

# GEROS Evaluation Quality Assurance Tool

Version: September 2021

## Gender and the COVID-19 response in MENA

REPORT RATING SUMMARY			
<b>Overall Rating</b>		51%	Fair
●●●●●	Exceptional (96% - 100%)	5	
●●●●○	Highly Satisfactory (87.5% - 95.99%)	4	
●●●○●	Satisfactory (62.5% - 87.49%)	3	
●●○●●	Fair (35% - 62.49%)	2	Meets UNICEF/UNEG standards for evaluation reports in some regards, but not all. Decision makers may continue to use the evaluation with caution, but substantive improvements are possible.
●○○○○	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
<b>Title of the evaluation report</b>		Gender and the COVID-19 response in MENA	
<b>Report sequence number</b>		MENARO, Jordan/59/2022/18419	
<b>Region</b>		MENA	
<b>Year of report</b>		2022	
<b>Office</b>		Jordan	
<b>Coverage (countries)</b>		Lebanon, Jordan and Yemen, partial coverage Morocco and Syria.	
<b>ToRs present</b>		Yes	
<b>Date of review (dd/mmm/yyyy)</b>		January 10, 2023	
<b>Name of review firm</b>		IOD PARC	
CLASSIFICATION OF EVALUATION REPORT			
<b>Management of evaluation (Managerial control and oversight of evaluation)</b>		Not clear from report	
<b>Unicef goal areas (Alignment with strategic plan priorities)</b>		5	
Every child survives and thrives		Yes	
Every child learns		Yes	
Every child is protected from violence and exploitation		Yes	
Every child lives in a safe and clean environment		Yes	
Every child has an equitable chance in life		Yes	
Gender equality (cross-cutting)		Yes	
Humanitarian action (cross-cutting)		No	
<b>Evaluation object</b>		Programme	
<b>Evaluation type</b>		Summative	
<b>Evaluation strategy</b>		Mixed methods	
<b>Evaluation design (primary method used)</b>		Other	
<b>Evaluation level</b>		Output & Outcome	
<b>Geographic scope</b>		Multi-country	
<b>Primary SDG(s) covered (number)</b>		5	
<b>EQA Summary:</b> <i>The rater will provide top line issues for this evaluation relevant for feedback to senior management (positive and negative), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight best practice/added value elements and the level of complexity of the evaluation.</i>			
<p>The findings section in particular is full of useful evaluative information to understand what progress and challenges have been made in integrating a gender perspective into the response to the COVID 19 crisis in the MENA region. However, the evidence underpinning these explanations and assessments is not entirely clear. Although the documentary analysis does appear in several of the findings to support these explanations and assessments, more strikingly, stakeholders' contributions to the findings are unclear. No inventory of stakeholders is given and sampling strategies are not contained in the report; an anonymised list of contributors is given but there is no indication of whether this was a sufficient sample; large enough and/or a varied enough group to support the findings.</p> <p>The introductory sessions have good points and bad points. The explanation of the context in which the interventions took place is very good. However, the definition and description of the interventions per se are very incomplete. The formulation of the evaluation questions, and especially of the sub-questions, is problematic. There are 83 sub-questions, some of which are very specific, but others are so broad that they could be the object of a separate evaluation.</p> <p>The lessons learned are very well formulated, the recommendations are generally actionable, but some are too ambitious and vague. It is very striking that no conclusions are included.</p>			
<b>Recommendations for Improvement:</b> <i>The rater will identify topline recommendations to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, resources will be cited to assist evaluation managers in overseeing future evaluations.</i>			
<p>In order of importance:</p> <ol style="list-style-type: none"> <li>1. It would be useful to make sure that there is evidence supporting the findings, i.e. why you know what you claim, is clearer.</li> <li>2. It is essential to include a 'conclusions section' that provides a second level of analysis.</li> <li>3. The analysis and ultimately the report would benefit from having a better quality evaluation matrix with much more focused sub-questions.</li> <li>4. The interventions being evaluated should be better defined, especially the logic of the interventions.</li> <li>5. The presentation of the report and its navigability would benefit from a functional table of contents and the inclusion of graphs and figures that add value.</li> </ol>			
SECTION RATINGS			
<b>SECTION A: EXECUTIVE SUMMARY (weight 5%)</b>		33%	
<b>Question 1.</b>	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).	No	The executive summary is four pages long. Although it is not of publishable quality, it is clearly understandable. Its usefulness for decision-making is limited mainly because it does not include conclusions. This makes the traceability between the findings and the recommendations unclear, which undermines the legitimacy of the recommendations.
ii	Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions on findings, lessons learned if requested, key recommendations) as per the ToR.	Partially	Most of these components are present in the executive summary, however there are two important ones that are not included. The most relevant are the conclusions, the second are the users and uses of the evaluation.

	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	The information included is sufficient to understand the overall rationale of the intervention, although, as in the body of the report, there is no specific information on the logic framework or the ToC . Another caveat, although not expressly requested, it would be advisable to ask for the inclusion of financial data in the executive summary, which would help the reader to better understand the scope of the intervention. The executive summary does not include information that is not present in the report, in fact most of the text in the summary is composed of extracts copied and pasted from the body of the report.
<b>SECTION B: BACKGROUND (weight 5%)</b>			21%	
<b>Question 2.</b>		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.	No	A shortcoming of the report is that neither the body of the report nor the annexes give details of these aspects of the programmes reviewed, i.e. in Jordan (MAKANI, Learning Bridges and WASH interventions out of camps); in Lebanon (LAHA Kits, Child Protection Programme (including its Qudwa Strategy) and Integrated Child Wellbeing Programme), and in Yemen (Jump Start Package & Community Health Empowerment). Annex 2 gives a very brief description of some aspects of the programmes in relation with the rationale behind their selection.
	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).	No	Throughout the report, the most relevant stakeholders are mentioned. They are also explicitly stated in the typology used for the KIIs. However, there is no specific section detailing who is who in the different programmes (or an overview of all of them). Hence, nor does it include a segregation of the different participants in the programmes.
<b>Question 3.</b>		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	In the chapter on the background, there is no explicit analysis of the relationship between the context(s) and the programmes being analysed, although this relationship can be easily inferred. Apart from this, the part dedicated to the context is complete. The most remarkable aspect is that it has been divided into blocks by country and that the gender analysis, although brief, is very relevant.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	No	The SDGs are mentioned only three times in the entire report and never in relation to targets and indicators.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	In most descriptions of country contexts, there is a good gender analysis highlighting the needs of women who are ultimately the right holders of the object of the evaluation. In addition, in the relevant places, there is also a differentiated analysis of the conditions and needs of the refugee population.
<b>Question 4.</b>		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).	No	There is an absence of a complete description of the programmes that have been analysed, there is also no stakeholder analysis. In 9.7. of annex 9, it is mentioned that the documentary analysis allowed the evaluation team to do a stakeholder analysis, but this is not included in the report.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	No	The contributions of UNICEF or other actors to the programmes analysed by the evaluation are not described.
<b>SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)</b>			42%	
<b>Question 5.</b>		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.	Partially	The rationale for the evaluation is well described. However, the uses and users of the evaluation are not clear, other than that it will be useful for UNICEF.
<b>Question 6.</b>		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 objectives are clear and do not deviate from what was proposed in the terms of reference.
	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	Some aspects of the scope are very clear, in particular the geographical scope. It is explained which area is covered by the evaluation and which countries and programmes have been used as examples in order to be able to infer more far-reaching findings and why they have been chosen. This description is given in the body of the report and in Annex 2. This annex also describes the thematic foci in each country, e.g. education, GBV, social protection, etc. It would be advisable to have at least one outline of this thematic scope in the body of the report. What is missing is the definition of the timeline in which the evaluation is framed.
<b>Question 7.</b>		Is the theory of change, results chain or logic well articulated?		

	i Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Partially	There is a reasonably clear theory of change on gender mainstreaming for the MENA region as part of the ToR. This ToC has been used as a basis to elaborate a contribution analysis by the evaluation. What is not presented is a theory or logic chain on how a gender-transformative approach was intended to be integrated into the actions stemming from the COVID 19 emergency.
	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).	No	It is important to note that what is being evaluated here is not a specific programme, but an approach (gender approach) applied to a series of programmes that were formulated to alleviate the COVID 19 crisis. However, what could have been done but has not been done, is a typological grouping of both outputs and outcomes of the programmes that were identified as case studies in order to understand what the trends in causal design were and how gender was integrated on paper.
	iii For theory-based evaluations, the theory of change or results framework is assessed, and if requested in the ToR, it is reformulated/improved by the evaluators.	No	In the first pages of the report it is stated that the ToC provided by the ToR has been used for a contribution analysis, which was the approach chosen by the evaluation. While there is no doubt that this ToC was the subject of a detailed study by the evaluation team, it is not obvious that the evaluation was based on such a theory or applied such an approach.  Annex 9, however, explains that the contribution approach was dropped in favour of a transformative gender focus and a human rights approach. These two approaches are not evaluation approaches, but essential aspects that must be present in whatever evaluative approach is chosen.
<b>SECTION D:</b>	<b>EVALUATION DESIGN AND METHODOLOGY (weight 20%)</b>	30%	
<b>Question 8.</b>	Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? <i>UNICEF evaluation standards refer to the OECD/DAC criteria - Relevance; Coherence; Effectiveness; Efficiency; Sustainability; Impact (not all are necessarily relevant for all evaluations). Evaluations should also consider equity and leaving no-one behind, gender and human rights based approach (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.</i>		
	i Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Partially	The evaluation questions have respected those proposed in the ToR. These questions, although a good starting point, could have been further refined during the inception period. A further point to note about the sub-questions that have been formulated is that they are too ambitious. There are 83 sub-questions that cannot be answered in a single evaluation unless the analysis is extremely superficial. Furthermore, some of the sub-questions could be the purpose of an entire evaluation. For example, the first sub-question asks "How did UNICEF's COVID-19 response influence policies at the national level in the areas of health and nutrition, child protection, education, WASH, and GBV (amid COVID-19 especially) (if it has)? How did UNICEF contribute to these policies from a gendered perspective?". Considering that the object of the evaluation is multi-country, this question could well be the purpose of a complex evaluation.  It is not very clear why the criteria that are the focus of the evaluation have been chosen. In general, the questions and sub-questions are aligned with the evaluation criteria that have been chosen, although with 83 sub-questions it is confusing to understand if these are duplications.
	ii In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions and/or other processes from which the analysis can be based and conclusions drawn.	Partially	The matrix presents the main questions, but not the sub-questions, a description of the evidence to be collected, the sources of information, and the methods of collecting and analysing it. However, neither assumptions nor indicators are presented, nor is it explained why they have not been integrated. It is mentioned in annex 9 that the benchmarks that have been used are those formulated in the " Gender Action Plan (2018-2021); the Five Actions for Gender Equality in the COVID-19 Response; and the CCC which forms the core of UNICEF'S policy and framework for humanitarian action " but they have not been integrated in the matrix.
<b>Question 9.</b>	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.	No	The definition of the methodology is very poor, which does not allow us to assess whether the methods used are robust enough to arrive at credible findings.
	ii Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Partially	The sources of information are those you would expect to find in a typical evaluation, i.e. a sample of stakeholders and a documentary review. However, in the absence of a stakeholder analysis or at least a stakeholder inventory, it is not possible to assess whether they are adequate in scope. It seems clear on the other hand that the qualitative sources have been from the second level, but it is not clear where they finally came from as in annex 9 the future tense is used, as if we were reading the inception report, which is confusing (example: the team will gather official government statistics).

	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	No	Probably one of the fundamental flaws of the methodology is that there is no definition of the sampling strategy. This is particularly shocking in a gender-specific evaluation because it is not explained to what extent the voices of women and other discriminated groups have been intentionally targeted in the strategies for choosing evaluation participants.
	iv	Clear and complete description of the methods of analysis.	Partially	It is mentioned that there has been a process of triangulation to arrive at the findings, but it is not explained whether this has been done for each sub-question and what tools have been used to do the triangulation, e.g. team meetings, qualitative analysis software, sense making meetings, etc.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Not Rated	In this evaluation this question is not relevant as it is not a programme with outputs and outcomes, but an approach that has been evaluated
	vi	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Partially	A couple of limitations are mentioned in the data collection of the assessment and the mitigation measures that were adopted in this respect. However, it seems clear that the list of limitations is not exhaustive. For example, while mention is made of the difficulty of obtaining data disaggregated by sex, there is no mention of the difficulty of obtaining balanced data for men and women and other groups such as LGBTBIQ+ or people with different (dis)abilities. Discriminated groups are often more difficult to attract. Indeed, paragraph 64 of the report makes a brief mention of the difficulty of reaching certain groups of people.
<b>Question 10.</b>		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	Most of the ethical issues mentioned are listed in Annex 4. These aspects relate entirely to the collection of information from participants. In other words, they refer to concepts such as consent, anonymity and data handling. Although UNICEF's ethical guidelines are mentioned, there is no mention of UNEG in this issue.
	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.	No	It is quite possible that the evaluation team observed all or most of these principles given the sensitive and cautious tone of the report, which respects the dignity of the informants. However, these issues are not explicitly mentioned in either the methodological or ethical protocol parts of the report.
<b>Question 11.</b>		Does the evaluation incorporate innovative practice that adds value to the evaluation process?		
	i	Innovation practice is used to improve the quality of evaluation process. This could be evident in several ways such as the design of the methodology (i.e. use of technology for data gathering, extensive participatory processes, systematic analysis processes such as collaborative outcomes reporting and incorporation of big data, specific strategies to address complexity such as outcome harvesting, strong child rights focus), or ways of sharing of evaluation results.	No	The evaluation has used standard practices that cannot be qualified as innovative.
<b>SECTION E:</b>		<b>EVALUATION FINDINGS (weight 25%)</b>	<b>63%</b>	
<b>Question 12.</b>		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 organisation of the findings chapter is confusing. It seems to be organised by evaluation questions (although this is not explicit) and does not seem to address sub-questions. This in itself is not a problem if it had indicated how the sub-questions have guided the analysis. In the formulation of these findings there is little mention of the supporting evidence.
	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Not Rated	The approach under evaluation did not have a ToC to begin with and has not been reconstructed during the course of the evaluation.
<b>Question 13.</b>		Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		

	i Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.	Partially	In general the findings have a great deal of evaluative content that is potentially useful for UNICEF. However, again with exceptions and in general, these assessments are not supported by evidence. That is, the report explains what has happened and, often, what implications it has for UNICEF, but it does not explain where these assessments are drawn from.  As mentioned previously, there are exceptions to this rule, for example "the extent to which gender is factored into the early inception phases of programming varies between COs (...) the CPP's Qudwa strategy. For example in Lebanon, where the strategy informs all sub-activities and initiatives under this programme, as well as serving as a framework under the sub-activities factor in for the duration of a response (UNICEF, MoSA and Magenta (2021), Qudwa Social and Behavioural Change Communications Plan to Prevent Violence against Children & Women, Child Marriage and Child Labour). Moreover, the influence and participation of Gender Specialists and GFPs was more pronounced in COs where leadership endorsed gender equality".  It is also important to note that although evidence from documentary review has been used with some frequency, there has been little analysis of the opinions of the people interviewed.
	ii Findings are clearly supported by, and respond to, the evidence presented, including both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.	Partially	It is mentioned under the methodology that the findings of the evaluation should use the benchmarks identified in the Gender Action Plan (2018-2021); the Five Actions for Gender Equality in the COVID-19 Response; and the CCC which form the core of UNICEF'S policy and framework for humanitarian action. However, in practice, although some of the content appears to be structured around concepts contained in these documents, it is not clear that the benchmarks are being used.  There are exceptions though. for example "Gender expenditure varies across the MENA COs; however, the majority were able to meet UNICEF's 15% benchmark".  The nuance of "it is not clear that the benchmarks are being used" is important because these documents that seem key to the analysis contained in the report are not annexed.
	iii Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyse the logical chain (progression -or not- from implementation to results).	Yes	The report has placed great emphasis on placing the findings in the organisational contexts in which they occur and how these institutional factors have or have not helped in their achievement.
<b>Question 14.</b>	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.	Not Rated	The approach assessed in the report does not have a RBM framework per se.
<b>SECTION F:</b>	<b>EVALUATION CONCLUSIONS &amp; LESSONS LEARNED (weight 10%)</b>	<b>50%</b>	
<b>Question 15.</b>	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).	No	The report does not include conclusions (!)
	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.	No	(see above)
<b>Question 16.</b>	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 fact that lessons learned have been formulated for each of the evaluation criteria discussed makes the link to the findings more obvious.  The cross-cutting nature of the object of the evaluation facilitates the application of the lessons learnt to a wide range of programmes beyond those that have been analysed.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	Lessons learned are presented in an unorthodox manner, i.e. at the end of each of the sections dealing with evaluation criteria. This makes them more numerous than those normally included in evaluation reports. However, they do justice to the weight given to this aspect in the evaluation objectives.
<b>SECTION G:</b>	<b>RECOMMENDATIONS (weight 15%)</b>	<b>63%</b>	
<b>Question 17.</b>	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 are clear. It is useful that for each of the recommendations it is detailed on which findings they are based.

	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	In general, recommendations are based more on what to do than on how to do it. However, there are exceptions, for example, in recommendation 8 "Train GFPs particularly in countries that do not have a Gender Specialist or senior gender full-time staff, in order for programming to be effective. To fill this gap, RO gender specialists could provide intensive two-three day training to COs/increase field visits".
	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).	No	There is no explanation as to how the recommendations have been developed.
<b>Question 18.</b>		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	They are presented in an unorthodox way after the findings of the evaluation criteria and with a summary at the end of the report. The presentation, although infrequent, is clear and each recommendation is prioritised.
<b>SECTION H:</b>	<b>REPORT STRUCTURE AND PRESENTATION (weight 5%)</b>		<b>50%</b>	
<b>Question 19.</b>		Does the evaluation report include all relevant information?		
	i	Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes; list of acronyms/abbreviations, page numbers.	Partially	The report contains most of these components, although there are areas for improvement, the two main ones being that the introductory pages do not contain a summary of the subject of the evaluation and that the table of contents does not correspond to the contents of the report. This makes the report difficult to navigate.
	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	Most of the annexes that would be expected are included. However, there is a lack of documents that are key guidelines for the analysis of the evaluation, in particular the UNICEF Gender Action Plan 2018 2021.
<b>Question 20.</b>		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).	Partially	There are some very useful aspects, such as the numbering of paragraphs and the different colours for the summary boxes. However, other important aspects such as the non-functional table of contents make the document difficult to 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.	Partially	The structure of the report follows a logical sequence as outlined by UNICEF, but with a number of differences. Most notably, it does not include conclusions. In addition, lessons learned and recommendations are presented after the findings for each evaluation criterion.
<b>Question 21.</b>		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 ToR did not specify the expected length of the report. However, the number of pages of the main body of the report, excluding annex and executive summary, seems reasonable given the complexity of the object of the evaluation.
	ii	Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	Not of publishable quality but reasonably well understood.
	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.	No	All graphics are relegated to an annex, only a few boxes have been used.
<b>SECTION I:</b>	<b>EVALUATION PRINCIPLES (weight 10%)</b>		<b>74%</b>	
<b>Question 22.</b>		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	Yes, the report mentions all of these UNICEF commitments at relevant points in the report.
	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	In one of the annexes, the persons who served as informants are listed anonymously and disaggregated by sex. Beyond that, there is no further explanation of the different groups that participated in the evaluation.
	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	As an explicit gender evaluation, one would expect no less.
	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	While more disaggregated analyses are always possible, these values are at the core of the analysis in this report.
<b>Question 23.</b>		Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards with detail provided below)	5	

i	GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Partially integrated	GEEW is naturally integrated into the analysis and the questions are gender-sensitive. However, there is no guarantee that GEEW related data could have been collected. It is not explicit in the report, not all informants are disaggregated by sex, so it is not known which voices contributed to the analysis. It is also not clear that the data collection settings were empowering or friendly to women in order to collect quality data.
ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	(see above)
iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	Since it is a gender-specific evaluation, this is to be expected.

### SWAP Rating Guidance

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