
REVIEW OF UNICEF QUALITY ASSURANCE SYSTEMS FOR EVALUATION
AT THE CENTRAL LEVEL

REPORT

Submitted by
Jean Serge Quesnel

December 2019

ACKNOWLEDGEMENT

I am grateful to George Laryea-Adjei, Director of UNICEF Evaluation Office, for having commissioned me to undertake this Review. Also special thanks go to Fabio Sabatini, Senior Adviser – Research & Evaluation of the Evaluation Office, who acted as Review Manager and to Laura Olsen, Evaluation Specialist who was the Review Co-manager.

With appreciation, I also wish to acknowledge the contribution of the other members of the Reference Group. They are Denis Jobin, Senior Evaluation Specialist and Jane Mwangi, Evaluation Specialist of the Evaluation Office.

I wish to express my gratitude for the collaboration and many contributions that I have received from all informants.

Jean Serge Quesnel
Consultant



Table of Content

	Page
Acknowledgement	3
Table of Contents	5
Acronyms Used	6
Executive Summary	7
UNICEF Quality Assurance System at the Central Level	9
UNICEF Framework for Quality Assurance	9
Quality Assurance Model used at the Global level	10
Purpose and Objective of the Review	12
Methodology	13
Findings	14
Relevance	14
Effectiveness	15
Efficiency	18
Sustainability, Cross-cutting and Emerging Issues	21
Good Practices in Quality Assurance in other organisations	22
Conclusions	23
The extent to which Quality Assurance achieved its objective	23
Conceptual Clarity	23
Adequacy and utility of Quality Assurance Approaches, Tools, Mechanisms and Related Guidance	25
Opportunities to strengthen Quality Assurance	25
Key Stakeholders' understanding and perceptions of Quality Assurance	26
Good Practices and Lessons	26
Recommendations	27
Annexes	29
Annex I Term of reference	31
Annex II Logic Model of Requisites to Evaluation of Quality	35
Annex III Key Questions and Data Collection Matrix	47
Annex IV List of Informants interviewed	49
Annex V Interview Protocol	51
Annex VI List of Documents reviewed	53

ACRONYMS

AIMS	Associates for International Management Services
CO	Country Office
EAPR	East Asia and Pacific Region
ECAR	Europe and Central Asia Region
EISI	Evidence Information System Integration
EO	Evaluation Office
EQA	External Quality Assessment
ESAR	East and Southern Africa Region
GEEW	Gender Equality and the Empowerment of Women
GEROS	Global Evaluation Reports Oversight System
HQ	Headquarters
IE	Influential Evaluations
LACR	Latin America and Caribbean Region
LTA	Long Term Agreement
MENA	Middle East and North America
OECD/DAC	Organisation for Economic Cooperation/ Development Assistance Committee
QA	Quality Assurance
QCPR	Quadrennial Comprehensive Policy Review
RO	Regional Office
ROSA	Regional Office for South Asia
SDGs	Sustainable Development Goals
UN- SWAP	UN System-wide Action Plan
UNEG	United Nations Evaluation Group
UNFPA	United Nations Population Fund
UNICEF	United Nations Children’s Fund
UNODC	United Nations Office on Drugs and Crime
VOPE	Voluntary Organisations for Professional Evaluation
WCAR	West and Central Africa Region

1. EXECUTIVE SUMMARY

In 2019, the Evaluation Office has commissioned three reviews pertaining to quality assurance in UNICEF. The first is the External Review of UNICEF's Global Evaluation Reports Oversight System (GEROS). The second is the Review of Quality Assurance at Central Level and the third is the Review of Quality Assurance at the Decentralized Levels.

This is the first time that a review is done of the Quality Assurance System at the Central Level of UNICEF. It examines how the Evaluation Office ensures the quality assurance of the evaluations that it commissions. The final versions of all evaluation reports are included in the Global Evaluation Reports Oversight System (GEROS) and their quality is rated. The assessments, done after finalization of evaluations, are sent to the evaluation managers and are included in an annual meta-evaluation report. The results are made available to the UNICEF Board as part of the annual report on UNICEF's evaluation function.

The purpose of the present review is to strengthen the current Quality Assurance/Quality Control System of the Evaluation Office, per the UNEG standards and UNICEF Evaluation Policy procedures. Therefore, the key objectives of the assignment are:

- a. To map and describe current practices for Quality Assurance within the Evaluation Office with commentary on strengths and weaknesses;
- b. To describe the current practice for a number of comparator organisations both UN entities and others;
- c. To propose options for improved for Quality Assurance based on analysis of the Evaluation Office and others' practices.

To assess the extent to which the stated objectives were achieved, based on specific evaluation questions, the reviewer used a mixed-method approach including document review and analysis (e.g. strategies, guidelines), and interviews with different stakeholders.

The main finding of the Review is there is no standard practice being routinized with regards to quality assurance. Evaluation managers have the responsibility to ensure quality assurance and use different approaches to do it. There was a concern expressed about the degree of objectivity of self-QA. As well, ethical reviews could be more systematized. There is a worry that a rigid quality assurance system would add steps to the evaluation process and cause delays. At HQ, the tendency is to work in silos restricting horizontal sharing of knowledge and evaluation experience. For evaluations conducted by the Evaluation Office, GEROS is a post-factum feedback tool and it could be used better as real-time Quality Assurance feedback.

Based on the conclusions, the Review make eight recommendations for the improvement of quality assurance in the Evaluation Office. They are:

1. The Director of the Evaluation Office should issue a Directive on UNICEF's Quality Assurance management system. The Directive would make clear the rationale and expected outcomes of quality assurance, the accountabilities, roles and responsibilities with respect to all steps of the evaluation process that will be the object of quality assessment.
2. Consideration should be given to ask members of the Evaluation Advisory Panel to serve as external reviewers of evaluations conducted by the Evaluation Office.
3. The use of peer review should be systematized and routinized.
4. The Evaluation Office should look at the alignment of resources dedicated to evaluation considering the 1% indicator contained in the Evaluation Policy and ensure that enough resources are dedicated to quality assurance.
5. The Evaluation Office should develop a Guide on Quality Assurance, grouping together the various existing guides, harmonizing them to strengthen UNICEF's evaluation standards applicable to different types of evaluation.
6. The availability of basic training on evaluation management and quality assurance should be increased across the organisation. Recognition should be made to those who have completed the training programme.
7. The Evaluation Office should create a pool of qualified evaluators who could assist evaluation managers in ensuring evaluation quality. To avoid conflict of interest, long term agreements should be put in place to facilitate access to professional competence dedicated solely to quality assurance.
8. GEROS rating templates should be adapted to different types of evaluation (strategic, thematic, impact, programme, project, rapid and government-led).

2. UNICEF QUALITY ASSURANCE SYSTEM AT THE CENTRAL LEVEL

UNICEF Framework for Quality Assurance

In UNICEF revised Evaluation Policy¹ it is stated that a high-quality evaluation is the outcome of a carefully planned and executed process. It is designed and managed in line with the relevant UNEG standards and, more specifically, the standards set out in the Evaluation Policy. It addresses issues of significance, and evaluation managers build a supporting constituency eager to receive and use it. It should be well resourced and undertaken by a technically excellent team of evaluators.

The revised Evaluation Policy indicates that every evaluation has a designated evaluation manager (e.g., evaluation specialist or staff member in a country office whose functions include evaluation; a regional evaluation adviser; or an evaluation specialist in the Evaluation Office). The designated evaluation manager supervises the selection of the evaluation team, manages the consultants and has the authority to hold them to a high standard of performance. The evaluation manager is ultimately responsible for the quality of the evaluation. Hence, any staff member who has a role in managing evaluations ought to have competence in quality assurance. Noteworthy is the fact the conduct of an evaluation is a shared responsibility, as explained below in the report at section 6.2 Conceptual Clarity.

The revised Evaluation Policy also says that quality assurance should be undertaken for terms of reference, inception reports and evaluation reports and carried out by an external expert or staff member with no role in managing the evaluation. Regional offices provide quality assurance for evaluations managed by the Country Offices, and the Evaluation Office provides such services for those managed by the Regional Offices. The evaluation manager ensures that the appropriate entities carry out adequate quality assurance without incurring delays. Those entities providing quality assurance (i.e., Regional Offices and the Evaluation Office) are responsible for ensuring that recommendations for quality improvement are acted upon. The head of the office managing the evaluation signs off upon completion and presents it to the Evaluation Office for final quality assessment.

The quality of all evaluation reports may be assessed by a specialist external to UNICEF. The Evaluation Office is responsible for ensuring that a credible quality assessment process is established and that its results are shared with the commissioning office and made available to UNICEF management and the Executive Board. Periodically, the Office undertakes reviews of the quality-assessment system at all levels of the organisation.

The use of evaluation is dependent upon the credibility of the evaluation, which in turn is dependent upon rigour, impartiality and professionalism. The value of evaluations is limited unless timely steps are taken to implement their recommendations and incorporate lessons into decision-making systems. UNICEF offices must prepare a formal

¹ Revised evaluation policy of UNICEF, 2018, Executive Board Document E/ICEF/2018/14, paragraph 52

management response to evaluation recommendations and make arrangements to maximize the use of evaluation results. Managers within UNICEF are to create opportunities to feed evaluation recommendations into key decisions. The Evaluation Office maintains a system for Management to record and report on the status of implementation of management response actions.

Annexe II proposes a broad logic model of requisites to evaluation quality assurance. In the logic model, the factors at play are the demand for and offer of evaluations of quality, an enhancing environment for evaluation, a quality-driven institutional framework and competent evaluation professionals. Quality assurance needs to include all the steps of conducting an evaluation. The steps are the scoping and evaluability, the involvement of stakeholders, the terms of reference, the selection of evaluators, the inception report, and oversight of the data collection, the evaluation report and the use of the evaluation.

Quality Assurance Model used in UNICEF at the global level

The Director of the Evaluation Office establishes a comprehensive assessment system to gauge the performance of the evaluation function and the quality and coverage of UNICEF evaluations undertaken throughout the organisation. The Director ensures that evaluations at all levels meet the standards set out in the evaluation policy. The Director maintains systems to monitor and report on management responses to evaluation recommendations and comprehensive databases to facilitate public access to all UNICEF evaluations and management responses.

The Director of the Evaluation Office reports directly to UNICEF Executive Director who safeguards the integrity of the evaluation function, fosters a culture of learning, continuous improvement and accountability, creates demand for and use of evaluation evidence. The Director of the Evaluation Office submits to the Executive Board reports on the implementation of the Evaluation Policy, on the status and effectiveness of the evaluation function. The Executive Board issues decision conveying expectations and guidance on improving performance.

The *Audit Advisory Committee*, which includes an external expert on evaluation, advises the Executive Director on matters pertaining to the oversight of the evaluation, including upholding the norms and standards set out in the evaluation policy. The Director of the Evaluation Office also seeks advice on quality assurance from the *Evaluation Advisory Panel*, an independent, external body, appointed by UNICEF Executive Director. The *Global Evaluation Committee* creates opportunities for the use of evaluation evidence and follows up on the implementation of management responses. The Committee is chaired by a Deputy Executive Director, as assigned by the Executive Director, with membership comprising all other Deputy Executive Directors, relevant divisional directors, the Director of Internal Audit, the Director of the Office of Research and all regional directors.

As well there were externally conducted reviews of UNICEF which include the quality of evaluation. The recent reviews are the following:

2019 – ImpactReady, An Independent Review of UNICEF Evaluation Report Quality and Selected Trends from 2016 - 2018

2019 – AIMS², *External Assessment of UNICEF’s Global Evaluation Oversight System (GEROS)*

2017 – DAC/UNEG Peer Review of the Evaluation Function of UNICEF

2016 – MOPAN³, 2015-16 Assessment of UNICEF

2013 – OECD DAC⁴ Network on Development Evaluation *Review of UNICEF’s Development Effectiveness: 2009 – 2011*

In 2010 the Evaluation Office established the Global Evaluation Quality Assurance System (GEROS). Its purpose is to ensure that UNICEF’s evaluations uphold the high-quality set for them. The GEROS review process and the information provided in annual meta-evaluation report help UNICEF to monitor its progress, identify its strengths and become aware of areas for improvement. GEROS relies on an external independent team of evaluation experts to assess evaluation quality. GEROS performs the role of Quality Control.

To facilitate transparency and access to evaluation reports and their use, the Evaluation Office created a database called the Evidence Information Systems Integration (EISI). It combines the functions of the Planning for Research, Implementation, Monitoring and Evaluation system (PRIME), the Evaluation and Research Database (ERDB) and the Management Response Tracking System (EMR).

Within the Evaluation Office

The Director of the Evaluation Office is accountable for all evaluations commissioned by EO and provides quality assurance oversight. Quality assurance for each evaluation is the main responsibility of evaluation manager who collaborates with stakeholders in doing so. Evaluation managers ensure quality assurance of their own production as well as the work performed by evaluation staff under their supervision. To doing quality assurance, evaluation managers rely on reference groups consisting of UNICEF staff, stakeholders’ representatives and external expertise. Final evaluation reports are submitted to GEROS for quality assessment, control and reporting.

² AIMS – Associates for International Management Services

³ MOPAN – Multilateral Organisations Performance Assessment Network

⁴ OECD-DAC – Development Assistance Committee of the Organisation for Economic Cooperation and Development

3. PURPOSE, OBJECTIVES AND SCOPE OF THE REVIEW

The purpose of the review is to strengthen the Quality Assurance/Quality Control System of the Evaluation Office, as per the UNEG standards and UNICEF Evaluation Policy procedures. Therefore, the key objectives of the assignment are:

- d. To map and describe current practices for Quality Assurance within the Evaluation Office with commentary on strengths and weaknesses;
- e. To describe the current practice for a number of comparator organisations both UN entities and others;
- f. To propose options for improved for Quality Assurance based on analysis of the Evaluation Office and others' practices.

Evaluation Office Quality Assurance/Quality Control System is designed at the planning and reporting stage. For the reporting stage, the GEROS system is well established and has been the subject of a separate review. The present review aims at Quality Assurance/Quality Control that are done during the planning phase of an evaluation (e.g. for Terms of Reference or Inception Reports), during implementation (conducting the evaluation and the QC check points during field visits, data collection, etc.) and the finalisation of the evaluation (quality of the report, use of the findings and recommendations).

The review covers the Evaluation Office's work in corporate evaluations, including humanitarian evaluations. It does not include the quality assurance systems for evaluations at the Decentralized levels of UNICEF, which is the subject of another review.

4. METHODOLOGY

The review intends to map the current Quality Assurance systems existing in the Evaluation Office, examine the various Quality Assurance mechanisms being used and how they have contributed to the improvement of the quality of evaluation products, and what the final outcomes are in terms of quality of evaluations. The review also examines the role of the actors providing required quality assurance services, in accordance with the provisions of the UNICEF evaluation policy. Lessons are drawn from best practices from comparator evaluation functions.

The review relied heavily on the reporting produced by the Global Evaluation Reports Oversight System (GEROS), the official documentation of UNICEF and other organisations of the United Nations System and the views of UNICEF evaluation officers involved in the quality assurance processes.

Within UNICEF, in addition to the Executive Board, as stakeholders there are:

- The Evaluation Office;
- Headquarters Programme Managers, both whose programmes are evaluated or who are expected to use evaluation results to improve existing programmes or design new ones;
- Regional Office staff who are commissioning or using the results of evaluations;
- Country offices that have commissioned or undertaken evaluations.

Also, the review is of interest to the Audit Advisory Committee, the Evaluation Advisory Panel and the Global Evaluation Committee.

Document review

The document review examined the annual meta-evaluations, the annual evaluation reports to the Executive Committee, the decisions of the Executive Committee, the Revised UNICEF Evaluation Policy, the UNICEF Procedure on the Implementation of the 2018 UNICEF Evaluation Policy, various UNICEF evaluation guides. The documents listed in Annex VI were also reviewed.

Interviews

All Regional Evaluation Advisors were interviewed, as well as evaluation officers of the Evaluation Office. The interviews were scheduled ahead of time. Skype (or phone call) was used and lasted approximately 60 minutes each. As can be seen from the semi structured interview protocol in Annex V, interviews covered a broad range of topics from role and use to satisfaction with and recommendations for improving quality assurance processes.

Ethical Considerations

The review followed the ethical standards and code of conduct set out in the UNEG ethical guidelines for evaluators. This included ensuring the confidentiality of all information obtained during the data collection process. Triangulation of data ensured conclusions are not influenced by any one party.

Limitations of the review

There were no major limitations to the review. The key elements, mapping the current Quality Assurance systems and obtaining data on their use, were possible. There was a short fall in having access to detailed information on financial resources invested in Quality Assurance.

5. FINDINGS

5.1 Relevance

Finding 1.1 – The coverage of evaluations conducted by the Evaluation Office is relevant.

In 2018, the Evaluation Office commissioned evaluations⁵ that assessed objects of evaluation covering a wide range of thematic areas, with humanitarian action and education comprising the greatest percentage of the region's portfolio (at 14 percent each) followed by health (13 percent); gender equality (12 percent); child protection (12 percent); WASH (12 percent); social inclusion (10 percent); nutrition (8 percent); and HIV/AIDS (5 percent). The thematic breakdown of evaluations commissioned by the Evaluation Office in 2018 was overall quite similar to the global average. However, EO commissioned nearly twice as many humanitarian evaluations. The 2018 portfolio was comprised of a generally even breakdown per 2018 – 2021 Strategic Plan goal area, with the largest percentage (20 percent) of evaluations assessing cross-cutting priorities.

In 2018, 46 percent of EO evaluations were programme evaluations, which was in line with the global average (49 percent). The office commissioned a much higher percentage of thematic evaluations compared to the global average, with 36 percent (4/11 evaluations) of its evaluation profile dedicated to this area (compared to 5 percent at the global level). The office also commissioned one strategy evaluation and one system evaluation.

⁵ Source: UNICEF GEROS Meta-Analysis Headquarters (2018)

The Evaluation Office also focused more on learning with 91 percent of its portfolio consisting of either formative and formative and summative evaluations compared to the global average of 72 percent.

Finding 1.2 – Quality Assurance does improve the evaluation deliverables

There is a consensus amongst evaluation professionals of the Evaluation that Quality Assurance does improve the evaluation deliverables. There is recognition that Quality Assurance is a sound investment.

5.2 Effectiveness⁶

Finding 2.1 – The effectiveness of the Quality Assurance System at the Central Level needs to be closely monitored and strengthened.

The effectiveness of the Quality Assurance System at the Central Level can be exemplified by the ratings given by GEROS to the evaluation reports commissioned by the Evaluation Office.

Table 1. below illustrates the results.

Table 1: Percentage of Reports per Assessment Rating for HQ from 2016 – 2018

Ratings	2016	2017	2018
Highly Satisfactory	12.5%		9.1%
Upper Satisfactory	37.5%		54.5%
Lower Satisfactory	37.5%	50%	27.3%
Upper Fair	12.5%	50%	9.1%
Lower Fair	---	---	---
Unsatisfactory	---	---	---

Source: GEROS Meta-Analysis 2018 – EO Profile

By looking at a breakdown of evaluations per assessment rating, we can see from the Table 1. above, 91 percent of the evaluations commissioned in 2018 by the Evaluation Office were rated as satisfactory or higher, which constitutes a 41 percent increase from the previous year. Also, whereas no evaluations rated in the upper satisfactory and highly satisfactory categories in 2017, 64 percent of EO reports did in 2018.

⁶ This section relies heavily on the GEROS Meta-Analysis 2018 Profile for Headquarters

If we look at the assessment criteria used by GEROS to rate the evaluation report we see that the Evaluation Office was strongest in 2018:

- at presenting the evaluation purpose, objectives and scope (3.36/4.0),
- at providing a complete and informative background section (3.27), and
- at presenting actionable recommendations (3.18) as well as complete yet succinct executive summaries (3.18).

The Evaluation Office had globally the highest average rating for evaluation principles (3.0/4.0) in 2018. On the other hand, globally reports from the Office ranked 2nd lowest, just above ESAR, at presenting the evaluation methodology (2.73). This was the lowest scoring criterion for EO in 2018 followed by conclusions and lessons learned (2.82).

With regards to the UN System-Wide Action Plan for Gender Equality and Empowerment of Women (SWAP), the Evaluation Office globally ranked 5th, with an average score of 64%, which is approaching SWAP requirements. SWAP has three criteria (Criteria 1 – Scope and Design; Criteria 2 – Methodology and Criteria 3 – Gender Analysis). If we look closer at the evaluations commissioned by the Evaluation Office, we find that they struggled most with SWAP criterion 2 by only partially integrating SWAP principles into the use of a gender responsive methodology, methods and tools, and data analysis techniques. EO reports ranked 2nd lowest globally for this criterion (1.55), only above ROSA (1.44), and 8.3 percent below the global average for SWAP criterion 2. On average, the Evaluation Office only satisfactorily integrated SWAP principles into criteria 1 (evaluation scope and design) and 3 (gender analysis) at an average score of 2.09 each.

Finding 2.2 – Overall strengths and weaknesses⁷

In sum, the overall strengths in 2018 of the evaluations commissioned by the Evaluation Office are:

- Clear Evaluation Purpose, Objectives and Scope: EO evaluations in 2018 were generally strong at clearly laying out evaluation objectives and at identifying specific evaluation users and uses (i.e. UNICEF departments and government ministries as opposed to UNICEF or national governments in general). They also generally provided a good description of both what is and is not included in the evaluation scope in terms of thematic, geographic, and time-bound

⁷ Source: UNICEF GEROS Meta-Analysis 2018 Headquarters Profile

characteristics.

- **Background Section:** Reports in EO generally presented strong background sections that provided a good description of the object of the evaluation along with an informative explanation of the national/regional context in which the object of the evaluation operated.
- **Actionable Recommendations:** EO evaluations in 2018 were generally strong at developing recommendations that were correctly drawn from the findings and conclusions and that were actionable, as they identified the stakeholder groups for action in each case. Recommendations were also duly prioritized.
- **Strong Executive Summary:** HQ executive summaries were in general complete, thorough and yet succinct with a length of not more than 6 pages to effectively inform decision-makers.

In 2018, the overall weaknesses of the evaluations commissioned by the Evaluation Office are the following:

- **Methods:** While most EO evaluations appear to use appropriate evaluation designs and methods, reports often do not adequately explain why these designs and methods are well suited for meeting the evaluation purpose and objectives. Additionally, reports often do not explain why evaluations diverted away from the standard OECD/DAC evaluation criteria. Evaluation limitations are usually identified but not all reports provide mitigation strategies to address the limitations. Even though reports are becoming stronger at describing how the evaluation followed ethical principles, it is common for evaluation reports to inadequately describe the ethical obligations of the evaluators.
- **Conclusions:** In general, EO reports could benefit from conclusions that present a more forward-looking analysis that goes beyond the information presented in the findings and that discusses potential future implications.
- **Lessons Learned:** EO reports in general struggled to provide correctly identified lessons learned that contributed to the organisation's institutional knowledge and whose relevance could make them applicable to similar initiatives taking place or projected to take place in other contexts. In some cases, evaluations did not identify any lessons learned whatsoever. In some other cases, lessons learned were confused with best practices that primarily highlight success stories as opposed to critical moments of learning from what did not work well. In total,

only 36.4 percent of EO evaluation reports presented satisfactory lessons learned.

5.3 Efficiency

Finding 3.1 – The model currently used by the Evaluation Office for quality assurance is neither systematized nor routinized

In theory, the design of the quality assurance system in place in the Evaluation Office should answer to current needs. However, there is no standard practice being followed with regards to quality assurance. In collaboration with other stakeholders involved in the evaluation process, Evaluation managers share the responsibility to ensure quality assurance and use different approaches to do it. There is a concern expressed about the degree of objectivity of self-QA. As well, ethical reviews could be done more systematically. In sum, Quality Assurance should be more systematized and routinized. On the other hand, there is a worry that a rigid quality assurance system would add steps to the evaluation process and cause delays.

Finding 3.2 – There is no compendium of guides and tools to facilitate quality assurance.

The list of guidance tools mentioned during the interviews was:

Guidance tool	Usefulness Rating Base 5	Times mentioned N=11
UNEG Norms and Standards (2016)	5	8
UNICEF Adapted UNEG Quality Checklist for Evaluation Terms of Reference (2017)	4	6
UNICEF Adapted UNEG Evaluation Reports Standards (2017)	4	7
UNEG Quality Checklist for Evaluation Reports (2010)	3	3
UNEG Improved Quality of Evaluation Recommendations – Checklist (2018)	4	4
Guidance Note for Conducting Evaluability Assessment in UNICEF (2019)	3	1
UNEG Integrating Human Rights and Gender Equality in Evaluations (2015)	3	2
GEROS Assessment Templates	5	9
UNICEF Terms of reference for Management and Governance (Humanitarian) (2016)	4	1
UNICEF Humanitarian Policies	3	2

ALNAP guides	3	1
University Evaluation Text Books	4	4
Literature reviews	3	1
Country visits to be able to contextualize.	4	1

As shown above, there was no consistency in the reference of the existence of guidance tools available nor were they cited as being used routinely. Most interviewees mentioned three or four of the above. They were the Geros checklists and UNEG adapted guides.

Finding 3.3 – There is quite a difference of staff time dedicated to quality assurance

The amount of time dedicated to quality Assurance by the professional staff of the Evaluation Office varies significantly. The percentage of time reported during the interviews is as follows:

Staff time dedicated to Quality Assurance:

Interviewee 1	10%
Interviewee 2	70%
Interviewee 3	100%
Interviewee 4	70%
Interviewee 5	30%
Interviewee 6	40%
Interviewee 7	40%
Interviewee 8	Not/Applicable
Interviewee 9	20%
Interviewee 10	70%

This points to the fact that there is no common perception and understanding of the roles and responsibilities concerning quality assurance. Some interviewees conveyed the impression that they held themselves fully responsible for quality assurance whereas others indicated that they share the responsibility with others involved in the evaluation process.

Finding 3.4 – Strengths and weaknesses of the current practices of Quality Assurance in the Evaluation Office

Some strengths of the current Quality Assurance practices in the Evaluation Office are:

- Quality Assurance is fluid and flexible. As a system, QA is not mechanical. Evaluation officers must exercise judgement.
- Use of Reference Groups that include UNICEF staff, external experts and stakeholders.
- Transparency with colleagues in programme divisions (ex.: EMOPS, Child Protection, WASH)
- There is continuous thinking about quality assurance during the process of evaluation, yet no standard approach is followed.
- There is recognition that to have good quality, there is a need of having competent professionals involved in the evaluation process. Thus, critical is the choice of the evaluation consultants.

Some weaknesses of the current Quality Assurance practices are:

- There are not many *highly satisfactory* evaluations.
- Quality Assurance could start with the question: Are we doing the right evaluation? There is a need for Quality Assurance when determining strategically the choice of evaluations to be conducted.
- The Quality Assurance process is mostly hierarchical and not sufficiently horizontal.
- Informal consultations do not yield a rigorous system of Quality Assurance. Some rely on network of friends. More complex evaluation warrant more quality assurance. Ad hoc QA comments are not enough. Within EO, there is a lack of a set of *second eyes* for Quality Assurance.
- There is no incentive, recognition and accountability for a Quality Assurance contribution toward an evaluation conducted by a colleague of the Office. A voluntary model is not sustainable.
- The "*busyness*" is such that it is not conducive to engaging the community of stakeholders in quality assurance, sharing and use of the evaluation. After the completion of the evaluation little time is dedicated to its dissemination. There is a need to have more buy-in for quality assurance concerning the use of evaluations.
- In the Evaluation Office there is the conviction that not enough time is available for quality assurance.
- EO staff competencies in Quality Assurance are quite uneven.

Finding 3.5 – Cost of external reviewers

External reviewers are hired on a limited basis. They are sectoral and thematic specialists. Their expertise focuses on the substantive part of the evaluation less so on the evaluation methodology. Their rates are high according to international specialist's standards.

5.4 Sustainability, Cross-cutting and Emerging Issues

Finding 4.1 – Lack of peerage

Within the Evaluation Office, the tendency is to work in silos restricting horizontal sharing of knowledge and evaluation experience. There is no incentive, recognition and accountability for a Quality Assurance contribution toward an evaluation conducted by a colleague of the Office. A voluntary model is not sustainable.

Finding 4.2 – There is a missed opportunity of using Quality Assurance for evaluation capacity building

Sharing knowledge is not *a forte* in the Evaluation Office. Evaluation Capacity Building is mostly done through mentoring. In Reference Groups, quality assurance enables thoughtfulness. Quality Assurance does not add up as evaluation capacity building if used as check against standards. Quality Assurance is not intentionally designed for evaluation capacity building. Nevertheless, ECB may occur as a result of the experiencing QA.

Within EO, GEROS is not used to its best ability as a tool for evaluation capacity building. If GEROS became a real-time feedback mechanism, it would contribute to ECB.

Finding 4.3 – Usefulness of GEROS

GEROSisation of evaluation highlights the compliance to report standards of one evaluation product, the final evaluation report. It is not wholesome looking at quality considering the interim products, nor does it include the usefulness of evaluations.

For EO-led evaluations, GEROS is a post-factum feedback tool and it could be used well as real-time Quality Assurance feedback tool.

5.5 Good Practices in Quality Assurance in other Organisations

ILO has done a series of i-eval Think Pieces on evaluation quality assessment. ILO has undertaken several initiatives to improve the quality of evaluation reports, such as the Evaluation Manager Certification Programme.

The Office of Evaluation of the **FAO** performs biennial ex-post quality assessment of evaluations purposely selected to have a balance regarding geographic and thematic coverage. FAO assesses the quality of the Terms of Reference and the overall evaluation process in addition to the evaluation report.

In order to ensure an evaluation report's quality and consistency, **UNDP** Independent Evaluation Office has engaged a pool of evaluation and thematic experts to undertake the quality assessment. These QA reviewers are senior development professionals with a detailed knowledge of evaluation in their thematic areas, an understanding of UNDP strategic approach in general and as well as in evaluation and have strong regional and country level knowledge and experience. To ensure evaluation quality assessment uniformity and consistency, the pool of QA reviewers is oriented to the application of quality assessment tools. The Independent Evaluation Office periodically verifies the quality assessment process to ensure inter-rater reliability.

UNFPA has set up its evaluation quality assurance and assessment system (EQAA) using overall and internationally agreed evaluation principles: Independence & Impartiality, Credibility and Utility.

The Independent evaluation unit of **UNODC** contracts external consultants to systematically assess, at the end of the year, the quality of all published evaluation reports and a third of the first draft evaluation reports. In this way, the process of revision of draft reports is also assessed.

UN-Women Global Evaluation Report Assessment and Analysis System (GERAAS) provides written justifications for scores that are given. This reduces subjectivity in a Quality Assurance system that uses qualitative judgements to determine scores.

WIPO includes in its evaluation manual a comprehensive and detailed checklist for assessing the quality of evaluation reports.

WFP provides a decision tool giving guidance on choosing either an evaluation or a review in view of purpose and expected use.

The **World Bank** is presently examining its quality assurance system to make its formal process simpler.

6. CONCLUSIONS

6.1 The extent to which Quality Assurance achieved its objective

The quality of evaluation reports has varied over the years and has improved in 2018.

The following table shows the results achieved by the Evaluation Office over the years

Table 1

Percentage of reports rated satisfactory or highly satisfactory from 2011 to 2018

Year	Number of Evaluation	Percentage rated satisfactory or above	Global Rating
2011	7	80%	42%
2012	16	100%	62%
2013	7	100%	69%
2014	9	100%	74%
2015	4	53%	50%
2016	10	80%	74%
2017	5	60%	72%
2018	11	91%	87%

Source: GEROS Meta-Analysis 2018 Headquarters Profile

The trend analysis from 2011 until today indicates that the quality of EO evaluation reports was well above the global average from 2011 – 2014 (with 100 percentage of evaluations rated as satisfactory or highly satisfactory in 2012, 2013, and 2014) until the overall quality was halved in 2015, reaching a low of only 50 percent of evaluations rated satisfactory or higher. Report quality then fluctuated considerably from 2016 to 2018 by rising to 80 percent in 2016, dropping to 60 percent in 2017, and then again rising to 91 percent in 2018. The percentage of evaluations rated satisfactory or higher commissioned by EO was 4 percent above the global average in 2018.

6.2 Conceptual Clarity

Quality Assurance and Quality Control

Four interviewees said that the distinction between Quality Assurance & Quality Control is unclear. Only two could describe the distinctions as being:

Quality Assurance

During the evaluation process assess if the steps and products meet the standards. The main steps being:

- 1- Clarify what is to be evaluated.
- 2- Engage stakeholders.

- 3- Assess resources and evaluability.
- 4- Determine the evaluation questions.
- 5- Determine appropriate methods of measurement and procedures.
- 6- Develop an evaluation plan.
- 7- Collect the data.
- 8- Process data and analyze results.
- 9- Interpret and disseminate results.
- 10- Apply the evaluation findings.
- 11- Gate keeping throughout the process: watch for deviations from standards.

Quality Control

Examine if the evaluation report reflects compliance to all the standards. Geros is used for this purpose.

There is no great volume of evaluations conducted annually by the Evaluation Office. De facto, for the interviewees, the distinction between QA and QC is fuzzy. For the Evaluation Office, a big Quality Control system is perceived as not being useful due to the small number of evaluations being conducted.

When doing Quality Assurance there is a need for a balance between substantive thinking and compliance to standard.

Responsibility and Accountability

During the interview process there was a strong propensity to merge the concepts of responsibility and accountability. They are distinct.

Responsibility

The state of having the duty, to do whatever it takes to complete the task, is known as responsibility. Responsibility expects a person to be reliable and dependable to complete the tasks assigned to him/her. Responsibility is something assigned to you by someone in a higher authority. Responsibility can be shared with another. Responsibility can be before and/or after a task.

Accountability

The condition, wherein a person is expected to take ownership of one's actions or decisions, is called accountability. Accountability requires a person to be liable and answerable for the things, he/she does.

Accountability comes with the individual's own acceptance of their obligation. Accountability is answerability, blameworthiness, liability, and the expectation of account-giving. Accountability emphasizes the sole ownership of the actions and decisions. Accountability is something you hold a person to only after a task is done or not done.

In the context of UNICEF, an evaluation manager is responsible for the evaluation. However, the accountability rests with the Office Manager who commissions and approves the final report.

6.3 Adequacy and utility of Quality Assurance Approaches, Tools, Mechanisms and Related Guidance

The design of the current Quality Assurance system of the Evaluation Office is suitable. However, there is a need for greater systematisation of its use.

As revealed by Finding 3.2, there is no compendium of guides and tools to facilitate quality assurance.

Quality Assurance is an important subject yet there are few opportunities to talk about it. There is lots of pressure to deliver evaluation in a short time frame and hence little time is dedicated to Quality Assurance per se.

6.4 Opportunities to strengthen Quality Assurance

There are many options that the Evaluation Office could strengthen & routinize Quality Assurance & Quality Control

The Office could prepare a Directive on Quality Assurance to explain expected outcomes of QA, to clarify roles and responsibilities. EO could make clear the role of a senior Quality Assurance reviewer who would oversee the production of all EO evaluations. The EO Directive on QA should be simple, to-the-point, such as the one issued by DRP on Quality Assurance in Research

There should be recognition that there is not only one way to do Quality Assurance. Therefore, there is a need to adapt Quality Assurance to different types of evaluations and contexts.

There ought to be recognition that professional time has to be allocated for Quality Assurance. Greater use of Peer review within EO would be beneficial. As well, there is a need for more systematic ethical reviews

During the conduct of EO led evaluations, GEROS could be used as a real-time feedback tool.

It would be useful if the Evaluation Office had Guidance on the hiring of a competent evaluation team.

In the development of the new online evaluation training programme a module could be dedicated to evaluation quality assurance. For new evaluation staff, an induction kit on quality assurance would be beneficial.

6.5 Key Stakeholders' understanding and perceptions of Quality Assurance

The roles, responsibilities and accountabilities for Quality Assurance are unclear. Evaluation managers are left alone in performing Quality Assurance. Quality Assurance oversight is perceived as not being performed. Quality Assurance is seen differently depending on the roles being played in an evaluation such as team member, as evaluation team lead, as evaluation manager.

Sector and thematic specialists are more concerned about the substantive content of the evaluation and do not pay much attention to the quality assurance dimensions of the evaluation approach.

6.6 Good Practices and Lessons

Regarding Quality Assurance systems/processes used by comparator organisations and how they compare with UNICEF's practices, four interviewees said that they are not familiar with QA systems in other organisations.

Some referred to UNDP, UNFPA, and WFP. Only one interviewee could describe the Quality Assurance systems existing in the organisations.

Perception is that in other organisations, evaluation staff is involved in quality assurance and does not rely on external reviewers to the extent that UNICEF does.

7. RECOMMENDATIONS

1. The Director of the Evaluation Office should issue a Directive on UNICEF's Quality Assurance management system. The Directive would make clear the rationale and expected outcomes of quality assurance, the accountabilities, roles and responsibilities with respect to all steps of the evaluation process that will be the object of quality assessment.
2. Consideration should be given to ask members of the Evaluation Advisory Panel to serve as external reviewers of evaluations conducted by the Evaluation Office.
3. The use of peer review should be systematized and routinized.
4. The Evaluation Office should look at the alignment of resources dedicated to evaluation considering the 1% indicator contained in the Evaluation Policy and ensure that enough resources are dedicated to quality assurance.
5. The Evaluation Office should develop a Guide on Quality Assurance, grouping together the various existing guides, harmonizing them to strengthen UNICEF's evaluation standards applicable to different types of evaluation.
6. The availability of basic training on evaluation management and quality assurance should be increased across the organisation. Recognition should be made to those who have completed the training programme.
7. The Evaluation Office should create a pool of qualified evaluators who could assist evaluation managers in ensuring evaluation quality. To avoid conflict of interest, long term agreements should be put in place to facilitate access to professional competence dedicated solely to quality assurance.
8. GEROS rating templates should be adapted to different types of evaluation (strategic, thematic, impact, programme, project, rapid and government-led).

ANNEXES

Annex I	Terms of reference
Annex II	Model of Requisites to Evaluation of Quality
Annex III	Key Questions and Data Collection Matrix
Annex IV	List of informants interviewed
Annex V	Interview Protocol
Annex VI	List of documents reviewed

ANNEX I - TERMS OF REFERENCE

TERMS OF REFERENCE UNICEF EVALUATION OFFICE QUALITY ASSURANCE SYSTEMS EXTERNAL ADVISOR

2. INTRODUCTION

UNICEF Evaluation Office (EO) is seeking to recruit an individual consultant who would examine and document models of quality assurance (QA) for evaluation (as applied at the central level), gauging the relevance and usefulness of those models in UNICEF EO's operating context and providing options for a QA function. This assignment is intended to generate information that the EO can use to guide its decision-making and contribute to higher quality evaluations. This document outlines the scope of the assignment for an individual consultant who will work under the guidance of the Director, EO. Successful candidates will have experience in the management of central evaluation units and familiarity with QA approaches and tools.

3. BACKGROUND

Enhancing quality is a key objective for the EO along with increasing the number and timeliness of evaluations. The UNICEF Revised Evaluation Policy (2018) sets expectations for the organization in regards to QA (e.g. should be undertaken for terms of reference, inception reports and evaluation reports and carried out by an external expert or staff member with no role in managing the evaluation). Regional offices provide QA for evaluations managed by the country offices, and the EO provides such services for those managed by the regional offices. Individual evaluation managers ensure that the appropriate entities carry out adequate QA without incurring delays. Those entities providing QA (i.e., regional offices and the EO) are responsible for ensuring that recommendations for quality improvement are acted upon.

A post-hoc quality assessment process (Global Evaluation Report Oversight System or GEROS) is in place for all UNICEF evaluation reports which relies on an external contractor to provide independent assessment. Through GEROS, the EO is responsible for ensuring that a credible quality-assessment process is established and that its results are shared with the commissioning office and made available to UNICEF management and the Executive Board. UNEG standards on QA systems give the responsibility to the head of evaluation to "ensure that there is an appropriate QA system.

Typically invoked at the design and finalization stages of evaluation, an appropriate QA mechanism looks at both the evaluation process and its products. Depending on the construct of the evaluation function, the mechanism can be operated with internal peer review or external review. In either case, the head of evaluation should ensure the objectivity of the review. Alternatively (or additionally), QA could be provided by an internal or external expert providing guidance and oversight throughout the evaluation process. (UNEG 5.1)" and UNEG standard 4.2 "Evaluation guidelines the head of evaluation is responsible for ensuring the provision of appropriate evaluation guidelines. Although guidelines may need to be

prepared for different types of evaluations or for different types of users, the guidelines should generally cover: The roles and responsibilities in setting up, managing, conducting, quality controlling, reporting and disseminating evaluations; The process of evaluation; Stakeholder involvement; Guidance on methodologies and quality control; and Reporting, dissemination and the promotion of learning.”

Further accountabilities for QA are outlined in the UNICEF Procedure on the Implementation of the 2018 UNICEF Evaluation Policy. These are:

- EO will provide technical guidance for purposes of enhancing the quality of evaluations and related quality assurance processes at all levels of the organization
- The EO should actively support capacity development for staff at all levels of the organization – in technical skills as well as in professional conduct
- Heads of Offices and Evaluation Specialists should ensure that the provisions of the Evaluation Policy for QA are enforced. In accordance with technical guidance issued by the EO, Regional Offices should strengthen quality assurance for evaluations managed by the Country Offices, and the Evaluation Office should do the same for those managed by the Regional Offices
- The EO should also continue to ensure that the system for providing feedback to Offices on the quality of evaluations (the Global Evaluation Reports Oversight System - GEROS) is periodically reviewed and remains credible.

4. PURPOSE, OBJECTIVES AND SCOPE OF THE ASSIGNMENT

The purpose of the assignment is to strengthen the actual EO QA/QC system, as per the UNEG standards and UNICEF Evaluation Policy procedures. Therefore, the key objectives of the assignment are:

- a. To map and describe current practices for QA within the EO with commentary on strengths and weaknesses
- b. To describe the current practice for a number of comparator organizations both UN entities and others
- c. To propose options for improved for QA based on analysis of EO and others’ practices. Options should include: consideration of resource requirements, timeline and sequencing of required steps.

Evaluation Office QA/QC is designed at the planning and reporting stage. For the reporting stage, the GEROS system is well established and currently under review. The consultant is not expected to review the GEROS system but will rather use the findings for the ongoing review and liaise with the Review GEROS team as necessary. This review is rather aimed at the forms of QA/QC that are provided during the planning phase of an evaluation (e.g. for Terms of Reference or Inception Reports) or during implementation (managing and role of manager and QC check points during field visits, data collection, etc.) and can be most immediately used for quality improvement.

The review should cover those parts of the Evaluation Office’s work in corporate evaluations, including humanitarian evaluations (L3-L3-L1) and its support for the work of the Regional Offices.

5. KEY QUESTIONS AND METHODS

Specific questions for examination will be developed jointly between the EO and the consultant. Overall, the main inquiries include the following:

- What practices are most commonly employed for QA within the UNICEF EO?
- What are strengths, weaknesses, threats and opportunities in regards to the existing EO QA practices?
- What QA systems or processes are used by comparator agencies? How do those systems compare to practices within the EO?
- What options should the EO consider to strength and routinize QA and quality control system?
- Which staff should be in charge of the QA process? What skills and competencies are required?

The review will include all aspects of the Evaluation Office QA functions including: a) for global corporate or thematic operations, b) for humanitarian evaluations, and c) for EO support to RO for QA of regionally-conducted evaluations. Comparator organizations should include other UN agencies (e.g. UNDP, UNFPA, WHO, WFP, ILO, UN Environment, etc.) as well as non-UN entities (e.g. GEF, Bill and Melinda Gates Foundation, Rockefeller Foundation, IFRC). The consultant will work closely with the EO to identify key variables for comparison between these organizations.

6. MANAGEMENT AND CONDUCT

The review will be conducted by an external consultant to be recruited by UNICEF EO. The consultant will operate under the direct supervision of a Senior Evaluation Officer, supported by an Evaluation Specialist. The Evaluation Office will be responsible for the day-to-day oversight and management of the review.

PROFILE

The review will be conducted by an individual consultant. That consultant is expected to have the following credentials:

- Strong management track record and commitment to delivering timely and high-quality products
- Extensive expertise (at least 10 years) in the management and oversight of Central evaluation units of large multi-lateral organizations
- Extensive experience with contemporary QA approaches and tools, ideally with more than one entity
- Familiarity with UNICEF's programming, policy and advocacy work and experience in evaluating multi-sectoral initiatives would be an asset
- Good interpersonal and communication skills; ability to interact with various stakeholders and to concisely express ideas and concepts in written and oral form.

DELIVERABLES

The consultant will be expected to submit:

- An implementation plan for the review inclusive of methods, tools and variables for comparison as well as a workplan and timeline
- A draft review report inclusive of options for review and discussion within the EO
- A final review report which takes EO input into consideration
- A presentation of the findings and proposed options

BUDGET AND TIMING

The review should be carried out between July and October 2019. It is estimated that the consultant could complete the assignment within 25 working days.

ANNEX II – LOGIC MODEL OF REQUISITES TO EVALUATION OF QUALITY

1. An appropriate Context is necessary for evaluations of quality.

The context conducive to evaluation quality is determined by a culture of learning and accountability, meaning the degree to which information is sought about past performance, the extent to which there is a drive to continuously improve and to be responsible for actions taken, resources spent, and results achieved. Such culture is embedded in tacit norms and behaviour, the understanding of what can and should - or should not – be done. Through a set of values and attitudes supporting evaluative (critical) thinking within an organisation, individuals use information to act; to take higher risks but, also to develop a greater sense of personal accountability and responsibility; and to consult and support each other.

In a macro-framework fostering the quality of evaluation five key factors⁸ are at play. They are:

A Demand for quality evaluation: This is the capability of policy-makers and senior managers to request sound and trustworthy evidence with the aim of using it in strategic decision-making processes.

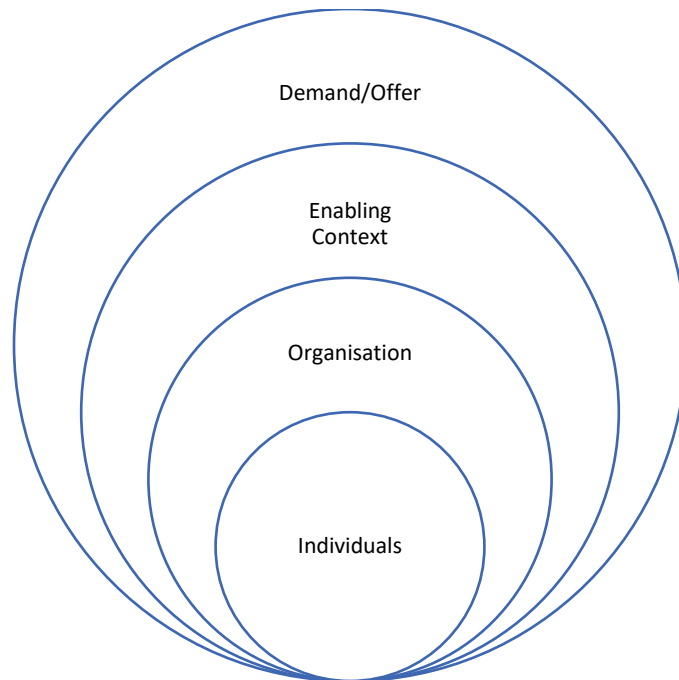
A Supply of quality evaluation: It is the capacity of professional evaluators to provide sound and trustworthy evaluative evidence.

An Enabling Environment that provides a context that fosters (or hinders) the performance and results of organisations and individuals.

A Quality driven Institutional Framework that provides a system and structures to perform and attain results individually as well as collectively as an organisation.

Competent individuals whose knowledge and skills are essential to perform and manage processes and relationships.

⁸ Source: UNICEF, Evalpartners, IOCE (2013) Evaluation and Civil Society Stakeholders' perspectives on National Evaluation Capacity Development



1.1 Capacity to demand and use information from evaluation⁹:

- Capacity within government institutions and civil society organisations to incorporate and use information from evaluation as part of the normal process of business (e.g. capacity to critically gauge evaluative evidence, to access timely evidence etc.).
- Governments and civil society stakeholders are clear about where and how evaluation information can and will be used within government (e.g. planning, policy or programme development, decision-making, budgeting). This can evolve over time.
- Policy-makers, government bodies, and CSOs have an appreciation of evaluation concepts and use of evaluation information.
- Adequate incentives within organisations and the system to ensure that evaluation information is used, and that evaluations report credible information in a timely fashion.
- Reinforces the need within organisations for formal or informal mechanisms and forums for reporting and sharing evaluation information.
- Laws governing access to information would increase transparency and the potential for evaluation information to be made available to the media, civil society among others and facilitate their participation in the national system.
- The technical capacity and infrastructure to undertake evaluation.

⁹ Source: Adapted from: UNEG Task Force on National Evaluation Capacity Development. (2012). National Evaluation Capacity Development: Practical tips on how to strengthen National Evaluation Systems.

1.2 Capacity to supply information from evaluation:

- Availability of skilled personnel to gather, analyze and report on the quality, value and importance of different levels and types of performance of government policies and programmes, including potential partners within the country, such as universities, research institutes, think tanks, among others.
- A national statistical agency to facilitate a national data development strategy and assist ministries and agencies in capturing and storing data.
- Existence of credible and relevant data (disaggregated by sex, age, ethnicity, etc.) and information-gathering systems.
- Infrastructure to ensure a systematic, comprehensive and credible approach to evaluation. This would include policies and standards intended to: clarify roles, responsibilities and accountabilities for performance monitoring and evaluation; establish expectations across the system for evaluation, monitoring and timing, and a high level of performance reporting; and, set out quality standards for conducting evaluations.
- Organisational structures to conduct and/or manage evaluation exercises.
- A policy center to provide policy direction, oversight and assistance for the system-wide development of evaluation.

1.3 Enabling Environment¹⁰: the enabling environment provides a context that fosters (or hinders) the performance and results of individuals and organisations

- Strong evaluation culture:
 - evaluative (critical) thinking;
 - protective environment;
 - managers and other decision-makers value and use evaluation.
- Public administration committed to transparency and managing for results and accountability, through results-based public budgeting and evidence-based policymaking.
- Political will to institutionalize evaluation.
- Existence of adequate information and statistical systems.

¹⁰ Source: Adapted from EvalPartners, 2010, Moving from policies to results by developing national capacities for country led monitoring and evaluation systems.

- Legislation and/or policies to institutionalize monitoring and evaluation systems.
- Duty bearers, i.e. Governments and national authorities, have capacity and willingness to be accountable for results.
- Rights holders i.e. citizens and CSO, have capacity to demand that policy and programmes are monitored and evaluated.
- National VOPE exists, has the necessary capacities, is recognized and given a consultative role.
- National evaluation standards and norms developed and implemented

1.4 Institutional Level: the institutional framework in which individuals work needs to provide a system and structures to perform and attain results individually as well as collectively as an organisation

- Evaluation policy exists and is implemented.
- An evaluation unit with a clearly defined role, responsibilities and an independent place in the institutional structure exists and is functional.
- Quality Assurance system exists and is functional.
- Independence of funding for evaluations.
- Adequate (number & expertise) staffing of the evaluation function.
- System to plan, undertake and report evaluation findings in an independent, credible and useful way exists.
- System to establish a formal management response to the recommendations of evaluations feeding back to policies and programmes, follow-up of implementation, exists.
- Open dissemination of evaluation results.
- Knowledge management systems in support of the evaluation function exist and are used.
- Technical competence for conducting evaluations exists.

1.5 Individual Level: the individual whose knowledge, skills and competences are essential to perform task and manage processes and relationships

- Senior management capacity to:
 - strategically plan evaluations;

- assure the development of relevant and appropriate Terms of Reference for evaluations;
 - manage evaluation for independence and credibility;
 - promote the use of evaluation findings;
 - follow up recommendations.
- At mid-management level, understanding of the role of evaluation as a tool for effectively achieving development results.
 - Identify and support leaders or natural champions.
 - Behavioral independence and professional competences of those who manage and/ or conduct evaluations.
 - Promote capacity development and involvement of a growing number of evaluators.

2. What is the difference between quality assurance and quality control?

The International Organization for Standardisation (ISO)¹¹, an independent, non-governmental organisation with a global network of national standards bodies with one member per country has issued ISO 9000 which defines quality assurance and quality control systems.

Quality Assurance is the process which guarantees that the preferred level of quality is fulfilled at each phase of the process. *Quality Control* ensures that the end-product quality is, as per the set standards.

Quality Assurance is proactive focussing on the process, preventing defects. Quality Control is reactive focussing on the end-product ensuring that it meets the standards. Quality Assurance is prevention- based system whereas Quality Control a detecting-based system. The "quality assurance" and "quality control" terms are often used interchangeably. Quality control, however, has its own specific function, and it pertains to the actual fulfillment of the expected quality requirements.

In UNICEF, Quality Assurance is performed by evaluation managers, supported by their hierarchical supervisors, by peers and by reference groups involving stakeholders. Quality Control of the evaluation reports is managed by the Evaluation Office who uses a corporate system, the Global Evaluation Reports Oversight System (GEROS).

3. What are the elements of a quality assurance system?

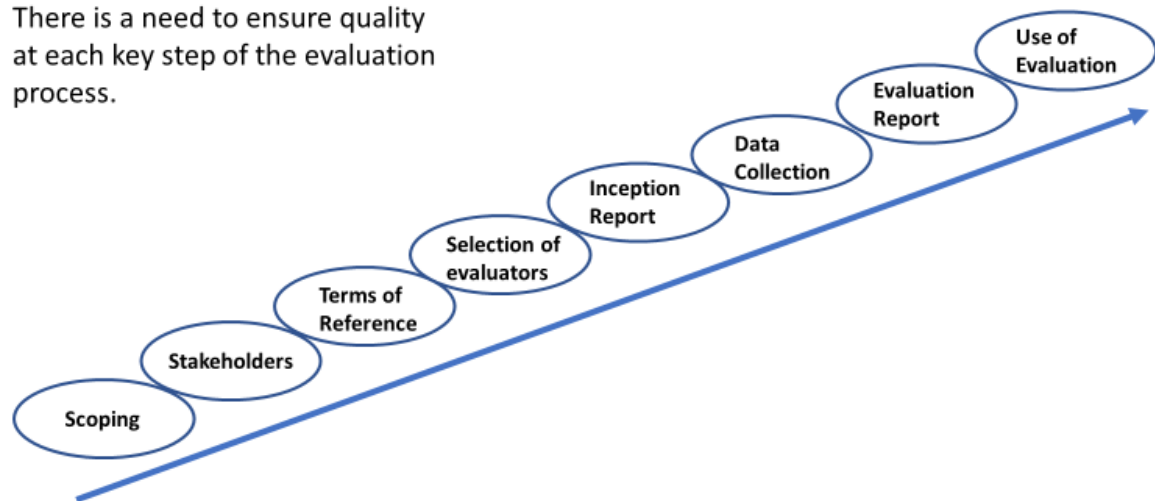
The purpose of a quality assurance system is the maintenance of a desired level of quality, especially by means of attention given to every stage of the process of production of an

¹¹ <https://www.iso.org/home.html>

evaluation. In other words, Quality Assurance encompasses the processes and procedures that systematically monitor the different aspects of conducting an evaluation according to professional standards. Quality assurance focus on having a good process to produce evaluations with the quality already built-in, rather than going through an unmonitored production process and trying to “inspect quality” to an evaluation that’s already been finished.

Quality during the evaluation process

There is a need to ensure quality at each key step of the evaluation process.



3.1. To have an evaluation of quality, one must ask the right question at the outset.

To have evaluations of quality, one needs to scope well the evaluations so that they are truly used and useful for decision making. There are six lessons learned in this regard:

1. *Be open and keep it simple.* Evaluation must be made public and understandable, so that there is pressure to actually use them.
2. *Where's the beef?* Do the evaluations contain information that can be used by decision makers? If final conclusions for impact and action cannot be made, can we still learn other lessons in the process?
3. *Timing is everything.* Evaluations should be done and completed at strategic times so that decision makers have the necessary data when they are making decisions. Incomplete but timely information may be more valuable for decision making than conclusive post-mortems.

4. *Make evaluations constructive.* Evaluations are more effective when they are a support to get interventions right rather than as an inquisitional judgment. People must believe in the power of evaluations, so they won't hide 'unfavorable' or 'negative' information during the evaluation process.
5. *Ask the right questions.* When creating evaluations, start by determining what is needed by decision makers. Identify the relevant research questions first, and then methodology and technology can follow.¹²

As well, an evaluability assessment may be the right thing to do.

3.2 Stakeholders Involvement

In order to make the evaluation relevant and useful, the involvement of stakeholders is necessary. Their involvement will allow the scoping to respond to their needs. It will open doors to access to information and will facilitate participation for information gathering. Transparency with stakeholders will give credibility and foster rigour of the evaluation process. It will ensure buy-in of the findings and of the improvement measures being proposed. Ultimately, it is the Stakeholders who will make use of the evaluation.

There are various degrees of stakeholder involvement, from discreet consultation with each stakeholder, to open interactive forum including all.

A convenient way of involving stakeholders is to include stakeholder representative in the reference group of the evaluation.

3.3 Terms of reference

A Terms of Reference (TOR) is a plan or blueprint outlining the key elements of the purpose, scope, process and products of an evaluation, including management and technical aspects as necessary. Developing a TOR is a critical early step in any evaluation. In the narrowest sense, it is the basis for contractual arrangements with external consultants, though it should first be developed as a means of clarifying expectations, roles and responsibilities among different stakeholders, providing the plan for overall activity, including follow-up. The time and effort spent in preparing a good TOR has big returns in terms of the quality relevance and usefulness of the product.

TORs are important:

For all stakeholders – as they explain the agreed expectations in terms of the parameters and process of the exercise, and are a guide to each stakeholder's specific role;

¹² <http://blogs.worldbank.org/governance/value-evaluations-asking-right-questions> by Mario Marcel

For the evaluation team – as they ensure that expectations are clear and the objectives are met;

For external and internal teams – as external teams may require more detail on background context and on intended audiences and uses; internal teams may simply need to clarify the parameters of the assignment.

TORs are often developed in stages. Stakeholders will focus on the purpose and evaluation questions. A further developed version is used for recruiting external consultants. It requires more detail on existing information sources, team composition, procedures and products, and may describe methodology and a calendar of activities in broad terms. The TOR may then be further refined once an evaluation team is on board, with a careful review of the purpose and key questions and corresponding methodology.

A key reference to ensure the quality of Terms of reference is the guide *UNICEF-Adapted UNEG Quality Checklist for Evaluation Terms of Reference (Updated June 2017)*

3.4 Selection of the evaluators

If you hire professionals, typically either a consulting firm or a group from a university, they'll have ways of doing research, prejudices about what an evaluation should cover, and interests, all of which may or may not coincide with what you need or want. Stakeholder and in-house evaluators also, depending upon the individuals involved and the groups they represent will have their own prejudices, interests, and needs which will influence how they view the evaluation process and the evaluation itself. The same goes for volunteers who might donate their time to conduct or assist with an evaluation. It is important to select people whose agenda matches your own, or who will put their own agendas aside and respond directly to the needs and desires of the organisation. Only then will you have an evaluation that best serves your purposes.

SOME CHARACTERISTICS TO LOOK FOR

Willingness to leave one's own agenda

Often researchers, particularly those attached to universities, may have their own reasons for embarking on an evaluation. It may fit into a doctoral dissertation or a book that a professor is writing, or a piece of long-term research that will eventually be published. They may also have very strong prejudices about what kind of research they want to use, or what they expect to find, and a need to prove their prejudices correct. And they may have issues about power and about their standing in relation to that of members of a planning team or organisational staff. If the researchers' needs mesh perfectly with yours, then there's no problem. But if they don't quite fit, there can be a very serious problem. If you're paying for a service, you should get what you're paying for, and not simply what the researcher wants to give you. It's important to be clear about this on all sides at the beginning, and to make sure by writing it into a contract or through some other formal mechanism that the professionals are willing to do what meets your needs, not what meets theirs.

Ability to communicate with a broad range of people

Evaluators will have to deal with people from all walks of life, of all political, religious, and philosophical persuasions, and probably of many ethnic, language, and racial groups. If they are to gather accurate information, they will have to be perceived as trustworthy by all those groups and will have to be able to generate a certain level of comfort with everyone.

Cultural sensitivity

In multi-cultural contexts, mutual respect, understanding and acceptance of how others see the world is crucial to the functioning of evaluators. Do they understand, or are they willing to learn to understand, the cultures of others? Assess the candidates' prior experiences working with the sectors and populations you serve and their understanding of the socioeconomic, demographic, historical, cultural and political factors that shape the conditions in which the strategy, initiative or program is operating. This means ensuring culturally competent evaluators - those who understand how various groups perceive an intervention, communicate their views and experiences. Culturally competent evaluators are particularly effective because they: a) Keep an open mind. B) Avoid making assumptions about a group of people. C) Understand how their own cultural background, biases and worldviews could affect their interactions with program participants and other stakeholders. d) Gather appropriate data to draw conclusions by using methods that are respectful of other cultures. v Look for candidates whose experience suggests the capability to devise and manage a variety of evaluation designs and tasks, and who can clearly articulate which design is most appropriate for your effort.

Ability to treat everyone with the same degree of respect

How evaluators approach people reflect on UNICEF. If they don't treat everyone respectfully, they're not going to get accurate--or any--information and they're going to complicate UNICEF relationship with Partners and stakeholders.

Absolute commitment to keeping all individuals' information confidential

Whether evaluators are paid professionals or not, for ethical, practical, and legal reasons, it is necessary to guarantee that any information gathered in the course of an evaluation will be kept confidential. This will help evaluators obtain more accurate data. It will protect the evaluators and the organisation from lawsuits that could be brought by individuals injured in some way by the information they provided.

Commitment to the evaluation process

This means trying to do the best evaluation possible, with an eye toward its actual usefulness for the organisation. Commitment means believing enough in the process to take the evaluation seriously and use it to improve the policy, strategy, program, and project.

Ability to understand the purpose of the evaluation

It's important to understand not only what the evaluation is supposed to assess, but also how its form and purpose are related.

Willingness to listen and learn

This covers a wide spectrum of people and behaviors. It also includes everyone's being aware that they have things to learn about the process, one another, and evaluation in general,

and, by the same token, that they have valuable knowledge, skills, and information to contribute.

Ability to work in a group

This can be a concern regarding groups that encompass several cultures and classes.

Knowledge of different kinds of evaluation techniques

Evaluators must have the capability to devise and manage a variety of evaluation designs, and to clearly articulate which design is most appropriate. They use both quantitative (numbers and statistics) and qualitative (facts, stories, anecdotes, analysis of situations and events, etc.) research, and know when each is appropriate.

Ability to do Quality Assurance

External reviewer firms must have own internal Quality Assurance processes

Past performance

Have they done evaluation before? What kinds of techniques have they used? Do they have good references? Can you see examples of past evaluations they've done?

In summary

The individual or team you select to conduct the evaluation will do much to determine the usefulness of the evaluation itself. If you work with professionals, you need to consider their credentials, experience, and level of expertise. Whether you choose professional evaluators, volunteers, or some combination of the two, you need to think carefully about how their needs and interests fit in with those of UNICEF, about the range and quality of their communication skills, about their willingness to work as partners, and about their fit with one another. Once you've put together a good team, you are well on your way to carrying out an accurate and valuable evaluation.

3.5 Inception Report

An inception report of an evaluation is prepared by an evaluator after an initial review of relevant documentation. It sets out the conceptual framework to be used in an evaluation, the key evaluation questions and methodology, including information on data sources and collection, sampling and key indicators. The inception report also includes a timeline for the evaluation and drafts of data collection instruments.

The inception report describes the object of evaluation and states the purpose and scope of the evaluation. It will address the evaluation framework; discuss the overall approach and the risks and limitations. The evaluation methodology will be specified. The data collection and data analysis methods will be explained. The data sources will be presented including the sampling methods. An evaluation matrix will show how the evaluation questions will be answered, using what sources of information. A work plan will be detailed, together with the matrix of responsibility. It will give a timeline and discuss logistics for the conduct of the evaluation. In appendices there will be the draft data collection instruments.

A reference for inception reports is the *Guidelines for Inception Reports* by United Nations Educational, Scientific and Cultural Organisation (2008, March).

3.6 Data Collection

There are many ways to collect data, some more costly and time-consuming than others. Therefore, before starting collecting data, it is imperative to consider the following questions:

What are the evaluation questions? Which methods will help answer them and provide the most reliable and valid data? Who is providing the data? Who are the potential respondents? How many respondents are needed? How much time is budgeted for data collection and analysis? When is the right time to collect the data? What is the evaluation budget? How much of it is for data collection and analysis versus reporting and dissemination of findings?

Depending on the data required to answer the evaluation questions, quantitative methods such as surveys or assessments could be right. Or qualitative methods such as interviews or focus groups might suit better the needs for information.

3.7 Evaluation Report

A key reference to ensure the quality of Evaluation Report is the guide UNICEF-Adapted UNEG Evaluation Reports Standards (Updated June 2017). It highlights the following questions:

- Is the report well structured, logical, clear and complete?
- Does the report present a clear and full description of the 'object' of the evaluation?
- Are the evaluation's purpose, objectives and scope fully explained?
- Does the report present a transparent description of the design and methods used in the evaluation that clearly explains how the evaluation addresses the evaluation criteria, yields answer to the evaluation questions, and achieves evaluation purposes?
- Do the findings respond directly to the evaluation criteria and questions detailed in the scope and objectives section of the report; and are they based on evidence derived from data collection and analysis methods described in the methodology section of the report?
- Do conclusions present reasonable judgements based on findings and substantiated by evidence and provide insights pertinent to the object and purpose of the evaluation?
- Are recommendations relevant to the object and purpose of the evaluation, are they supported by evidence and conclusions, and were they developed with involvement of relevant stakeholders?

3.8 Use of evaluation

Recommendations are usually linked to the concept of utility as they are used for improving projects, programmes, policy or strategy. Recommendations are intended for use as a basis for management decisions to improve performance, results or impact. Recommendations are made to suit the needs of the decision-makers that will be using them. As a result, recommendations in a mid-term evaluation can look quite different to those in an evaluation of an intervention that is close to its end.

The United Nations Evaluation Group's (UNEG) created a checklist of expectation and requirements to improve the quality and utility of evaluation recommendations. The guide entitled *Improved Quality of Evaluation Recommendations Checklist* was published in June 2018.

Four main areas affect a recommendation's development and ultimately its use. These are: a) the process of formulating recommendations; b) the structure of the recommendations; c) the content of the recommendations; and d) the implementation and follow-up of the recommendations.

ANNEX III – KEY QUESTIONS AND DATA COLLECTION MATRIX

Key Questions	Sources of information			
	Documentary review	UNICEF documents	GEROS	Interviews
What practices are most commonly employed for quality assurance at within the Evaluation Office?		XX		XX
What are the strengths and weaknesses of existing quality assurance practices?				XX
What quality assurance systems or processes are used by comparator organisations?	XX			XX
How do comparator organisations' systems compare to UNICEF's practices?				XX
What options should the Evaluation Office consider to strength and routinize the system of quality assurance and quality control?				XX
Which staff should oversee the quality assurance process?				XX
What skills and competencies are required?	XX			XX

ANNEX IV - LIST OF STAKEHOLDERS INTERVIEWED

From the Evaluation Office

Denis JOBIN	Senior Evaluation Specialist
Mariel KISLIG	Evaluation Specialist
Kathleen LETSHABO	Evaluation Specialist
Lovemore MHURIYENGWE	Knowledge Management Specialist
Jane MWANGI	Evaluation Specialist
Ada OCAMPO	Senior Evaluation Specialist
Laura OLSEN	Evaluation Specialist
Beth Ann PLOWMAN	Senior Evaluation Specialist
Fabio SABATINI	Senior Adviser – Research & Evaluation
Adrian SHIKWE	Evaluation Specialist
Tina TORDJMAN-NEBE	Evaluation Specialist

From the Regional Offices

Lori BELL	ECARO	Europe and Central Asia Regional Office
Samuel BICKEL	ROSA	South Asia Regional Office
Mirella HERNANI	LACRO	Latin America and Caribbean Regional Office
Urs NAGEL	ESARO	Eastern and Southern African Regional Office
Riccardo POLASTRO	EAPRO	East Asia and Pacific Regional Office
Robert STRYK	MENA	Middle East and North Africa Regional Office
Michele TARSILLA	WCARO	West and Central Africa Regional Office

Other informants

Josef VAESSEN	Adviser on evaluation methods – World Bank
Anne-Marie DAWSON	Evaluation Specialist – Rater - UNIVERSALIA

ANNEX V - INTERVIEW PROTOCOL

UNICEF

Review of Quality Assurance Systems for Evaluation at the Central Level Interview Guide

1. Purpose of the Review

The Review commissioned by UNICEF Evaluation Office will:

- i) examine and document models of quality assurance (QA) for evaluation, as applied at the central level,
- ii) gauge the relevance and usefulness of those models in UNICEF EO's operating context and
- iii) provide options for a Quality Assurance function.

This Review intends to generate information for the UNICEF evaluation function to use to improve its quality assurance processes and, ultimately, the quality of evaluations as well as their timeliness. It aims at quality assurance, from the planning for an evaluation through the final evaluation report. It does not include the quality assessment system (GEROS) which was subject of a separate review nor does it include the quality assurance system at the Decentralized levels, also subject of another review.

2. Objectives of the Review

The objectives of the Review is to strengthen the actual EO QA/QC system, as per the UNEG standards and UNICEF Evaluation Policy procedures. Therefore, the key objectives of the assignment are:

- a. To map and describe current practices for QA within the EO with commentary on strengths and weaknesses
- b. To describe the current practice for a number of comparator organisations both UN entities and others
- c. To propose options for improved for QA based on analysis of EO and others' practices. Options should include: consideration of resource requirements, timeline and sequencing of required steps.

Evaluation Office QA/QC is designed at the planning and reporting stage. For the reporting stage, the GEROS system is well established and was the subject of recent review. The present review is aimed at the forms of QA/QC that are provided during the planning phase of an evaluation (e.g. for Terms of Reference or Inception Reports), during implementation (managing and role of manager and QC check points during field visits, data collection, etc.) and can be most immediately used for quality improvement.

The review will include all aspects of the Evaluation Office QA functions including:

- a) for global corporate or thematic operations,
- b) for humanitarian evaluations, and
- c) for EO support to RO for QA of regionally-conducted evaluations.

Main Questions

1. What practices are commonly employed for quality assurance within UNICEF Evaluation Office?
2. What are the reference documents used for doing quality assurance?
3. What are the strengths and weaknesses of existing QA practices within UNICEF Evaluation Office?
4. To what extent quality assurance processes change/improve the evaluation deliverables?
5. Has quality assurance been useful for evaluation capacity enhancement?
6. What are the staff time required for quality assurance systems and the cost of external reviewers?
7. What are the roles and responsibilities of the key players in the QA processes?
 - a. Which staff should be in charge of the quality assurance function?
 - b. What skills and competencies are required for quality assurance?
8. What do you think should be the complementarities between the quality assurance processes and quality control system in the Evaluation Office and the equivalent processes and system existing at the Regional Offices Level?
9. What quality assurance systems or processes are used by comparator organisations? How do those systems compare to UNICEF's practices within UNICEF Evaluation Office?
10. What options should the Evaluation Office consider in order to strengthen and routinize its quality assurance processes and quality control system?

ANNEX VI - LIST OF DOCUMENTS REVIEWED

ALNAP (2016) *Evaluation of Humanitarian Action Guide*

ALNAP, HS Alliance (2016) *Self-assessment tool: Core Humanitarian Standard on Quality & Accountability*

ALNAP (2003) *Quality Proforma*

Centre for Development and Emergency Practice (CENDEP) and USAID (2018) *Good practice guide: embedding inclusion of older people and people with disabilities in humanitarian policy and practice*

CGIAR (2018) – *Guidance on finalization, Feedback and Decision Making*

CGIAR (2015) *Standards for Independent External Evaluations*

CGIAR (2012) **Policy** - *CGIAR External Evaluation*

CLEAR - Regional Centers for Learning on Evaluation and Results (2014) *Quality Assurance Framework for Implementing Capacity Building Programs*

DAC/UNEG *Peer Review of the Evaluation Function of UNICEF* and UNICEF Evaluation Management Response (2017)

EPRS | European Parliamentary Research Service (2017) *Evaluation in the European Commission – Rolling Checklist and State of Play*

EvalPartners, IOCE, UNICEF (2013) *Evaluation and Civil Society – Stakeholders’ perspective on National Evaluation Capacity Development*, published by UNICEF in collaboration as well with CLEAR, IEG-World Bank, OECD, UNEG, UN-Women.

FAO (2019) *Programme Evaluation Report 2019*

FAO (2017) *Guidelines for the assessment of gender mainstreaming*

FAO (2015) *OED Evaluation Manual*

FAO (2015) *OED Guidance - Responsibilities and procedures for management responses and follow-up reports on evaluations*

FAO (2011) *Quality Assurance Framework for evaluation in FAO*

FAO (2011) *Tools for Quality Assurance of OED evaluation terms of reference and reports*

FAO (2010) *Charter for the FAO Office of Evaluation*

GEF (2019) *The GEF Evaluation Policy*

GEF (2018) *Sixth Comprehensive Evaluation of the GEF*

GEF (2011) *Terminal Evaluation Review Criteria and Indicators*

GEF (2010) *GEF Monitoring and Evaluation Policy*

GEF (2009) *The ROTI Handbook - Towards Enhancing the Impacts of Environmental Projects*

Group URD (2018) *Quality and Accountability COMPASS*

ICRC (2017) *Handbook on Data Protection in Humanitarian Action*

IADB (2019) *18 Annual Report of the Office of Evaluation and Oversight of the Inter-American Development Bank*

IADB (2019) *Evaluation Policy Framework*

IADB (2019) *Review of Knowledge Generation and Dissemination in the Inter-American Development Bank*

ILO-EVAL (2017) *Independent evaluation of the ILO's Evaluation Function* by ImpactReady and Lattanzio

ILO-EVAL (2018) *ILO Evaluation Quality: Challenges and Potential Strategies Reflections Beyond the Quality Appraisal* i-eval Think Piece No 13 by Salvador Bustamante, Estela López & Inmaculada Román Edited by C. Russon

ILO -EVAL (2017) *Evaluation quality assessment methodology in the UN system and changes to the ILO's quality appraisal methodology* i-eval Think Piece No 12 by Iker Llabres

ILO (2017) *ILO Evaluation Policy*. GB.331/PFA/8

ILO -EVAL (2016) *Quality Assessments of ILO Project Evaluations: What are the next steps to better evaluations?* i-eval Think Piece No 10 by Brad R. Watts

ILO - Watts, B., Coryn, C., Feikowsky, E., Wilson, L. and Mateu, P. (2015) *Independent Quality Appraisal of ILO Project Evaluations*. The Evaluation Center, Western Michigan University

ILO_EVAL (2014) *Leveraging appraisal findings to improve evaluation quality*, I-eval Think Piece No 4 by Kelly Robertson and Daniela Schroeter

ILO_EVAL (2014) *External Quality Appraisal: Implications for evaluation quality and utilization*. i-eval THINK Piece No. 8 by Judith Friedman and Naomi Blight

ILO-EVAL (2014) *Checklist 6: Rating the Quality of Evaluation Reports*. i-eval Resource Kit

ILO (2011) *Results-based strategies 2011–15: Evaluation strategy – Strengthening the use of evaluations*. GB.310/PFA/4/1(Rev.)

ILO (2005) *Evaluation. A new policy and strategic framework for evaluation at the ILO.*

GB.294/PFA/8/4

ISO – International Organization for Standardization (2015) 5th Edition – ISO 9000-2015 *Quality Management Systems – Requirements*

ISO (2015) *Quality Management Principles*

JIU Prom-Jackson, S. and Bartsiotas, G. (2014) *Analysis of the Evaluation Function in the United Nations System*. JIU. Geneva

JIU Prom-Jackson, S. and Bartsiotas G. (2013) *Review of the Management of Implementing Partners in the United Nations System Organisation*

Mathiason John R. (2019) *Assessing the Quality of International Evaluations*, CIPA, Cornell University

MOPAN (2017) *2015-16 Assessments United Nations Children’s Fund (UNICEF) Institutional Assessment Report*

Norwegian Refugee Council NRC (2016) *Monitoring & Evaluation: Minimum Standards and Guidelines*

OECD DAC (2010) *Guidelines and Reference Series – Quality Standards for Development Evaluation*

OECD DAC Network on Development Evaluation (2013) *Review of UNICEF’s Development Effectiveness: Final Report 2009 – 2011*

Oxfam (2019) *Evaluation Guidelines*

Oxfam (2017) *Gender Leadership in Humanitarian Action – Training Manual*

Ribi Joan (2017) *Evaluation of European Commission Policy and Programmes. Overarching Principles and Quality Assurance*, SEVAL

Surbhi S. (2019) *Difference Between Quality Assurance and Quality Control*, Key Differences

The Netherlands (2009) *Evaluation policy and guidelines for evaluations*

UNAIDS (2019) *Evaluation Policy*

UNDP (2019) *The revised UNDP evaluation policy*

UNDP (2019) *UNDP Evaluation Guidelines*

UNDP (2019) *Evaluation Quality Assessment*

UNDP (2016) *The UNDP Evaluation Policy*

UNDP (2014) *Independent Evaluation Office response to the review of the UNDP evaluation policy*

UNDP (2014) *Guidance for Conducting Mid-Term Reviews of UNDP-Supported, GEF-Financed Projects*

UNDP (2012) *Guidance for Conducting Terminal Evaluations of UNDP-Supported, GEF-Financed Projects*

UNDP (2011) *A companion guide to the handbook on planning monitoring and evaluating for development results for programme units and evaluators*

UNDP (2009) *Handbook on Planning, Monitoring and Evaluating for Development Results*

UN Economic and Social Council (2019) - *Annual report for 2018 on the evaluation function in UNICEF*

UN Economic and Social Council (2018) - *Annual report for 2017 on the evaluation function in UNICEF*

UN Economic and Social Council (2017) - *Annual report for 2016 on the evaluation function in UNICEF*

UN Economic and Social Council (2016) - *Annual report for 2015 on the evaluation function in UNICEF*

UNEG (2018) *Improved Quality of Evaluation Recommendations Checklist*

UNEG (2016) *Norms and Standards for Evaluation*

UNEG (2015) *Integrating Human Rights and Gender Equality in Evaluations*

UNEG (2015) *Strengthening Decentralized evaluation – Concept Note*

UNEG (2010) *UNEG Quality Checklist for Evaluation Reports*

UNEP (2018) *Quality Assurance of Evaluation Office Reports*

UNEP (2017) *Quality Assessment of the Evaluation Report*

UNEP (2016) *Evaluation Policy*

UN ESCAP (2007) *Monitoring and Evaluation Framework*

UNESCO (2016) *Synthetic Review of Evaluations in the UNESCO System*. Evaluation Office. IOS/EVS/PI/148 REV.2.

UNESCO (2015) *Evaluation Strategy 2016 – 2018*

UNESCO (2014) *Evaluation Policy 2014-2021*. 196 EX/24.INF.

UNESCO (2008) *Guidelines for Inception Reports*

UNFPA (2019) *Evaluation Policy*

UNFPA (2019) *Evaluation Handbook -How to Design and Conduct a Country Programme Evaluation at UNFPA*

UNFPA (2018) *UNFPA Independent External Strategic Review*

UNFPA (2016) *Evaluation Quality Assessment Grid Template*

UNFPA (2016) *Explanatory notes for UNFPA Evaluation Quality Assessment Grid*

UN General Assembly (2016) – *Resolution 71/243. Quadrennial comprehensive policy review of operational activities for development of the United Nations system*

UNHCR (2016) *UNHCR Policy on Evaluation*

UN-Habitat (2013) *Evaluation Policy*

UN-Habitat (2011) *Reporting Quality Criteria*

UN-Habitat (2011) *Format for Evaluation Reports*

UNICEF (2019) *External Review of UNICEF's Global Evaluation Reports Oversight System (GEROS)* Associates for International Management Services (AIMS), UNICEF Evaluation Office, New York

UNICEF (2019) *GEROS Meta-Analysis 2016-2018: An Independent Review of UNICEF Evaluation Report Quality and Selected Trends from 2016-2018*

UNICEF (2019) *UNICEF Guidance on Gender Integration in Evaluation*

UNICEF (2019) *Terms of Reference for External Peer Reviewer* prepared by Tina Tordjman-Nebe

UNICEF (2019) *Guidance Note for Conducting Evaluability Assessment in UNICEF*

UNICEF (2018) *Strategic Plan 2018-2021*

UNICEF (2018) *Revised Evaluation Policy of UNICEF*

UNICEF (2018) *Corporate Implementation Framework for the UNICEF Evaluation Policy*

UNICEF (2018) *UNICEF Procedure on the Implementation of the 2018 UNICEF Evaluation Policy*

UNICEF (2018) *Plan for Global Evaluation 2018-2021*

UNICEF (2018) *Evaluation Management Response – Guidance for UNICEF Staff* by ImpactReady

UNICEF (2018) Evaluation Office – *Progress Update on EISI*

UNICEF (2018) *Influential Evaluations: A Selection of UNICEF Evaluations That Led to Learning and Change* by Samantha Cocco-Klein

UNICEF (2017) *Adapted UNEG Quality Checklist for Evaluation Terms of Reference*

UNICEF (2017) *Adapted UNEG Evaluation Reports Standards*

UNICEF (2017) *Global Evaluation Report Oversight System UNICEF Staff Handbook*

UNICEF (2016) *External Assessment of UNICEF's Global Evaluation Reports Oversight System (GEROS)* by Associates for International Management Services (AIMS)

UNICEF (2016) *UNICEF GEROS Meta-Analysis 2015, An independent review of UNICEF evaluation report quality and trends, 2009-2015* by ImpactReady

UNICEF (2015) *GEROS Global Meta-Evaluation 2014*

Other Reports concerning GEROS

GEROS- Evaluation Reports for UNICEF (2018)

GEROS Evaluation Quality Assurance Tool template (EQA) (2013 - 2018)

GEROS – Assessments/ Evaluation Reports EAPR (2018)

GEROS – Assessments/ Evaluation Reports ECAR/CEE/CIS (2018)

GEROS – Assessments/Evaluation Reports ESAR (2018)

GEROS – Assessments/Evaluation Reports HQ (2018)

GEROS – Assessments/Evaluation Reports LACRO (2018)

GEROS – Assessments/Evaluation Reports MENA (2018)

GEROS – Assessments/Evaluation Reports ROSA (2018)

GEROS – Assessments/Evaluation Reports WCARO (2018)

GEROS – Final EQA for UNICEF (2018)

GEROS – Evaluation Report for Azerbaijan (2017)

GEROS – UNICEF Handbook Summary (2017)

GEROS – Final list of Evaluation Reports (2017)

UNICEF (2015) *UNICEF Procedure for Ethical Standards in Research, Evaluation, Data Collection and Analysis* – Documents number CF/PD/DRP/2015-001

UNICEF (2015) *UNICEF Procedure for Quality Assurance in Research* by the Division of Data, Research and Policy – Document number CF/PD/DRP/2015-002

UNICEF (2014) Evaluation Office - *A Guide on how to use the New Global Evaluation Database and Global Evaluation Management Response Tracking System*

UNICEF (2014) *Standard Operating Procedures for timely preparation and implementation of management responses to evaluations commissioned by the Evaluation Office*

UNICEF (2013) *Global Evaluation Reports Oversight System (GEROS)*. Evaluation Office

UNIDO (2018) *Evaluation Policy*

UNIDO (2015) Director General’s Bulletin: *Evaluation Policy*. UNIDO/DGB/(M).98/Rev.1.

UNODC (2019) *Independent Quality Assessment of UNODC Evaluation Reports 2018*

UNODC (2017) *Independent Quality Assessment of UNODC Evaluation Reports 2016*

UNODC (2017) *Evaluation Handbook* Guidance for designing, conducting and using independent evaluations at UNODC

UNODC (2015) *Evaluation Policy*

UNODC (2010) *Inception Report for Final or Mid Term Evaluation*

UN-OIOS (2018) United Nations Secretariat-*Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation*

UN-OIOS/IED (2014) *Inspection and Evaluation Manual*

UNSDG (2019) *United Nations Sustainable Development Cooperation Framework (UNSDCF) 6th Draft*

UN Women (2018) *Quality assurance tool for training for gender equality*

UN Women (2017) *Global Evaluation Report Assessment and Analysis System. Meta Evaluation Report 2016*

UN Women (2015) *How to Manage Gender-Responsive Evaluation – Evaluation Handbook*

UN Women (2015) *Global Evaluation Report Assessment and Analysis System (GERAAS)*

UN Women (2015) *Meta-Evaluation Report – GERAAS 2014 by Impact Ready*

UN Women (2012) *UN Women Evaluation Policy*

USAID, UNICEF (2017) *Interagency Gender-Based Violence Case Management Guidelines*

USAID (2017) *Strengthening Evidence-Based Development – Five Years of Better Evaluation Practices at USAD 2011-2016*

USAID (2016) *Evaluation Utilization at USAID*

USAID (2013) *Meta-Evaluation of Quality and Coverage of USAID Evaluation 2009-2012*

WFP (2018) *Evaluation Top 10 Lessons for Quality Policy in WFP*

WFP (2018) *CEQAS Central Evaluation Quality Assurance System Guidance Materials*

WFP (2018) *Annual Evaluation Report for 2017*

WFP (2017) *Annual Evaluation Report for 2016*

WFP (2016) *Annual Evaluation Report for 2015*

WFP (2017) *Monitoring and Evaluation Guidelines*

WFP (2017) *Guidance Materials for Country Portfolio Evaluations*

WFP (2015) *Evaluation Policy (2016–2021) WFP/EB.2/2015/4-A/Rev.1.*

WFP (2014) *Guidance Materials for Policy Evaluations*

WFP (2014) *Guidance Materials for Strategic Evaluations*

WFP (2013) *Evaluation Quality Assurance System (EQAS)*

WHO (2018) *Evaluation Policy*

WHO (2013) *Evaluation Practice Handbook*

WIPO (2016) *Evaluation Policy*

WIPO (2016) *Evaluation Manual*

World Bank (2019) *World Bank Group Evaluation Principles*

World Bank (2018) *Five ways to think about quality in evaluation* by Jos Vaessen

World Bank (2015) *Managing Evaluations: A How-To Guide for Managers and Commissioners of Evaluations*

World Bank (2012) *Designing a results framework for achieving results: a how-to guide*