-context in the background and a detailed gender analysis embedded in the findings. Recommendation 6 for future WASH programming was to incorporate a gender- and disability-inclusive approach in CLTS.

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.

- 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?
- Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?
- Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?
- 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.

- 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?
- Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?
- Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?
- Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?
- 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.

- 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?
- 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?
- Are unanticipated effects of the intervention on human rights and gender equality described?
- 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?