


GEROS Evaluation Quality Assurance Tool


Version 2018.1


Reviewers: complete all cells highlighted in Yellow


		Notes for QA
Title of the Evaluation Report	EVALUATION OF UNICEF'S Coverage and Quality in Complex Humanitarian Emergencies	
Report sequence number	2018/018	
Region	HQ	
Year of Report	2018	
Office	Evaluation Office	
Coverage (countries)	Global	
ToRs present	Yes	
Date of Review	5/18/2019	
Name of reviewer	ImpactReady	
Classification of Evaluation Report		
Management of Evaluation (<i>Managerial control and oversight of evaluation decisions</i>)		UNICEF managed
SPOA Correspondence <i>(Alignment with SPOA focus area priorities)</i>	Health	Yes
	HIV/AIDS	Yes
	WASH	Yes
	Nutrition	Yes
	Education	Yes
	Child protection	Yes
	Social inclusion	Yes
	Gender equality (cross-cutting)	Yes
	Humanitarian action (cross-cutting)	Yes
Evaluation object		Thematic area
Evaluation type		Formative
Evaluation strategy		Case study
Evaluation design		Mixed Methods

Evaluation level		Outcome
Geographic Scope		Multi-region/Global
SECTION A: BACKGROUND (weight 5%)		92%
Question 1.	Is the object of the evaluation clearly described?	100%
	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status	Yes
	Clear and relevant description of intended beneficiaries 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 (as appropriate to the purpose of the evaluation)	Yes
	Description of the relative importance of the object to UNICEF (e.g. in terms of size, influence, or positioning)	Yes
Question 2.	Is the context of the intervention clearly described?	100%
	Clear and relevant description of the context of the intervention (policy, socio-economic, political, institutional, international factors relevant to the implementation of the intervention)	Yes
	Clear and relevant description (where appropriate) of the status and needs of the target groups for the intervention	Yes
	Explanation of how the context relates to the implementation of the intervention	Yes
Question 3.	Is the results chain or logic well articulated?	100%
	Clear and complete description of the intervention's intended results	Yes
	Intervention logic presented as a coherent theory of change, logic chain or logic framework	Not Rated
Question 4.	Are key stakeholders and their contributions clearly identified?	67%
	Identification of implementing agency(ies), development partners, primary duty bearers, secondary duty bearers, and rights holders	Mostly
	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF	Mostly
••••	Executive Feedback on Section A	Overall rating for Section
The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements		Satisfactory
<p>The object of the evaluation is well described. The report includes a description of the object of the evaluation, i.e. coverage and quality of humanitarian responses, along with the contexts in which some of the humanitarian responses take place around the world. The report provides a complete description of the coverage of many humanitarian responses including the budgets involved, although other implementing actors (other agencies and INGOs) and their roles and contributions could be discussed in further detail. Also, the report describes the numbers of beneficiaries of the humanitarian responses over the last years and the most important challenges of operating in contexts of humanitarian crises. This helps the reader's understanding of the particularities of measuring quality and coverage of the initiatives undertaken under these circumstances. Finally, the analytical framework of the evaluation is clearly outlined.</p>		 3
Additional comments for Section A (recommendations for improvement)		Weighting
It would be useful to include further discussion of other implementing agencies and their mandates as well as INGOs, national governments, etc., and how UNICEF has collaborated with them.		0.05


SECTION B: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)		100%
Question 5.	Is the purpose of the evaluation clearly described?	100%
	Specific identification of how the evaluation is intended to be used and to what this use is expected to achieve	Yes
	Identification of appropriate primary intended users of the evaluation	Yes
Question 6.	Are the objectives and scope of the evaluation clear and realistic?	100%
	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	Yes
	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 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
••••	Executive Feedback on Section B	Overall rating for Section
The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements		Highly Satisfactory
<p>The purpose of the evaluation is clearly stated and the intended use and primary users are explicitly listed. The purpose of this evaluation is described as generating practical solutions that can inform how UNICEF improves the coverage and quality of its humanitarian response. Furthermore, the intended use and intended primary users of the evaluation are duly described and the scope of the evaluation is explained in terms of its temporal coverage (the period 2015-2018). Additionally, the analytical framework used is described in order to circumscribe what the evaluation was set out to explore. Similarly, the geographic scope of the evaluation, the country-case studies and a rationale for the decisions made in this regard are provided.</p>		4
Additional comments for Section B (recommendations for improvement)		Weighting
This section observes good practices. No further improvement is required.		0.05
SECTION C: EVALUATION METHODOLOGY (weight 15%)		97%
Question 7.	Does the evaluation provide a relevant list of evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? UNICEF evaluation standards refer to the OECD/DAC criteria. Not all OECD/DAC criteria are relevant to all evaluation objectives and scopes. Standard OECD DAC Criteria include: Relevance; Effectiveness; Efficiency; Sustainability; Impact. Evaluations should also consider equity, gender and human rights (these can be mainstreamed into other criteria). Humanitarian evaluations should consider Coverage; Connectedness; Coordination; Protection; Security.	100%
	Clear and relevant presentation of the evaluation framework including clear evaluation questions used to guide the evaluation	Yes
	If the framework is OTHER than UNICEF standard criteria, or if not all standard criteria of the chosen framework are included, the reasons for this are clearly explained and the chosen framework is clearly described	Yes
Question 8.	Does the report specify methods for data collection, analysis, and sampling?	92%
	Clear and complete description of a relevant design and set of methods that are suitable for the evaluation's purpose, objectives and scope	Yes

	Clear and complete description of the data sources, rationale for their selection and sampling strategy. This should include a description of how diverse perspectives are captured (or if not, provide reasons for this), how accuracy is ensured, and the extent to which data limitations are mitigated	Yes
	Clear and complete description of the methods of analysis, including triangulation of multiple lines and levels of evidence (if relevant)?	Yes
	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias?	Mostly
Question 9.	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:	100%
	Explicit reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability)	Yes
	Description of ethical safeguards for participants appropriate for the issues described (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm)	Yes
	ONLY FOR THOSE CASES WHERE THE EVALUATION INVOLVES INTERVIEWING CHILDREN: explicit reference is made to the UNICEF procedures for Ethical Research Involving Children	Yes
•••-	Executive Feedback on Section C	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements	Satisfactory
	The report provides a discussion regarding the rationale for conducting the present evaluation, i.e. the need to better understand, analyse and address the challenges of achieving coverage and quality in crises. Since the object of the evaluation is UNICEF's humanitarian responses, the report does not use the standard OECD/DAC criteria, i.e. relevance, effectiveness, efficiency, sustainability and impact. Instead, the report explains that an analysis framework and key evaluation questions are used to guide the assessment and a full evaluation matrix is included in Annex 2, which includes all of the evaluation questions and sub-questions as well as the indicators and data collection sources and methods used. The report explains that the evaluation used mixed methods including a document review, Key Informant Interviews, Focus Group Discussions, as well as desk and field-based case studies. Similarly, the data analysis methods and synthesis used are explained as well as the quality assurance process through triangulation of evidence from different sources and levels. Sampling methods are also explained and a very complete explanation of the criteria used in this regard is provided in Annex 9. On the other hand, while methodological and operational limitations are correctly identified and explained, mitigation strategies for these are not consistently provided. Finally, the report discusses the ethical considerations applied during the conduct of the evaluation and a link is provided that outlines the ethical obligations of the evaluators and research ethics involving children.	 3
	Additional comments for Section C (recommendations for improvement)	Weighting
	It is good practice to provide mitigation strategies for any limitations faced by the evaluation in order to overcome data gaps and avoid bias.	0.15
SECTION D: EVALUATION FINDINGS (weight 20%)		92%
Question 10.	Do the findings clearly address all evaluation objectives and scope?	100%

	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions and criteria	Yes
	Reference to the intervention's results framework in the formulation of the findings	Yes
Question 11.	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.	75%
	The evaluation clearly presents multiple lines (including multiple time series) and levels (output, outcome, and appropriate disaggregation) of credible evidence.	Yes
	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.	Yes
	Unexpected effects (positive and negative) are identified and analysed	No
	The 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
Question 12.	Does the evaluation assess and use the intervention's Results Based Management elements?	100%
	Clear and comprehensive assessment 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)	Yes
	Clear and complete assessment of the use of monitoring data in decision making	Yes
• • • -	Executive Feedback on Section D	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements	Satisfactory
	The report does a good job at presenting particularly robust findings that are supported by a great deal of evidence that is organized in a clear fashion and drawn from different sources, referring to both strengths and areas for improvement. The findings are organized so as to respond to key evaluation questions, and the initiative's performance is consistently supported by relevant concrete examples of country case studies and with reference to the results framework. On the other hand, the report does not present a description and analysis of unintended positive or negative effects. Finally, one of the strongest points of this evaluation report is the description and deep analysis it provides of the humanitarian responses' M&E system. The report provides a number of concrete examples of its performance in different contexts including an analysis of the way this system informed decision-making and, especially ways in which monitoring and evaluation processes and mechanisms could be strengthened.	 3
	Additional comments for Section D (recommendations for improvement)	Weighting
	Even if the object of the evaluation may not lend itself to a deep analysis of unintended effects, it is good practice to include an assessment of any occurrence of such unexpected results as they may inform of any issues that were not captured by the design of the responses or in the scope of the evaluation.	0.2
SECTION E: EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 15%)		83%
Question 13.	Do the conclusions present an objective overall assessment of the intervention?	100%
	Clear and complete description of the strengths and weaknesses of the intervention that adds insight and analysis beyond the findings	Yes
	Description of the foreseeable implications of the findings for the future of the intervention (if formative evaluation or if the implementation is expected to continue or have additional phase)	Yes
	The conclusions are derived appropriately from findings	Yes

Question 14.	Are lessons learned correctly identified?	67%
	Correctly identified lessons that stem logically from the findings, presents an analysis of how they can be applied to different contexts and/or different sectors, and takes into account evidential limitations such as generalizing from single point observations.	Mostly
••••	Executive Feedback on Section E	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements	Satisfactory
	The report presents conclusions that are correctly drawn from the evidence presented in the findings section. Furthermore, conclusions provide a much-needed overview of the strengths and areas for improvement in the coverage and quality of UNICEF's humanitarian responses. The conclusions have a clear forward-looking perspective and discuss foreseeable implications on future programming in this area. Finally, the report presents lessons learned that are formulated so as to signify what wider relevance they could have in other initiatives and different contexts, although they are scattered through the findings section as opposed to being clearly presented within their own section.	 3
	Additional comments for Section E (recommendations for improvement)	Weighting
	For the sake of clarity, it is recommended that lessons learned from the evaluation be presented within their own section.	0.15

SECTION F: RECOMMENDATIONS (weight 15%)	83%
--	------------

Question 15.	Are recommendations well grounded in the evaluation?	67%
	Recommendations are logically derived from the findings and/or conclusions	Yes
	Recommendations are useful to primary intended users and uses (relevant to the intervention and provide realistic description of how they can be made operational in the context of the evaluation)	Yes
	Clear description of the process for developing recommendations, including a relevant explanation if the level of participation of stakeholders at this stage is not in proportion with the level of participation in the intervention and/or in the conduct of the evaluation	No
Question 16.	Are recommendations clearly presented?	100%
	Clear identification of target group for action for each recommendation (or clearly clustered group of recommendations)	Yes
	Clear prioritization and/or classification of recommendations to support use	Yes
••••	Executive Feedback on Section F	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements	Satisfactory
	Recommendations are logically derived from the information provided in the findings and conclusions and are robust in the detail they provide about the way UNICEF should go about their implementation. Also, the report is strong at presenting recommendations that are clustered into five different strategic areas and provide clear identification of the target organisational level for action in each case. Conversely, the process followed in developing the recommendations as well as the level of stakeholder participation in their validation are not discussed in the report.	 3
	Additional comments for Section F (recommendations for improvement)	Weighting

Good practices recommend to include a description of the way recommendations were developed and formulated in order to provide more credibility. Also, the level of stakeholder involvement in the validation of these normally affects the level of ownership by stakeholders and helps to ensure that recommendations are realistic and actionable.

0.15

SECTION G: EVALUATION STRUCTURE/PRESENTATION (weight 5%)	75%
---	------------

Question 17.	Does the evaluation report include all relevant information?	83%
--------------	--	------------

	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	Mostly
--	--	---------------

	Annexes should include, when not present in the body of the report: Terms of Reference, Evaluation matrix, list of interviewees, 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, copy of the results chain, information about the evaluator(s)	Yes
--	--	------------


Question 18.	Is the report logically structured?	67%
--------------	-------------------------------------	------------

	The structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles	Partly
--	---	---------------

	Context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations	Yes
--	--	------------

••• -	Executive Feedback on Section G	Overall rating for Section
--------------	--	-----------------------------------

	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements	Satisfactory
--	---	---------------------

	The report observes for the most part the standard order of sections. However, the report is excessively long (over 120 pages). The opening pages of the report include most of the necessary elements to readily inform the reader including the name of the evaluation object, name of the commissioning organization, date of submission of the report, as well as a table of contents, and acronyms. On the other hand, the name of the evaluator(s) or the evaluation firm and the timeframe to be evaluated are not included. Finally, the report provides nine (9) annexes in a separate document which provide valuable information such as the evaluation matrix, the ToR, bibliography, list of people interviewed, etc., which greatly contribute to increase the credibility of the evaluation.	 3
--	--	--


Additional comments for Section G (recommendations for improvement)	Weighting
--	------------------

The opening pages should include all of the necessary elements to ensure that readers can readily grasp the timeframe to be evaluated, the fact that this is a global evaluation, and the names of the evaluators or evaluation firm. Also, the evaluation report should ideally not exceed 60 pages in order to observe good practices, with any extra information being annexed to the report.	0.05
---	-------------

SECTION H: EVALUATION PRINCIPLES (weight 15%)	79%
--	------------

Question 19.	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?	89%
--------------	--	------------

	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	Yes
--	--	------------

	Clear description of the level of participation of key stakeholders in the conduct of the evaluation, and description of the rationale for the chosen level of participation (for example, a reference group is established, stakeholders are involved as informants or in data gathering)	Mostly
	Stylistic evidence of the inclusion of these considerations can include: using human-rights language; gender-sensitive and child-sensitive writing; disaggregating data by gender, age and disability groups; disaggregating data by socially excluded groups.	Yes
Question 20.	Does the evaluation assess the extent to which the implementation of the intervention addressed gender, equity & child rights?	92%
	Identification and assessment of the presence or absence of equity considerations in the design and implementation of the intervention	Yes
	Identification and assessment of the presence or absence of gender in the design and implementation of the intervention	Yes
	Explicit analysis of the involvement in the object of right holders, duty bearers, and socially marginalized groups, and the differential benefits received by different groups of children	Yes
	Clear proportionality between the level of participation in the intervention and in the evaluation, or clear explanation of deviation from this principle (this may be related to specifications of the ToRs, inaccessibility of stakeholders at the time of the evaluation, budgetary constraints, etc.)	Mostly
Question 21.	Does the evaluation meet UN SWAP evaluation performance indicators?	Approaches requirements
Note: this question will be rated according to UN SWAP standards		5
UN-SWAP criterion 1:	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 1 point
UN-SWAP criterion 2:	A gender responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated 1 point
UN-SWAP criterion 3:	The evaluation findings, conclusions and recommendations reflect a gender analysis.	Fully integrated 3 points
•••-	Executive Feedback on Section H	Overall rating for Section
The rater will briefly assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). The rater should include one labelled sentence for each of the UN-SWAP criteria.		Satisfactory
<p>The report explicitly states the use of a rights-based approach and the use of a language that reflects this throughout the report is consistent. For instance, the report includes a detailed description of the legal basis for humanitarian action and International Humanitarian Law and International Human Rights Law are explicitly referred to and described. On the other hand, while the report states that an evaluation reference group was in place, the impossibility of having their feedback in a timely manner is stated among with the limitations of this report. The report correctly assesses the extent to which equity and gender considerations were integrated in the design and implementation of the initiative. Furthermore, even though a gender perspective is not evident from the scope of analysis through the inclusion of gender-related questions, the findings, conclusions and recommendations include a meaningful gender analysis which cascades down through conclusions and recommendations.</p>		 3
Additional justifications for UN-SWAP ratings		
UN-SWAP criterion 1:	The scope of the analysis could have more explicitly explained the way a gender perspective was integrated into the evaluation.	

UN-SWAP criterion 2: The methodology does not sufficiently discuss the way a gender-sensitive approach was used to data collection and analysis.

UN-SWAP criterion 3: The findings, conclusions and recommendations consistently present a gender analysis.

Additional comments for Section H (recommendations for improvement)

Weighting

It is important to include a more explicit description of the ERG, i.e. its composition and its contributions to the evaluation purpose. Additionally, the gender sensitive elements of the evaluation scope and methodology could be strengthened. To learn more about gender sensitive evaluations, please see:
http://www.betterevaluation.org/en/resources/guide/gender-responsive_evaluation_handbook

0.15

SECTION I: EXECUTIVE SUMMARY (weight 5%) 56%

Question 22. Can the executive summary inform decision-making? 56%

	An executive summary is provided that is of relevant conciseness and depth for primary intended users	Partly
	Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key findings, key conclusions, key recommendations)	Mostly
	Includes all the necessary information to understand the intervention and the evaluation AND does not contain information not already included in the rest of the report	Mostly

•••• Executive Feedback on Section I Overall rating for Section

The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements

Fair

The report presents an executive summary that provides a brief overview of the purpose of the evaluation, its methodology, key findings, conclusions and recommendations. Furthermore, the executive summary only includes information that is developed in further depth in the body of the report. On the other hand, the executive summary does not include a description of UNICEF's humanitarian response at global scale and does not provide a sample of the types of humanitarian work carried out around the globe. Furthermore, the executive summary is too lengthy (15 pages) which by far exceeds the recommended length and challenges the purpose of such a summary.



2

Additional comments for Section I (recommendations for improvement)

Weighting

The idea of an executive summary is to provide high-level users with a succinct yet concise overview of the nature of the object of the evaluation as well as the main elements included in the body of the report. Therefore, the executive summary should be short (6 pages max.) and ensure that all of the elements are provided in order to effectively inform high-level users. In this case, the executive summary should include a relevant description of some (or at least one) of the humanitarian responses that UNICEF carries out in order to exemplify the ways in which their coverage and quality are measured or the challenges in doing so.

0.05

Weighting checksum (should equal 1) 1

OVERALL SCORE (max=4, min=0)

•••• Satisfactory 3

Meets UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with confidence

Lessons for improving the management and performance of future evaluations

This is overall a good quality report and can be used with confidence by decision-makers. It is admirable that UNICEF has undertaken the task of evaluating such a complex and non-traditional object such as the coverage and quality of humanitarian responses. The information contained in this report is supported by a robust and sustained analysis of UNICEF's humanitarian monitoring system that provides important information to contribute towards future strengthening. In the future, the evaluation manager may wish to include a question within the ToRs to identify unexpected effects and the manager may wish to ask the evaluation team to present lessons learned within their own separate section. It is also essential that the Executive Summary remain of a reasonable length. In the future, the evaluation manager should insist that the Executive Summary be kept at no longer than 6 pages.