### GEROS Evaluation Quality Assurance Tool
**Version 2016.4**

<table>
<thead>
<tr>
<th>Reviewers: complete all cells highlighted in Yellow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title of the Evaluation Report</strong></td>
</tr>
<tr>
<td><strong>Report sequence number</strong></td>
</tr>
<tr>
<td><strong>Region</strong></td>
</tr>
<tr>
<td><strong>Year of Report</strong></td>
</tr>
<tr>
<td><strong>Office</strong></td>
</tr>
<tr>
<td><strong>Coverage (countries)</strong></td>
</tr>
<tr>
<td><strong>ToRs present</strong></td>
</tr>
<tr>
<td><strong>Date of Review</strong></td>
</tr>
<tr>
<td><strong>Name of reviewer</strong></td>
</tr>
</tbody>
</table>

### Classification of Evaluation Report

**Management of Evaluation** *(Managerial control and oversight of evaluation decisions)*

<table>
<thead>
<tr>
<th>SPOA Correspondence (Alignment with SPOA focus area priorities)</th>
<th>Health</th>
<th>HIV/AIDS</th>
<th>WASH</th>
<th>Nutrition</th>
<th>Education</th>
<th>Child protection</th>
<th>Social inclusion</th>
<th>Gender equality (cross-cutting)</th>
<th>Humanitarian action (cross-cutting)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Joint managed with organizations</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

**Evaluation object**

- Project

**Evaluation type**

- Summative and formative

**Evaluation strategy**

- Mixed methods

**Evaluation design**

- Case study

**Evaluation level**

- Impact

**Geographic Scope**

- National

---

### SECTION A: BACKGROUND (weight 5%)

**Question 1.** Is the object of the evaluation clearly described? **Rating 56%**

- Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status

**Yes**
<table>
<thead>
<tr>
<th>Question 2.</th>
<th>Is the context of the intervention clearly described?</th>
<th>33%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Partly Clear and relevant description of the context of the intervention (policy, socio-economic, political, institutional, international factors relevant to the implementation of the intervention)</td>
<td>Partly</td>
</tr>
<tr>
<td></td>
<td>Clear and relevant description of the intervention</td>
<td>Partly</td>
</tr>
<tr>
<td></td>
<td>Explanation of how the context relates to the implementation of the intervention</td>
<td>Partly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 3.</th>
<th>Is the results chain or logic well articulated?</th>
<th>83%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes Clear and complete description of the intervention's intended results</td>
<td>Mostly</td>
</tr>
<tr>
<td></td>
<td>Intervention logic presented as a coherent theory of change, logic chain or logic framework</td>
<td>Mostly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 4.</th>
<th>Are key stakeholders and their contributions clearly identified?</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mostl Clear and relevant description of the status and needs of the target groups for the intervention</td>
<td>Mostly</td>
</tr>
<tr>
<td></td>
<td>Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF</td>
<td>Partly</td>
</tr>
</tbody>
</table>

**Executive Feedback on Section A**

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.

**Overall rating for Section A**

<table>
<thead>
<tr>
<th>Weighting</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>Fair</td>
</tr>
</tbody>
</table>

**Additional comments for Section A (recommendations for improvement)**

Il serait bien si le rapport aurait fourni aussi une description de la situation des bénéficiaires. Une description plus ample du contexte politique et socio-économique du Madagascar aurait aussi été envisageable. En outre, il aurait été pertinent aussi de préciser l'importance relative de l'objet pour UNICEF n'est pas précisée.

**SECTION B: EVALUATION PURPOSE, OBJECTIVES AND SCOPE**

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.

**Overall rating for Section B**

<table>
<thead>
<tr>
<th>Weighting</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Satisfactory</td>
</tr>
</tbody>
</table>
**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.

**Rating:** 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

Not Rated

**Question 8.** Does the report specify methods for data collection, analysis, and sampling?

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

Mostly

Clear and complete description of the methods of analysis, including triangulation of multiple lines and levels of evidence (if relevant)?

Mostly

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:

- Explicit reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability)
- 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)

ONLY FOR THOSE CASES WHERE THE EVALUATION INVOLVES INTERVIEWING CHILDREN: explicit reference is made to the UNICEF procedures for Ethical Research Involving Children

**Rating:** 83%

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)

Mostly

**Not Rated**

Executive Feedback on Section C

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

**Overall rating for Section** Satisfactory

Les critères et questions d’évaluation, la méthodologie utilisée, le processus d’échantillonnage et ses limitations sont bien abordés dans cette section. Cependant, la conception de la recherche (étude de cas, comparatif etc.) n’est pas identifiée.

**Additional comments for Section C (recommendations for improvement)**

Le rapport aurait bien fait de décrire la conception de la recherche (research design).

**Weighting** 0.15

**SECTION C: EVALUATION METHODOLOGY (weight 15%)**

**Weighting** 0.05

N/A

**SECTION D: EVALUATION FINDINGS (weight 20%)**
<table>
<thead>
<tr>
<th>Question 10.</th>
<th>Do the findings clearly address all evaluation objectives and scope?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Findings</strong></td>
<td>marshal sufficient levels of evidence to systematically address all of the evaluation’s questions and criteria</td>
</tr>
<tr>
<td>Reference to the intervention’s results framework in the formulation of the findings</td>
<td>Mostly</td>
</tr>
<tr>
<td><strong>Question 11.</strong></td>
<td>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?</td>
</tr>
<tr>
<td>The evaluation clearly presents multiple lines (including multiple time series) and levels (output, outcome, and appropriate disaggregation) of credible evidence.</td>
<td>Yes</td>
</tr>
<tr>
<td>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.</td>
<td>Yes</td>
</tr>
<tr>
<td>Unexpected effects (positive and negative) are identified and analysed</td>
<td>Partly</td>
</tr>
<tr>
<td>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).</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Question 12.</strong></td>
<td>Does the evaluation assess and use the intervention’s Results Based Management elements?</td>
</tr>
<tr>
<td>Clear and comprehensive assessment of the intervention’s monitoring system (including completeness and appropriateness of results/performance framework -including vertical and horizontal logic; M&amp;E tools and their usage)</td>
<td>Partly</td>
</tr>
<tr>
<td>Clear and complete assessment of the use of monitoring data in decision making</td>
<td>No</td>
</tr>
</tbody>
</table>

**Executive Feedback on Section D**

Overall rating for Section D: Fair

Dans l’ensemble, les résultats sont à la hauteur des objectifs de l’évaluation et sont axés sur une analyse fidèle des données. Les indicateurs de performance utilisés sont crédibles et facilement mesurables.

Néanmoins, cet aspect est contrepéché par le manque d’une analyse de la présence ou l’absence de changements imprévus malgré le fait que les critères d’évaluation comprennent des questions pouvant servir à cet effet. De même, le rapport néglige l’examen et l’évaluation approfondis des systèmes de suivi et d’évaluation utilisés dans le cadre du projet.

**Additional comments for Section D (recommendations for improvement)**

Enquêter sur la présence ou l’absence de changements imprévus. Analyser et décrire de manière approfondie les systèmes de suivi et d’évaluation utilisés dans le cadre du projet.

**SECTION E: EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 15%)**

<table>
<thead>
<tr>
<th>Question 13.</th>
<th>Do the conclusions present an objective overall assessment of the intervention?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear and complete description of the strengths and weaknesses of the intervention that adds insight and analysis beyond the findings</td>
<td>Mostly</td>
</tr>
<tr>
<td>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)</td>
<td>Partly</td>
</tr>
<tr>
<td>The conclusions are derived appropriately from findings</td>
<td>Mostly</td>
</tr>
</tbody>
</table>

**Executive Feedback on Section E**

Overall rating for Section E: Fair

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.

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.
**Weighting**

**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?  
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

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.

**Overall rating for Section F**  
Satisfactory

---

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

**Question 17.** Does the evaluation report include all relevant information?  
100%  
Yes

---

**Additional comments for Section F (recommendations for improvement)**

Inclure une description détaillée de l’élaboration des recommandations et du rôle joué par les partenaires.
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? 83%
The structure is easy to identify and navigate (for instance, with numbered sections, clear titles and subtitles)
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

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

Overall rating for Section

Le rapport inclut les informations les plus importantes. La structure est correcte et permet la lecture logique du document. Les annexes sont complètes et utiles.

All 4

Additional comments for Section G (recommendations for improvement)

N/A

Weighting

0.05

SECTION H: EVALUATION PRINCIPLES (weight 15%)

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?

No

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

Partly

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?

No

Identification and assessment of the presence or absence of equity considerations in the design and implementation of the intervention

Partly

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.)

Partly

Question 21. Does the evaluation meet UN SWAP evaluation performance indicators?

Approaches requirements

Note: this question will be rated according to UN SWAP standards

Evaluation Criteria and Evaluation Questions specifically address how GEEW has been integrated into the design, planning, implementation of the intervention and the results achieved.

Partially integrated 1 point

A gender-responsive Evaluation Methodology, Methods and tools, and Data Analysis Techniques are selected.

Satisfactorily integrated 2 points

The evaluation Findings, Conclusions and Recommendation reflect a gender analysis.

Partially integrated 1 point

Executive Feedback on Section H

Overall rating for Section

Highly Satisfactory
Question 22. Can the executive summary inform decision-making?

| An executive summary is provided that is of relevant conciseness and depth for primary intended users | Yes |
| Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key findings, key conclusions, key recommendations) | Yes |
| 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 | Yes |

**Executive Feedback on Section I**

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

| Le résumé exécutif fournit une vision synthétique mais complète des toutes les sections du rapport. | | |

**Additional comments for Section I (recommendations for improvement)**

N/A

| Weighting checksum (should equal 1) | 1 |
| OVERALL SCORE (max=4, min=0) | Fair | 2.4 |
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