



# GLOBAL EVALUATION REPORT OVERSIGHT SYSTEM

*Summary UNICEF Staff Handbook*

unicef   
Evaluation Office

# GEROS Handbook

(Summary Version)

Version 3.2

Cover photo © UNICEF/UN013173



June 2017

UNICEF GEROS Handbook

- This Handbook aims to orient UNICEF staff and partners to the GEROS tool and to the GEROS system more generally.
  - This is part of a wider commitment to aid better understanding of UNICEF’s evaluation function and to support accountability.
  - The handbook helps to set a common standard and level of expectation between UNICEF staff and the independent GEROS assessment team about the GEROS process and the evaluation quality assessment
- Improving evaluation report quality is intended to contribute to:
  - Enhancing the usefulness of evaluations
  - Building UNICEF corporate knowledge and organizational learning
  - Strengthening evaluation capacity for better evaluation in the future.
- Evaluators may also find the Handbook to be useful as a clear articulation of UNICEF expectations regarding a good evaluation report.
  - Preparing evaluation reports that meet UNICEF standards can help ensure completeness, reduce the need for revisions and edits, and improve uptake of recommendations.

# Contents

- Chapter 1: Evaluation Quality Assurance
  - An overview of what evaluation quality assurance is and why it is important to UNICEF
- Chapter 2: GEROS Assessment Process
  - How the GEROS system works and what can be expected
- Chapter 3: Assessment Standards
  - Detailed information on each element of the evaluation quality assessment
- Chapter 4: Review Tools
  - Links to the key GEROS tools and further resources

June 2017

# *How to use this handbook*

Summary Version

- Chapter 1 (an introduction to evaluation quality assessment) helps to better understand the context of GEROS
- Chapter 2 (the GEROS assessment process) gives an overview of the key milestones during evaluation quality assessment and feedback
- Chapter 3 (assessment standards) gives a more detailed look at the standards used to assess the quality of reports
- Chapter 4 (tools) provides a set of links to GEROS templates, including quality checklists

June 2017

# Acronyms

<b>CEE/CIS</b>	Central and Eastern Europe and Commonwealth of Independent States [Region]
<b>CCC</b>	Core Commitments to Children
<b>CEDAW</b>	Convention on the Elimination of all forms of Discrimination Against Women
<b>CO</b>	Country Office
<b>CRC</b>	Convention on the Rights of the Child
<b>EAPR</b>	East Asia and Pacific Region
<b>ERDB</b>	Evaluation and Research Database
<b>EO</b>	Evaluation Office
<b>EPI</b>	Evaluation Performance Indicator (for UN-SWAP)
<b>EQA</b>	Evaluation Quality Assurance
<b>ESAR</b>	Eastern and Southern Africa Region
<b>GEEW</b>	Gender Equality and Empowerment of Women
<b>GEROS</b>	Global Evaluation Report Oversight System
<b>HQ</b>	Headquarters [New York]
<b>HRBAP</b>	Human Rights Based Approach to Programming
<b>LACR</b>	Latin America and the Caribbean Region
<b>MENA</b>	Middle East and North Africa [Region]
<b>PPP</b>	UNICEF Programme Planning Process
<b>OECD-DAC</b>	Organisation for Economic Cooperation and Development – Development Assistance Committee
<b>RBM</b>	Results Based Management
<b>RO</b>	Regional Office
<b>ROSA</b>	Regional Office for South Asia
<b>SPOA</b>	Strategic Plan Outcome Area
<b>TOR</b>	Terms of Reference
<b>UNEG</b>	UN Evaluation Group
<b>UN-SWAP</b>	UN System Wide Action Plan for gender equality
<b>WCAR</b>	West and Central Africa Region

June 2017

*An overview of what evaluation quality assurance is and why it is  
important to UNICEF*

# *CHAPTER I: EVALUATION QUALITY ASSESSMENT*

# The GEROS system

- The purpose of GEROS is to support strengthening of the evaluation function to meet and exceed United Nations Evaluation Group (UNEG) standards, UN System Wide Action Plan on gender equality (UN-SWAP) and other UNICEF commitments (including equity and human-rights based approaches).
  - GEROS aims to ensure accountability and to promote the use of robust evaluative evidence.
- The GEROS system consists of systematic and independent quality assessment of evaluation reports that have been uploaded to the corporate Evaluation and Research Database (ERDB) by UNICEF Country Offices, Regional Offices, HQ divisions and Evaluation Office.
  - Assessments of individual reports are fed back to commissioning offices and publicly available along with the evaluation reports online.
  - Annual meta analysis reports identify global, regional and thematic trends in the quality of evaluation reports.
- GEROS has three objectives (revised in 2016)
  - Objective 1: Enabling environment for senior managers and executive board to make informed decisions based on a clear understanding of the quality of evaluation evidence and usefulness of evaluation reports
  - Objective 2: Feedback leads to stronger evaluation capacity of UNICEF and partners
  - Objective 3: UNICEF and partners are more knowledgeable about what works, where and for who.

June 2017

# Evaluation quality assessment

- The evaluation function, guided by the evaluation policy and PPP, seeks to strategically contribute to UNICEF performance by providing good-quality evidence for:
  - decision making
  - policy advocacy and strategic positioning
  - organizational learning and accountability.
- Reflecting the decentralized nature of UNICEF, the majority of evaluations supported by UNICEF are managed at the decentralized level.
  - The decentralized nature of the evaluation function ensures that evidence generated is relevant to the local, national and regional context and therefore more likely to serve purposes that include informing national policies for children.
  - However, all decentralized evaluation functions pose the challenge of setting up a consistent corporate system to ensure good quality and credibility.
  - The Global Evaluation Reports Oversight System was designed to help address this challenge.
- Evaluation quality assessment is therefore important to:
  - Quantifying the quality of evaluation reports to help inform UNICEF managers (CO, RO, HQ) and partners on one aspect of the current performance of the evaluation function
  - Provide an incentive for evaluation managers, assessors and evaluators to make all efforts to improve the quality of evaluation reports before being submitted to GEROS
  - Provide constructive written feedback to help improve the credibility and use of future evaluations, thus contributing to the professionalization of the evaluation function over time
  - Identify trends in quality over time to inform evaluation strategies and policies
  - Help to identify good evaluation reports to be included in evaluation syntheses

June 2017

# The GEROS principles

- GEROS is an organization-wide system under the Evaluation Policy.
- The EQA (evaluation quality assessment) of final evaluation reports is managed by the Evaluation Office (EO).
- GEROS is complemented by, and independent from, a number of regional quality-assurance mechanisms, which provide assessments of draft ToR and draft reports against the UNICEF/UNEG ToR and Report checklists, giving real time feedback to allow quality improvement of final ToR and reports.
- Some UNICEF offices have established evaluation quality assurance mechanisms that can refer to the GEROS tools and standards.
- GEROS is informed by the UNEG Norms and Standards for Evaluation 2016
- UNICEF-adapted UNEG standards are the basis for quality assessment of final evaluation reports.
- GEROS integrates the UN System Wide Action Plan for gender equality and empowerment of women (UN-SWAP) Evaluation Performance Indicator (EPI).
- To ensure credibility and objectivity, the quality assessments of final evaluation reports is undertaken by an external and independent firm selected through an open bidding process.
- While the corporate Evaluation and Research Database (ERDB) contains evaluations, surveys, studies and researches, GEROS focuses on the assessments of evaluation reports.
- Evaluations are defined by using PPP Manual definitions and the UNICEF evaluation office Taxonomy.
- EO screens reports uploaded to the ERDB, and submits the reports identified as “evaluations” to the external firm for a final screening and evaluation quality assessment.
- EO may reclassify reports as evaluations when they have been misclassified

June 2017

# The GEROS quality rating scale

- Since 2016, GEROS uses a 5 point scale based upon the level to which reports meet assessment criteria
- UN System Wide Action Plan for gender equality (UN-SWAP) criteria are rated according to the methods set by the UN Evaluation Group; with the results integrated into the GEROS rating according to the following table.

Section & Overall Rating	UN-SWAP Rating	Implication	GEROS Score
Highly Satisfactory	Exceeds Requirements (3)	Exceeds UNICEF/UNEG standards for evaluation reports. Decision makers may use the evaluation with a high degree of confidence	4
Satisfactory	Meets Requirements (2)	Meets UNICEF/UNEG standards for evaluation reports. Decision makers may use the evaluation with confidence	3
Fair	Approaches Requirements (1)	Meets UNICEF/UNEG standards for evaluation reports in some regards, but not all. Decision makers may use the evaluation with caution. Substantive improvements in some areas are needed	2
Unsatisfactory	Misses Requirements (0)	Does not sufficiently meet the UNICEF/UNEG standards for evaluation reports. Decision makers cannot rely on the evaluation.	1
Missing		Important aspects of the evaluation that are required by the UNICEF/UNEG standards were found to be absent. The evaluation report is incomplete.	0
Not Rated	Not Applicable	Refers to aspect of the evaluation that was not rated for a legitimate reason. Normally this does not affect the quality of evaluation report.	0

June 2017

*How the GEROS system works and what can be expected*

11

# CHAPTER 2: GEROS ASSESSMENT PROCESS

# Standards for a good quality evaluation report

- An evaluation report is assessed as being of good quality when:
  - It is a credible, independent and useful report that addresses the evaluation purpose and objectives.
  - Can be used with confidence by decision-makers.
- Evaluation reports are assessed using the UNICEF-adapted UNEG evaluation report standards to assess the following core elements:
  - Clear and full description of the background of the evaluation
  - The evaluation’s purpose, objectives and scope are fully explained
  - Appropriate and sound methodology in line with the ethical standards and procedure
  - Findings, conclusions and recommendations are based on evidence, sound analysis and are useful
  - Lessons learned are correctly identified.
  - Well structured, logical and clear report
  - Meet or exceed UN-SWAP evaluation performance indicator for integration of gender equality

June 2017

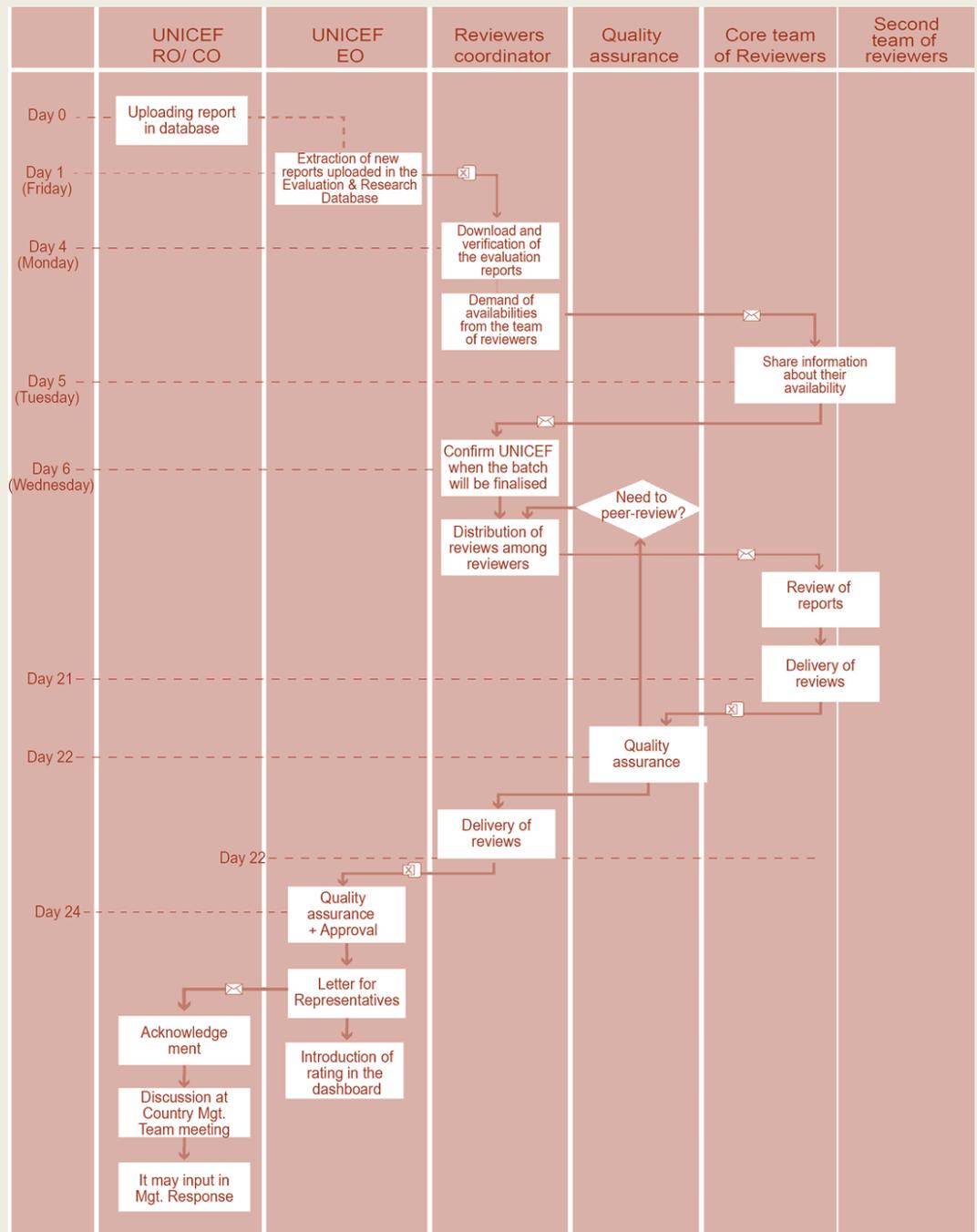
# Standards for a good quality GEROS assessment

- A good quality GEROS assessment (evaluation quality assessment) is:
  - Timely, submitted within the agreed timeframe
  - Complete and in the original language of the evaluation report
  - Accurate and reliable, giving examples where relevant. Does not second-guess the evaluation process itself.
  - Consistent between quantitative and qualitative ratings, and across the evaluation portfolio
  - Constructive in tone and proportionate in the application of standards to the scale and nature of the evaluation being assessed
  - In accordance with the ‘spirit and purpose’ of the GEROS standards. An important consideration in the process that criteria emphasize the quality of elements, not just the extent to which that element is present
  - Honest, independent and quality controlled

June 2017

# Workflow for the assessment process

The workflow for GEROS starts with the uploading of evaluation documents to the ERDB by the section in charge of the evaluation.



June 2017

# Weighting of quality assessment ratings

To ensure consistency of ratings while preserving the freedom for professional assessment by the assessors, the overall rating for an evaluation report is based on a weighted aggregation of the scores from eight EQA 'sections'.

The weightings are set and followed-up at the Annual Review Workshops for GEROS.

Each section is given a quality score (0-4) by an assessor based on a set of guiding questions. Each question has several indicators that are first scored 0-4. The overall section score should be guided by the average ratings of the indicators for each question in that section, but this is not automatic and the assessor maintains professional discrepancy.

The justification for the score is included in written feedback by the assessor; and the EQA tool displays the average scores for each question to support transparency.

Section	Weighting (%)
Background	5
Purpose	5
Methods	15
Findings	20
Conclusions / Lessons	15
Recommendations	15
Structure	5
Principles	15
Executive Summary	5
TOTAL	100

June 2017

# Contents of each assessment

- Background information (meta data)
  - Title of the Evaluation Report; Report sequence number; Region; Year of Report; Office; Countries covered; ToRs present; Date of Review; Name of assessor
- Classification of Evaluation Report (UNICEF taxonomy)
  - Management of Evaluation (Managerial control and oversight of evaluation decisions); Strategic Plan Outcome Areas (SPOA) Correspondence (Alignment with SPOA focus area priorities); Evaluation object; Evaluation type; Evaluation strategy; Evaluation design; Evaluation level; Geographic Scope
- Main assessment
  - Sections A-I; Overall Rating; Feedback on Evaluation Management
- Executive summary (auto-generated)
  - Overall rating; Section ratings; Executive feedback
- Data sheet (auto-generated)
  - All data from the single row for inserting in database for analysis

June 2017

# Style of qualitative comments

The 'house style' for GEROS comments is **constructive and clear**

GEROS assessments are prepared in the **original language** of the evaluation report (Ar, En, Es, Fr, Po, Ru)

**Narrative is aligned with the ratings for consistency.** Comments should use words and structure sentences that avoid dissonance between the comments and the ratings. If a section rates as "fair" then comments should not use a term like "very good" even if the context it is used in is accurate.

June 2017

- **Be a peer, not the police:** use words like "could", "may consider", is "advised to" and avoid instructional words such as "should" unless this is warranted by the complete absence of a UNICEF standard (such as included the ToR as an annex).
- **Be a guide to the UNEG standards:** use phrases such as "the UNICEF/UNEG standards require that..." instead of saying that something is missing or is needed – help the reader to hold themselves to account.
- **Acknowledge constraints and flag things that are contentious:** where the evaluation has clearly faced constraints because of the context, let the reader know that these are acknowledged and understood in the comments, even if the standards require the report to be rated low.
- **Use constructive feedback:** start the comment by stating the indisputable facts. Then explain the implications of those facts on evaluation report quality (with references). Finally, provide guidance on how the issue can be addressed next time. If possible, refer back to a positive example in the same report that can be built on, or give an example of a specific solution to similar challenges in the future.
- **Go beyond the indicators – feedback must add value:** avoid repeating the information from the indicators within the text. Try to elaborate underlying patterns or why the evaluation may have faced difficulties – for example if the recommendations are vague because the purpose was unclear.
- **Try to be specific:** where possible, refer to examples from the report that support your comment. Where there are gaps, share specific ideas for the future (such as "future evaluations of this nature may wish to consider cluster sampling"). Be familiar with the full range of approaches and techniques on [BetterEvaluation.org](http://BetterEvaluation.org) and all UNEG guidance

# EQA reports and follow-up

## *EQA report*

- The EQA process generates 5 products for use by UNICEF staff
  1. A full copy of each independent assessment (Excel and PDF)
  2. An executive summary of each independent assessment (PDF)
  3. A letter to the senior manager responsible for each evaluation explaining the EQA result
  4. The UNICEF Evaluation Dashboard
  5. An annual meta analysis report with feedback for each UNICEF region

## *Management response*

- Management responses should be uploaded within 60 days.
  - Management responses are independent of the EQA process and should not rely on the availability of a completed EQA
  - Management responses are required for all evaluations. Evaluation reports rated as Unsatisfactory by GEROS still require a management response. In these cases, the MR can identify specific findings that will be followed-up or triangulated with other data.

*Detailed information on each element of the evaluation quality  
assessment*

# CHAPTER 3: ASSESSMENT STANDARDS

# Meta-data (used for analysing overall trends)

## Assessment details

- Title of the Evaluation Report
- Report sequence number
- Region
- Year of Report
- Office
- Coverage (countries)
- ToRs present
- Date of Review
- Name of reviewer

June 2017

## Classification of the evaluation

- Management of Evaluation
- SPOA Correspondence
- Evaluation object
- Evaluation level
- Geographic Scope
- UNICEF Taxonomy
  - Evaluation type
  - Evaluation strategy
  - Evaluation design

# Assessment sections, questions and indicators

- ▶ SECTION A
- ▶ SECTION B
- SECTION C
- SECTION D
- SECTION E
- SECTION F
- SECTION G
- SECTION H
- SECTION I

June 2017

## SECTION A: BACKGROUND (weight 5%)

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

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

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)

Description of the relative importance of the object to UNICEF (e.g. in terms of size, influence, or positioning)

**Question 2. Is the context of the intervention clearly described?**

Clear and relevant description of the context of the intervention (policy, socio-economic, political, institutional, international factors relevant to the implementation of the intervention)

Clear and relevant description (where appropriate) of the status and needs of the target groups for the intervention

Explanation of how the context relates to the implementation of the intervention

**Question 3. Is the results chain or logic well articulated?**

Clear and complete description of the intervention's intended results

Intervention logic presented as a coherent theory of change, logic chain or logic framework

**Question 4. Are key stakeholders and their contributions clearly identified?**

Identification of implementing agency(ies), development partners, primary duty bearers, secondary duty bearers, and rights holders

Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF

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

**Question 5. Is the purpose of the evaluation clearly described?**

Specific identification of how the evaluation is intended to be used and to what this use is expected to achieve  
Identification of appropriate primary intended users of the evaluation

**Question 6. Are the objectives and scope of the evaluation clear and realistic?**

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

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)

# Assessment sections, questions and indicators

SECTION A

SECTION B

▶ SECTION C

SECTION D

SECTION E

SECTION F

SECTION G

SECTION H

SECTION I

June 2017

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

**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.

Clear and relevant presentation of the evaluation framework including clear evaluation questions used to guide the evaluation

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

**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

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

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

Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias?

**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

# Assessment sections, questions and indicators

SECTION A

SECTION B

SECTION C

▶ SECTION D

▶ SECTION E

SECTION F

SECTION G

SECTION H

SECTION I

June 2017

## SECTION D: EVALUATION FINDINGS (weight 20%)

Question 10.	Do the findings clearly address all evaluation objectives and scope?
	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions and criteria
	Reference to the intervention's results framework in the formulation of the findings
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.
	The evaluation clearly presents multiple lines (including multiple time series) and levels (output, outcome, and appropriate disaggregation) of credible evidence.
	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.
	Unexpected effects (positive and negative) are identified and analysed
	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).
Question 12.	Does the evaluation assess and use the intervention's Results Based Management elements?
	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)
	Clear and complete assessment of the use of monitoring data in decision making

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

Question 13.	Do the conclusions present an objective overall assessment of the intervention?
	Clear and complete description of the strengths and weaknesses of the intervention that adds insight and analysis beyond the findings
	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)
	The conclusions are derived appropriately from findings
Question 14.	Are lessons learned correctly identified?
	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.

# Assessment sections, questions and indicators

SECTION A

SECTION B

SECTION C

SECTION D

SECTION E

▶ SECTION F

▶ SECTION G

SECTION H

▶ SECTION I

June 2017

## SECTION F: RECOMMENDATIONS (weight 15%)

Question 15. Are recommendations well grounded in the evaluation?

Recommendations are logically derived from the findings and/or conclusions

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)

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

Question 16. Are recommendations clearly presented?

Clear identification of target group for action for each recommendation (or clearly clustered group of recommendations)

Clear prioritization and/or classification of recommendations to support use

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

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

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

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)

Question 18. Is the report logically structured?

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

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

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

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

Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key findings, key conclusions, key recommendations)

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

# Assessment sections, questions and indicators

SECTION A

SECTION B

SECTION C

SECTION D

SECTION E

SECTION F

SECTION G

▶ SECTION H

SECTION I

June 2017

## 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?
	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
	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)
	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.
Question 20.	Does the evaluation assess the extent to which the implementation of the intervention addressed gender, equity & child rights?
	Identification and assessment of the presence or absence of equity considerations in the design and implementation of the intervention
	Identification and assessment of the presence or absence of gender in the design and implementation of the intervention
	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
	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.)
Question 21.	Does the evaluation meet UN SWAP evaluation performance indicators?
	Note: this question will be rated according to UN SWAP standards
	GEEW is integrated in the Evaluation Scope of analysis and Indicators are designed in a way that ensures GEEW-related data will be collected
	Evaluation Criteria and Evaluation Questions specifically address how GEEW has been integrated into the design, planning, implementation of the intervention and the results achieved.
	A gender-responsive Evaluation Methodology, Methods and tools, and Data Analysis Techniques are selected.
	The evaluation Findings, Conclusions and Recommendation reflect a gender analysis

*Links to the key Geros tools and further resources*

# CHAPTER 4: ASSESSMENT TOOLS

# *GEROS Evaluation Quality Assurance template*

Click the icon to open the template,  
or [download here](#)



## Microsoft Excel Worksheet

June 2017

# Evaluation Norms and Standards

## Norms and Standards for Evaluation (2016)

<http://www.unevaluation.org/document/detail/1914>

## UNEG Quality Checklist for Evaluation Reports

<http://www.unevaluation.org/document/detail/607>

## Integrating Human Rights and Gender Equality in Evaluations

<http://www.unevaluation.org/document/detail/1616>

## UNICEF Report Standards Checklist

[https://www.unicef.org/evaluation/files/UNICEF\\_adapted\\_reporting\\_standards\\_updated\\_June\\_2017.pdf](https://www.unicef.org/evaluation/files/UNICEF_adapted_reporting_standards_updated_June_2017.pdf)

## UNICEF TOR Standards Checklist

[https://www.unicef.org/evaluation/files/UNICEF\\_UNEG\\_TOR\\_Checklist\\_updated\\_June\\_2017.pdf](https://www.unicef.org/evaluation/files/UNICEF_UNEG_TOR_Checklist_updated_June_2017.pdf)

June 2017

UNEG Norms 2005	UNEG Norms 2016
Norm 1: Definition	Norm 1: Internationally agreed principles, goals and targets
Norm 2: Responsibility for evaluation	Norm 13: Responsibility for the evaluation function
Norm 3: Policy	Norm 12: Evaluation policy
Norm 4: Intentionality	Norm 2: Utility
Norm 5: Impartiality	Norm 5: Impartiality
Norm 6: Independence	Norm 4: Independence
Norm 7: Evaluability	-
Norm 8: Quality of evaluation	Norm 3: Credibility
Norm 9: Competencies for evaluation	Norm 10: Professionalism
Norm 10: Transparency and consultation	Norm 7: Transparency
Norm 11: Evaluation ethics	Norm 6: Ethics
Norm 12: Follow-up to evaluation	Norm 14: Evaluation use and follow-up
Norm 13: Contribution to knowledge building	-
	Norm 8: Human rights and gender equality
	Norm 9: National evaluation capacities
	Norm 11: Enabling environment

# CONTACTS

**For further help and support**  
[evalhelp@unicef.org](mailto:evalhelp@unicef.org)

**UNICEF GEROS Manager**  
Ada Ocampo

**Handbook prepared by ImpactReady**  
Joseph Barnes | Sara Vaca



A full version of this handbook is also available from the Evaluation Office