GEROS Evaluation Quality Assurance Tool

Version 2016.4

<table>
<tr><td colspan="2"></td><td></td><td>Reviewers: complete all cells highlighted in Yellow</td></tr>
</table>

		Notes for QA
Title of the Evaluation Report	**Evaluation of Supply Division's Supply Community Strategy**	
Report sequence number	**2017/001**	
Region	**HQ**	
Year of Report	**2017**	
Office	**Supply Division**	
Coverage (countries)	**HQ/Global**	
ToRs present	**Yes**	
Date of Review	**26/05/2017**	
Name of reviewer	**ImpactReady**	
Classification of Evaluation Report		
Management of Evaluation *(Managerial control and oversight of evaluation decisions)*		**UNICEF managed**
SPOA Correspondence *(Alignment with SPOA focus area priorities)*	**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		**Strategy**
Evaluation type		**Formative**
Evaluation strategy		**Mixed methods**
Evaluation design		**Case study**
Evaluation level		**Impact**
Geographic Scope		**Multi-region/Global**
SECTION A: BACKGROUND (weight 5%)		Rating
Question 1.	Is the object of the evaluation clearly described?	78%
	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status	**Mostly**

	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/cites, sub-regions…) and in terms of numbers reached (as appropriate to the purpose of the evaluation)	**Mostly**
	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?	89%
	Clear and relevant description of the context of the intervention (policy, socio-economic, political, institutional, international factors relevant to the implementation of the intervention)	**Mostly**
	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	**Yes**
Question 4.	Are key stakeholders and their contributions clearly identified?	100%
	Identification of implementing agency(ies), development partners, primary duty bearers, secondary duty bearers, and rights holders	**Yes**
	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF	**Yes**
• • • -	**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**
The report provides a good description of the UNICEF SD, the development of the supply community, and the major pillars of the supply strategy. Information around the practicalities of how the strategy is expected to lead to change (including a description of the interventions) and stakeholder involvement is provided in the annexes (although the report should direct the reader towards this information). The changing strategic function of the SD is well outlined within the UNICEF context. However, a more detailed description of the diverse SD staff and their needs would be useful.		3
Additional comments for Section A (recommendations for improvement)		Weighting
The report could include a sentence to direct the reader to the annexes to find information on the program interventions and stakeholders. Also, a more detailed description of the diversity of the SD staff and their needs would be useful.		0.05

SECTION B: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)	Rating

Question 5.	Is the purpose of the evaluation clearly described?	17%
	Specific identification of how the evaluation is intended to be used and to what this use is expected to achieve	**Partly**
	Identification of appropriate primary intended users of the evaluation	**No**
Question 6.	Are the objectives and scope of the evaluation clear and realistic?	67%
	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	**Mostly**
	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)	**Mostly**
• • - -	**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		**Fair**

The report explains how the purpose of the evaluation is to "document the strategy's achievements and to provide guidance for its amelioration in the years ahead". It would be useful if the report could be more specific about how the findings and recommendations of the evaluation will be used and by whom. Four (4) clear evaluation objectives are presented and appear to be feasible, although quite ambitious. The document does not reference the evaluation scope but it can be assumed that the evaluation covers the entire strategy from its inception.	**2**

Additional comments for Section B (recommendations for improvement)	Weighting
It would be useful if the report could be more specific about how the findings and recommendations of the evaluation will be used and by whom. The report should take the time to explicitly identify the evaluation scope and to mention whether any changes have been made to the original ToRs.	0.05

SECTION C: EVALUATION METHODOLOGY (weight 15%)		Rating
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.	50%
	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	**No**
Question 8.	Does the report specify methods for data collection, analysis, and sampling?	75%
	Clear and complete description of a relevant design and set of methods that are suitable for the evaluation's purpose, objectives and scope	**Mostly**
	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	**Partly**
	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?	**Yes**
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:	50%
	Explicit reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability)	**No**
	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	**Not Rated**

• • - - **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	**Fair**
The evaluation uses a strong mixed methods approach, which includes desk research, a global survey, interviews with staff from different functions/field offices, interviews with comparable international organizations, and participatory workshops with Supply staff. However, the reasons for using this specific approach and data collection tools are not provided. A list of stakeholders interviewed is presented in the annexes but there is no dicussion around the sampling strategy used to identify these particular people. The evaluation uses standard OECD/DAC evaluation criteria but does not explain why the standard criterion of "sustainability" is not included. Ethical considerations regarding ethical safeguards for participants are briefly mentioned but nothing is said about the ethical obligations of the evaluators.	**2**

Additional comments for Section C (recommendations for improvement)	Weighting

The report could be strengthened by providing an explanation as to why a mixed methods approach was used and was appropriate for achieving the objectives of the evaluation. Also, the sampling strategy used to identify stakeholders interviewed should be presented and the report should explain why the standard criterion of "sustainability" was not applied to this evaluation. The report needs a more extensive section on ethics and needs to explain how evaluators followed ethical standards. It is useful to provide a link to the OECD/DAC ethical standards.			0.15

SECTION D: EVALUATION FINDINGS (weight 20%)	Rating
Question 10.	Do the findings clearly address all evaluation objectives and scope?	83%

	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	**Mostly**

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.	**Mostly**
	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	**Partly**
	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?	67%
	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	**Partly**

• • • -	**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**
Findings reference both qualitative and quantitative data and are particularly strong at identifying causal reasons for the accomplishment or non-accomplishment of results. Even though the strategy did not include a results framework, findings still discussed the extent to which outputs have contributed to outcome and impact-level results (although this assessment could have been done more explicitly). While some unexpected effects are discussed within the findings, they are not explicitly identified or labelled as such. Even though the strategy did not contain an M&E system, the report provides a suggested future M&E framework within the annexes. However, the it does not discuss how monitoring data should be used to inform decision making.		3

Additional comments for Section D (recommendations for improvement)	Weighting
The report's assessment of outcome and impact-level change would have been stronger if the evaluation had structured findings around the reconstructed theory of change model. It helps the reader identify unexpected effects if they are included within their own separate and clearly labelled sub-section. Additionally, the proposed M&E strategy would be even stronger if the report mentioned how monitoring data should be used to inform decision making.	0.2

SECTION E: EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 15%)	Rating
Question 13.	Do the conclusions present an objective overall assessment of the intervention?	78%

	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)	**Partly**
	The conclusions are derived appropriately from findings	**Yes**
Question 14.	Are lessons learned correctly identified?	100%

	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.	**Yes**

• • • -	**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**
Conclusions are structured around the evaluation criteria and discuss both strengths and areas for improvement. However, the report does not provide an in-depth discussion around the implications of the findings and the future of the strategy. Lessons learned from other similar initiatives are presented as a learning tool. They are correctly formulated and provide interesting insights that are relevant to the object of evaluation.		3
Additional comments for Section E (recommendations for improvement)		Weighting
Considering that this is a formative evaluation, the report should provide detailed information on the foreseable implications of the findings for the future implementation of the strategy.		0.15

SECTION F: RECOMMENDATIONS (weight 15%)		Rating
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?	83%
	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	**Mostly**

• • • -	**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 findings and conclusions and will likely be useful to the primary intended users. All of the recommendations are clearly targeted to the SD. While the recommendations are numbered, the report does not explicitly state that they are presented in order of priority. The report also does not explain the process followed for developing the recommenations, including the level of stakeholder participation in this process.		3
Additional comments for Section F (recommendations for improvement)		Weighting
It is important for the report to explain the process followed in developing recommendations as well as the level of stakeholder participation in this process in order to help the reader understand the extent to which the recommendations have been validated by stakeholders. Additionally, the report should include a sentence specifying that recommendations are presented in priority order (if this is, in fact, the case).		0.15

SECTION G: EVALUATION STRUCTURE/PRESENTATION (weight 5%)		Rating
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	**Yes**
	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)	**Mostly**

Question 18.	Is the report logically structured?		100%
	The structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles		**Yes**
	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 is well written and logically structured with one section logically preceding the next. The opening pages contain all of the necessary information and the annexes include the ToRs, an evaluation matrix, additional information on the program description and stakeholder involvement, a reconstructed theory of change model, a proposed M&E framework, and copies of the data collection tools. The annexes could be strengthened by also including a list of documents consulted, information on the sampling strategy used to select interview respondents, and information about the evaluators.		3

Additional comments for Section G (recommendations for improvement)	Weighting
The annexes could be strengthened by also including a list of documents consulted, information on the sampling strategy used to select interview respondents, and information about the evaluators.	0.05

SECTION H: EVALUATION PRINCIPLES (weight 15%)		Rating
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?	22%
	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	**No**
	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)	**Partly**
	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.	**Partly**
Question 20.	Does the evaluation assess the extent to which the implementation of the intervention addressed gender, equity & child rights?	22%
	Identification and assessment of the presence or absence of equity considerations in the design and implementation of the intervention	**No**
	Identification and assessment of the presence or absence of gender in the design and implementation of the intervention	**Partly**
	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	**Partly**
	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.)	**Not Rated**
Question 21.	Does the evaluation meet UN SWAP evaluation performance indicators?	Missing requirements
	Note: this question will be rated according to UN SWAP standards	3
	GEEW is integrated in the Evaluation Scope of analysis and Indicators are designed in a way that ensures GEEW-related data will be collected	**Partially integrated 1 point**
	Evaluation Criteria and Evaluation Questions specifically address how GEEW has been integrated into the design, planning, implementation of the intervention and the results achieved.	**Not at all integrated**
	A gender-responsive Evaluation Methodology, Methods and tools, and Data Analysis Techniques are selected.	**Partially integrated 1 point**
	The evaluation Findings, Conclusions and Recommendation reflect a gender analysis	**Partially integrated 1 point**

● - - -	**Executive Feedback on Section H**	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		**Unsatisfactory**

The evaluation does not sufficiently integrate and mainstream GEEW, equity, and human-rights considerations. Human rights are not at all referenced and issues of equity are only at times discussed in the evaluation findings. The global survey asks respondents to specify their sex, making it possible to create sex-disaggregated data. However, differences in responses between men and women are only referenced once in the report. The evaluation matrix does not include any specific questions that would facilitate the gathering of information and analysis of gender and equity-specific issues. The report indicates that a reference group guided the evaluation but the level of stakeholder participation throughout the evaluation process is not at all discussed.

1

Additional comments for Section H (recommendations for improvement)	Weighting
The evaluation should reference the rights of the people intended to benefit from the strategy and should mainstream equity and GEEW considerations throughout the evaluation process and report. This can be done by including specific questions in the evaluation matrix related to how well the strategy is addressing the different needs of women and men and sub-groups. Additionally, the different experiences of women and men should be discussed throughout the evaluation report, drawing on sex-disaggregated data. The report should also explain how stakeholders were involved in the managing and implementing the evaluation, including whether they were consulted during the formulation of the evaluation recommendations.	0.15

SECTION I: EXECUTIVE SUMMARY (weight 5%)	Rating
Question 22. Can the executive summary inform decision-making?	44%
An executive summary is provided that is of relevant conciseness and depth for primary intended users	**Mostly**
Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key findings, key conclusions, key recommendations)	**Partly**
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	**Partly**

• • – – **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 Executive Summary is a good length (6 pages) but does not include all of the necessary elements. The Executive Summary should include more information on the object of evaluation and some information on the operating context. Information on the evalution itself (including purpose, objectives, and methodology) should also be provided. Although the Executive Summary presents findings and recommendations, it is also necessary to present conclusions and lessons learned.	2

Additional comments for Section I (recommendations for improvement)	Weighting
The Executive Summary should include more information on the object of evaluation and some information on the operating context. Information on the evalution itself (including purpose, objectives, and methodology) should also be provided, along with conclusions and lessons learned. The Executive Summary should be a synthesized version of the report and should reflect the primary report sections.	0.05

Weighting checksum (should equal 1)   1

OVERALL SCORE (max=4, min=0)		
• • – –	**Fair**	**2.45**
	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.	

Lessons for improving the management and performance of future evaluations

The evaluation uses a strong mixed methods methodology and presents solid and useful findings. However, a more explicit description of the evaluation purpose (including who will use the evaluation and to what end) needs to be provided as well as a justification for methodological choices, including a sampling strategy to justify which stakeholders were selected for interviews. The report should also justify why the standard criterion of "sustainability" was not included among the evaluation criteria. Additionally, the report should discuss the ethical obligations of the evaluation team, preferrably providing a link to the UNEG Ethical Guidelines. Unexpected findings should be more clearly identified and discussed as well as the implications of the conclusions for future programming (especially considering that this is a formative evaluation). The report does not satisfactorily integrate and mainstream human rights, equity, and gender equality considerations. To do this, the evaluation matrix should include specific questions related to how well the strategy is addressing the different needs of women and men and sub-groups. Additionally, the different experiences of women and men should be discussed throughout the evaluation report, drawing on sex-disaggregated data. Although the Executive Summary is a good length, it needs to include information on the evalution itself (including purpose, objectives, and methodology) along with conclusions and lessons learned. The Executive Summary should be a synthesized version of the report and should reflect the primary report sections. Finally, the report's annexes could be strengthened by including a list of documents consulted, information on the sampling strategy used to select interview respondents, and information about the evaluators.