

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Summative Evaluation of the Simplified Birth Registration System in Mainland Tanzania

REPORT RATING SUMMARY			
Overall Rating	91%	Highly Satisfactory	
●●●●●	Exceptional (96% - 100%)	5	
●●●●●	Highly Satisfactory (87.5% - 95.99%)	4	Exceeds UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with a high degree of confidence
●●●●●	Satisfactory (62.5% - 87.49%)	3	
●●●●●	Fair (35% - 62.49%)	2	
●●●●●	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
Title of the evaluation report	Summative Evaluation of the Simplified Birth Registration System in Mainland Tanzania		
Report sequence number	United Rep. of Tanzania/63/2021/8934		
Region	ESAR		
Year of report	2021		
Office	UNICEF Tanzania Country Office		
Coverage (countries)	Tanzania		
ToRs present	Yes		
Date of review (dd/mm/yyyy)	January 30, 2022		
Name of review firm	DeftEdge		
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	No		
Every child has an equitable chance in life	Yes		
Gender equality (cross-cutting)	Yes		
Humanitarian action (cross-cutting)	No		
Evaluation object	Project		
Evaluation type	Summative		
Evaluation strategy	Mixed methods		
Evaluation design (primary method used)	Quasi-experimental		
Evaluation level	Output & Outcome		
Geographic scope	National		
Primary SDG(s) covered (number)	16		
EQA Summary:	<p>This is a highly satisfactory report that provides a comprehensive assessment of the Simplified Birth Registration System in Mainland Tanzania. The background section provides a thorough description of the relevant stakeholders and their roles, and the object of the evaluation is clearly described, though the context could be further elaborated on. The authors did well in describing the evaluation purpose, objectives, scope, and the intervention logic. The evaluation had a quasi-experimental design, collecting extensive qualitative and quantitative data from control and intervention groups. The sampling strategy is well described and ensures diverse stakeholders were reached. The data collection process and ethical guidelines are well-explained and professionally executed. The findings are well-validated, organized around the programme theory of change/results framework, and clearly respond to the evaluation questions. The lessons and recommendation sections are well-drafted and easy to comprehend, especially the recommendation section, which has a table where recommendations are categorized and prioritized. Child rights, gender equity and equality are an integral part of the evaluation.</p>		
Recommendations for Improvement:	<p>While the report is well written, the executive summary and overall report length are over the page limits specified in the ToR. In addition, there are many grammatical errors and issues with copy-editing that should be tended to. The description of the global, regional and country context is quite general and should more comprehensively analyze the relevant socio-economic, political, cultural and power-privilege dynamics in mainland Tanzania affecting the intervention. The evaluators state that the evaluation design was participatory, however, the extent of engagement of key stakeholders is not much more than standard non-experimental evaluations involving reference groups and interviewing / engaging duty bearers and rightsholders as key informants. It should be ensured that any approach or methodology noted should be clearly explained, and there should be evidence of its application in the evaluation report. Finally, it is recommended that conclusions are more thorough and provide a depth of analysis that moves beyond simply re-stating findings.</p>		
SECTION RATINGS			
SECTION A: EXECUTIVE SUMMARY (weight 5%)	50%	Comments on Rating	
Question 1.	Can the executive summary inform decision-making?		
i	No	At 8 pages, the executive summary exceeds the page limit of 3 pages specified in the ToR. The summary, while providing a sufficient level of detail to inform key users and decision-makers, is not well-edited as it has numerous spelling and grammatical errors and sometimes presents duplicate findings.	
ii	Yes	The summary provides a brief overview and context of the intervention, purpose, objectives and scope of the evaluation, evaluation design and methodology, followed by findings, conclusions, recommendations, and lessons. All of the essential elements are present in the summary document.	
iii	Partially	The summary document, though sometimes repetitive, is otherwise comprehensive and provides sufficient details to understand the intervention and the evaluation, in most cases. However, the frequent use of acronyms detracts from having a clear and full understanding of the intervention. The results of the evaluation are well covered and the section does not introduce new information.	
SECTION B: BACKGROUND (weight 5%)	93%	Comments on Rating	
Question 2.	Is the object of the evaluation clearly described?		
i	Yes	The background chapter is comprehensive in its coverage of key elements of the intervention under evaluation. Its sub-sections provide a clear and relevant description of the intervention from the country and global context, along with the programme objectives and expected results, timeline, implementation status, and programme resources allocated to each phase of implementation.	

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 neighborhoods, 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	The roles and types of the rightsholders (beneficiaries) and duty bearers (stakeholders) are carefully described in separate sections: programme stakeholders and programme participants. The roles of duty bearers are presented in Table 3. The proportion of children under 5 registered at birth is provided in Table 5 and disaggregated by region and gender.
Question 3.	Is the context of the intervention clearly described?		
i	Clear and relevant description of the context of the intervention (i.e. relevant policy, socio-economic, political, cultural, power/privilege, institutional, international factors) and how context relates to the implementation of the intervention.	Partially	The description of the global, regional and country context is quite general. The relevant socio-economic, political, cultural and power-privilege dynamics are only noted without much supportive statistical data or an analysis of trends across regions, socio-economic backgrounds, or other strata.
ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	SDG 16 and its target 16.9 are noted in several places and clearly linked to the object of the evaluation.
iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	A brief description of the status and needs of rightsholders of the intervention is provided, notably whether children are registered or not across regions.
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 primary duty bearers, development partners and rightsholders are clearly defined. Details on their roles, generally, and the extent of their engagement in the programme are provided in Table 3 and Appendix 3.
ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	The stakeholders and their roles are described well, along with the role of UNICEF, which is to support the Government of Tanzania in implementing new birth registration systems in regions of mainland Tanzania.
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	Section 2.1 notes that the evaluation will be summative in nature and generate evidence around the programme's achievements. It will also inform the roll-out of Phase 3, which is currently ongoing.
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 4 evaluation objectives are clearly specified. In addition, it is mentioned that objectives outlined in the ToR were reviewed and rephrased and were approved by the Evaluation Reference Group.
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).	Yes	The evaluation scope is clear and concise. There is a good description of the evaluation's thematic, chronological, and geographic scope. It is also highlighted that chronological scope was extended by few months from what was originally defined in the ToR.
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	A short description on the programme's theory of change is provided in section 1.4 along with a detailed graphic. The evaluation covers all outputs, outcomes and any (un)anticipated impacts, which are clearly defined in this section.
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 graphic organizes results around three pillars (demand, supply, environment) to clearly show the causal relationships between multiple outputs and one primary outcome. The output-level results are sometimes phrased as short-term outcomes, but the logical chain of progression between results is clear.
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 report contains an assessment of the ToC under findings on effectiveness in section 4.2.4. Limits to the programme's theory of change are noted, however, the evaluators did not reconstruct or improve the theory, which was an expected output of the inception phase per the ToR requirements.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	86%	Comments on Rating
Question 8.	Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? <i>UNICEF evaluation standards refer to the OECD/DAC criteria - Relevance; Coherence; Effectiveness; Efficiency; Sustainability; Impact (not all are necessarily relevant for all evaluations). Evaluations should also consider equity and leaving no-one behind, gender and human rights based approach (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.</i>		
i	Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Yes	The evaluation questions are appropriate for meeting the evaluation purpose, and all the questions are based on OECD/DAC criteria. Coherence was not included as a requirement of the ToR, though the evaluators justified the addition of 'human rights-based approach' as a non-DAC criterion to be considered alongside gender equality and equity.
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 in Annex 7 includes key evaluation questions, sub-questions, indicators, data collection and analysis methods, and information sources.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?		

	i	Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Partially	The evaluation notes to have used a mixed-methods and participatory approach, and each method is described satisfactorily. Figure 28 illustrates all of the evaluation methods. However, the extent of engagement of key stakeholders is not much more than standard good practice involving reference groups, interviewing / engaging duty bearers and rightsholders as key informants, and hosting a validation workshop to finalize recommendations. As such, the proposed approaches and methods are mostly sufficient, however it is important that evaluators clearly demonstrate how specific approaches were used in the context of the evaluation at hand.
	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	All data sources are appropriate and well-described. The evaluation used quantitative data (HH survey with treatment and comparison groups) and qualitative data (key informant interviews, focus group discussions, photographs, and validation workshop).
	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 for surveys is clear and well defined in the report, including both a treatment and comparison group. Purposive and/or random sampling was applied for selection of participants in surveys, focus groups and interviews. In Appendix 13, the evaluators describe the multi-stage sampling strategy further, including how diverse perspectives were captured. The graphics in the report also show the extent to which participants were representative of different groups (the proportion of rightsholders engaged, the proportion of rightsholders by category (mother, father, service provider, community leaders, single or young mothers, etc.).
	iv	Clear and complete description of the methods of analysis.	Partially	There is a clear section on data analysis; however, the description of quantitative analysis methods is mostly general and does not capture the full scope of quantitative analysis necessary for such a rigorous quasi-experimental design.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The methodology clearly allows for drawing causal connections through analysis of both treatment and comparison group data against the theory of change.
	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	Table 10 details the limitations and associated mitigation measures for each.
Question 10.		Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
	i	Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Yes	There is an explicit reference to the UNEG obligations of evaluators and UNEG ethical guidelines in a separate section. The evaluation team adhered to all six obligations of UNEG: Integrity, Accountability, Respect, Beneficence, Credibility, and Independence.
	ii	Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.	Yes	There is some reference to ethical safeguards within various sections of the methodology, including compliance with UNEG and UNICEF ethical norms and standards. A clear description of the ethical safeguards for participants is provided in Appendix 23 and Appendix 24.
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.	Partially	Extensive data collection was done with both a treatment and comparison group during the pandemic, which has been quite a challenge to do for many other evaluators. In addition, the evaluators noted that the evaluation "used a hybrid design including comparative and contribution analysis." While these methods are not clearly innovative, nor fully explained in the report methodology, a partial is applied for the extent of rigor in the evaluation design during a complex and challenging time for collecting data.
SECTION E: EVALUATION FINDINGS (weight 25%)			100%	
Question 12.		Do the findings clearly address all evaluation objectives and scope?		
	i	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Yes	The findings are well-organized and answer all the questions and sub-questions given in the evaluation matrix. In addition, there is sufficient evidence provided to substantiate all of the findings.
	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The findings are presented against the programme's results framework and ToC.
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 quantitative and qualitative data. Quantitative data from the treatment and control groups are presented and validated through findings from focus groups and interviews. In addition, each section has stakeholder (key informant) interview excerpts to corroborate the findings from the household level survey.
	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	Findings are adequately supported by the evidence available to evaluators. Both positive and negative findings are discussed. The layout of the findings section is such that each finding responds to the outcome and output related questions.

	iii	Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyze the logical chain (progression -or not- from implementation to results).	Yes	The causal factors are clearly mentioned. In addition, the reasons for the failures and achievements of the programme are clearly highlighted wherever available.
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 provides a detailed assessment of the programme's monitoring and evaluation system and its theory of change.
SECTION F:		EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	88%	
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	Conclusions are formulated by evaluation criteria and provide an additional summary titled "Conclusions on Way Forward."
	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 are appropriately derived from the findings and highlight both the strengths and weaknesses of the programme delivery. However, they are simply stated and do not add much insight or analysis 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	All the lessons are insightful and are drawn from the findings and conclusions. The lessons are relevant, practical, and have broader applicability.
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	The lessons are presented clearly and concisely. In addition, they have sufficient evidence and background details that the intended audience can use to improve future programmes of a similar nature.
SECTION G:		RECOMMENDATIONS (weight 15%)	100%	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	Recommendations are comprehensive, clear and are drawn from the evaluation findings.
	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 listed are practical and actionable for the intended users. They identify both the primary and supporting stakeholders responsible for taking 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 of developing recommendations is clearly described. The stakeholders were asked to share their experiences and aspirations for the future. Further, the evaluators also presented the recommendations to the ERG for their comments and suggestions.
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	Table 26 in the recommendations section provide a list of recommendations with their priority and the names of primary and secondary stakeholders who can take some action on these recommendations.
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	All elements are present in the opening pages.
	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	All elements are present in either the report body or the appendices.
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 structure of the report is easy to navigate. All sections and titles are labelled clearly and are 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.	Yes	The report conforms with UNICEF guidelines. Each of the sections follow a logical order (i.e. context and purpose prior to findings etc.).
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.	Partially	At 63 pages and eight pages of executive summary, the length of the report exceeds the specified limit in the ToR of 60 pages including the executive summary.
	ii	Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	It is written in an accessible way, but there are noticeable grammatical errors, spelling mistakes or typos, 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 report uses tables and graphs to convey key information, however they are not always clearly presented or labeled.
SECTION I:		EVALUATION PRINCIPLES (weight 10%)	95%	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	There is an explicit mention of CRC in the introduction section and the evaluation findings. In addition, the evaluation assessed the programme goals using child rights-based frameworks and child rights provisions specified under international and regional treaties.
	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	Appendix 8 on the evaluation stakeholders' roles, interests, and uses of evaluation provides detailed information in this regard.
	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, and duty-bearer and beneficiary terminologies are used. A special attention is given to ensure the evaluation methodology was gender sensitive and captured gender-disaggregated data.
	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	Child rights are the main focus of the evaluation. Children with disability are also included in the evaluation and were identified through survey questions.
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)	8	
	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	GEEW is integrated in the evaluation design through explicit reference to diverse groups, including women and girls, within the evaluation objectives as well as the addition of the non-DAC criterion of human rights-based approach, gender equality and equity to the evaluation, despite it not being included in the ToR. In addition, one broad and three sub-questions on gender equity and equality are included in the list of key evaluation questions.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Fully integrated	The steps taken to ensuring the methodology was gender responsive are described. There is a comprehensive description of the ethical practices used in the appendices. A gender disaggregated list of evaluation participants is provided in Table 9 and shows diversity in the stakeholders consulted.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Satisfactorily integrated	The findings discuss human rights-based approaches, gender equality, and gender equity by answering questions and using standardized frameworks designed specifically to understand the extent of gender mainstreaming in the design, implementation, and results. The evaluation also sought to provide an analysis of the programme against UN-SWAP indicators, however the relevance of this framework for assessing a specific programme's compliance is not clear or substantiated. Gender-specific conclusions and recommendations are also included. Good practice would be to also include gender-specific lessons learned.

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?