

Guidance on evaluation quality assurance and assessment

Towards excellence in evaluation



2024



Table of Contents

Foreword.....	3
1. Introduction.....	1
2. Principles to guide quality evaluations at UNFPA.....	2
Independence.....	2
Impartiality.....	2
Credibility.....	3
Utility.....	3
Other cross-cutting considerations.....	3
3. Translating the principles: Quality assurance and assessment (EQAA) in practice.....	4
Evaluation quality assurance.....	4
Evaluation quality assessment (EQA).....	5
Revisions to the EQA.....	5
The EQA process.....	6
4. The EQA grid.....	7
5. The EQA rating scale.....	8
Overall rating.....	8
Indicator ratings.....	9
UN-SWAP indicator ratings.....	9

Copyright © UNFPA 2024, all rights reserved.

This is a publication by the Independent Evaluation Office of UNFPA.

 unfpa.org/evaluation  evaluation.office@unfpa.org  [@unfpa_eval](https://twitter.com/unfpa_eval)  [@UNFPA_EvaluationOffice](https://www.youtube.com/@UNFPA_EvaluationOffice)

Foreword

The Independent Evaluation Office is committed to producing high-quality, contextualized, and relevant evaluations that inform better programming and policy decisions.

With the approval of the 2024 Evaluation Policy by the Executive Board, ensuring high-quality and credible evaluations has become not only a policy imperative, but also a fundamental prerequisite to strengthen accountability, evidence-based decision-making and learning. The revised Evaluation Quality Assurance and Assessment (EQAA) system introduces a more rigorous assessment framework, consistent with the 2024 Evaluation Policy, which calls for a higher standard and threshold for quality of evaluation. In addition, it responds to the 2023 independent peer review of the evaluation function and enhances its alignment with EQAA systems of sister United Nations agencies.

This guidance note aims to increase the knowledge and understanding of evaluation staff, evaluation managers, and external evaluation consultants on quality assurance and assessment at UNFPA. A key feature of this guidance note is an updated Evaluation Quality Assessment grid that will serve to independently assess all types of evaluations and raise the bar for the quality of evaluations at UNFPA.

I am confident that the revised system will serve as a valuable tool in ensuring that all UNFPA evaluations are of the highest quality possible, to improve decision-making to achieve results for women and girls around the world.

Marco Segone

Director, Independent Evaluation Office
UNFPA

1. Introduction

The evaluation function at UNFPA serves to strengthen the accountability, evidence-based decision-making and learning by providing high quality evidence. In order to do this and ensure that evaluations are credible, high quality and useful, the Independent Evaluation Office (IEO) put in place an evaluation quality assurance and assessment (EQAA) system.

The EQAA system has two mechanisms to address the quality of evaluation at UNFPA – quality assurance and quality assessment. **Quality assurance** takes place from the onset of the evaluation and continues through each phase of the evaluation. It is conducted at various levels, including by the evaluation team leader, the evaluation reference group, and the evaluation manager. **Quality assessment**, which serves to complement the quality assurance process, takes place after an evaluation is completed, whereby the evaluation report is systematically assessed against established standards and criteria by an external independent reviewer.

The EQAA system is essential to ensure that all evaluations at UNFPA are conducted in an impartial and credible manner, demonstrate a high methodological rigor, produce accurate and reliable evidence, and provide relevant and useful information for decision-making.

The purpose of this guidance is to provide clarity on the principles underpinning the quality of evaluations undertaken at UNFPA and how these principles are applied and operationalized through various quality assurance and assessment tools. The guidance also outlines the revisions that have been introduced to enhance the [Evaluation Quality Assessment grid](#), which is used to assess all types of evaluations in UNFPA.

The revisions introduced in this guidance recognise the need for a more comprehensive system that can be systematically applied to both centralized and decentralized evaluations. The guidance note is intended to increase the knowledge and understanding of UNFPA evaluation staff, evaluation managers, and external evaluation consultants on quality assurance and assessment at UNFPA.

2. Principles to guide quality evaluations at UNFPA

The [UNFPA Evaluation Policy 2024](#) underscores the importance of high-quality evaluations to improve decision-making, increase accountability, and more importantly, facilitate use. Accordingly, the IEO together with regional offices has made efforts in recent years to strengthen evaluation quality, including by developing additional tools and templates such as the Country Programme Evaluation (CPE) Management Kit, revamping the methodological framework for country programme evaluations covering each phase of an evaluation and provision of various evaluation capacity development initiatives. In addition, the IEO continuously strives to ensure the utility, relevance and effectiveness of its quality assessment grid through internal reviews and revisions of the tool. Taken together, such efforts contribute towards strengthening the evaluation culture and capacities at UNFPA while also fulfilling requirements of the Evaluation Policy.

The guiding principles of evaluation at UNFPA emanate from the United Nations Evaluation Group (UNEG) norms and standards and code of conduct for evaluation. These principles, outlined in the UNFPA Evaluation Policy 2024, form the foundation for its evaluation quality assurance and assessment system. In particular, the EQAA system focuses on core principles of independence, impartiality, credibility, and utility.¹

Independence

Independence in evaluation entails the ability of evaluators to work without the undue influence of any party. It includes, for instance, the freedom for evaluators to choose the design of the evaluation and to select suitable methods and tools for data collection and analysis. Ensuring the independence of an evaluation bolsters the credibility of the exercise as well as promotes the reliability of the evaluation report.

Impartiality

Impartiality requires that evaluators to not have been directly responsible for the policy setting, design or management of the subject being evaluated. Impartiality helps to ensure UNFPA evaluations maintain a level of objectivity while safeguarding the integrity of the evaluation process by preventing conflicts of interest.

¹ [UNEG Norms and Standards for Evaluation \(2016\)](#).

Credibility

Credibility requires the evaluation be conducted in a transparent and unbiased manner, employing inclusive approaches with a wide range of relevant stakeholders and a rigorous methodology that includes robust tools to collect and analyze data in a structured, ethical, and comprehensive manner. The quality of an evaluation is very much dependent on such elements and is further linked to the principles of independence and impartiality.

Utility

Utility requires that evaluations be understood as a tool to catalyze change, stimulate learning, and promote accountability. The utility of an evaluation is highly dependent on its quality, as well as relevant and timely contributions to organizational learning and decision-making processes.

Other cross-cutting considerations

In addition to these principles, the EQAA has evolved to include other cross-cutting issues and considerations such as human rights and gender equality, disability inclusion, leaving no one behind, social and environmental standards, and ethical dimensions. Incorporating cross-cutting issues into evaluation quality assessment is essential to ensuring that the evaluation process promotes a more inclusive, holistic human rights-based approach and that the evaluation report adequately addresses such issues that are directly linked to more inclusive and sustainable development.

- **Human rights and gender equality:** A high-quality evaluation should integrate human rights and gender equality considerations throughout the evaluation process and product. In this effort, UNFPA has adhered to UNEG guidance on integrating gender and human rights in evaluation², which are grounded in the principles of inclusion, participation and fair power relations, and has directly integrated the United Nations System-Wide Action Plan on Gender Equality and Women's Empowerment (UN-SWAP) on Gender Equality and the Empowerment of Women Evaluation Performance Indicator³ among its criteria to assess the quality of evaluation reports.
- **Disability inclusion:** Similarly, a high-quality evaluation should also mainstream disability inclusion into its evaluation processes and products⁴. For this purpose, this guidance aligns with the United Nations Disability Inclusion Strategy⁵ and the related technical notes to operationalize the strategy.
- **Leave no one behind:** Evaluation at UNFPA seeks to integrate the principles of leave no one behind/reaching the furthest behind into evaluation processes, approaches and results⁶ to

² [UNEG Guidance: Integrating Human Rights and Gender Equality in Evaluations.](#)

³ [UN-SWAP on Gender Equality and the Empowerment of Women Evaluation Performance Indicator.](#)

⁴ [Guidance on disability inclusion in UNFPA evaluations.](#)

⁵ [United Nations Disability Inclusion Strategy.](#)

⁶ [Guidance on integrating the principles of leaving no one behind and reaching the furthest behind in UNFPA evaluations.](#)

improve the overall quality through ensuring that evaluations use an intersectional lens to address the root causes of discrimination, social norms, structural barriers, and gender inequality.

- **Social and environmental standards:** UNFPA evaluations seek to integrate social and environmental standards⁷ in a systematic fashion to improve the quality of its evaluation reports through strengthening the evidence base on how UNFPA programmes and operations have minimized and potentially mitigated the adverse effect of social and environmental effects beyond the “do no harm” principle.
- **Ethical dimensions:** UNFPA also places great importance on ethical principles to be upheld when conducting evaluations, in line with UNEG Ethical Guidelines for Evaluation⁸ and the UNEG Code of Conduct for Evaluation⁹, including ethical dimensions in humanitarian contexts¹⁰.

3. Translating the principles: Quality assurance and assessment (EQAA) in practice

The evaluation principles elaborated in the previous section provide the foundation for IEO’s EQAA system. As such, the evaluation principles are translated directly into the tools and templates, as well as technical and capacity development support provided by IEO and regional offices to help facilitate and ensure high-quality evaluation reports.

Evaluation quality assurance

Evaluation quality assurance is a systematic process for ensuring that evaluations are conducted in a rigorous, impartial, and transparent manner, producing credible and actionable evidence to inform decision-making. Evaluation quality assurance occurs throughout the evaluation process, from planning and design to implementation, reporting, and dissemination.

The UNFPA Evaluation Handbook 2024, the CPE Management Kit, the guidance on humanitarian evaluation and a series of theme-based guidance documents issued by the IEO form the cornerstone of a comprehensive evaluation quality assurance (EQA) framework that ensures the consistent and rigorous application of evaluation standards at each phase of the evaluation, from developing terms of reference to selecting evaluation teams, conducting the design phase, and drafting and finalizing reports. This framework is further complemented by an e-learning course on evaluation launched by the IEO to deepen the knowledge and skills of M&E staff organizing, , managing, and utilizing high-quality evaluations. Together, these existing sets of guidance together with quality assurance and

⁷ [Guidance on integrating social and environmental standards into evaluations.](#)

⁸ [UNEG Ethical Guidelines for Evaluation.](#)

⁹ [UNEG Code of Conduct for Evaluation.](#)

¹⁰ [ALNAP Evaluation of Humanitarian Action Guide \(2016\).](#)

advisory support by regional M&E advisors and IEO fosters a culture of continuous improvement in evaluation quality, ensuring that UNFPA evaluations consistently meet the highest standards of rigour and relevance.

To avoid duplication of these already comprehensive resources, this guidance note will specifically focus on ex-post evaluation quality assessment.

Evaluation quality assessment (EQA)

Evaluation quality assessment is conducted after the evaluation takes place by independent assessors. It serves as an essential tool to ensure UNFPA evaluations are conducted rigorously, impartially, and transparently and produce credible, reliable and useful evaluation results.

The EQA provides an external assessment of the strengths and weaknesses of the final evaluation report to offer an indication of the relative reliability of the results and determine the extent to which the report can be used with confidence to feed into future programming and to serve other purposes.

Evaluation quality assessment covers UNFPA's two main types of evaluations, 1) centralized evaluations commissioned and conducted by the IEO and 2) decentralized evaluations¹¹ commissioned by country offices and regional offices, as well as business units in the Headquarters.

Revisions to the EQA

In response to UNFPA's commitment to advancing good evaluation practice, periodic adjustments have been made to the evaluation quality assessment portion of the EQAA since its introduction in 2013. Specifically, a review of the UNFPA EQAA was performed in 2015-2016 with the objective of proposing options for expanding and further strengthening the quality assurance and assessment systems. In late 2020, the EQA grid and guidance note was updated with a view towards refining the grid and adding disability inclusion and leave no one principles as an important element to be assessed.

In 2023, an independent peer review of UNFPA evaluation function¹² was conducted and included a recommendation on strengthening the EQAA system, including a review of the evaluation quality assessment criteria. Moreover, the recommendation suggested that, moving forward, all evaluations, including project, programme, and thematic evaluations that had previously been excluded, should be subject to evaluation quality assessment. The IEO accepted this recommendation and undertook a thorough review and revision of the EQA system to meet the needs of the organization for high-quality evaluations to support its programmatic, operations and policy decisions. UNFPA is also aligned to the evaluation quality assessment system of sister agencies to enhance and promote more coordinated,

¹¹ Decentralized evaluations include, in addition to Country Programme Evaluations (CPE) and Regional Programme Evaluations, programme and project evaluations (including joint evaluations) managed by the business units responsible for the programme/project being assessed. In humanitarian situations, decentralized evaluations also include evaluations of emergency responses where the scale, magnitude and level of complexity of the emergency is such that the country office can manage with existing resources while requiring limited or additional prioritized support from the regional office and headquarters.

¹² [Independent peer review of UNFPA evaluation function 2023.](#)

coherent efforts across the UN system in the context of UN reform. Lastly, the revised EQA will include all evaluations at the decentralized level, including country-level project and programme evaluations.

To this end, the revised EQA system is closely aligned to the overarching requirements set forth in the revised evaluation policy and drew important elements from the UNICEF Global Evaluation Reports Oversight System (GEROS)¹³ while retaining the particularities and unique requirements of the UNFPA evaluation function.

Toward this end, the revised EQA includes additional guiding questions and indicators and a new rating schema. A higher threshold for the top rating category will serve to raise expectations for a more robust, high quality evaluation report that can be used with increased confidence for accountability, learning and decision-making purposes. The overall rating for the report will now make use of a five-point rating system (excellent, highly satisfactory, satisfactory, fair and unsatisfactory) instead of the previous four-scale system (very good, good, fair and unsatisfactory).

Further adjustments have also been made on the grid by revising the weighting given to each section with increased importance given to: a solid assessment of the theory of change and the underlying assumptions to assess contribution to outcome level changes; the identification of appropriate evaluation questions to meet the evaluation objectives and purpose; emphasis on meaningful engagement of key stakeholders in the validation of findings; and, co-development of recommendations.

This guidance and related EQA tools will be periodically revisited by the IEO to ensure continued usefulness in line with the evolving needs of the organization as well as internationally recognized good practices in evaluation.

The EQA process

Given that the IEO is the custodian of the EQAA system, the evaluation quality assessment process will be coordinated by designated staff of the UNFPA IEO.

To ensure the independence of the process, external reviewers will conduct the evaluation quality assessment for each evaluation report. External reviewers should not have had any prior involvement with the intervention evaluated, nor with the team who conducted the evaluation being reviewed. Moreover, evaluation reports should generally be subject to assessment by two external reviewers to ensure consistency and internal validity.

¹³ Global Evaluation Reports Oversight System (GEROS): The UNICEF evaluation quality-assessment system. In the spirit of UN reform, UNFPA decided to leverage the UNICEF LTA with an external firm contracted to conduct its evaluation quality assessments as well as align its evaluation quality assessment tools to UNICEF GEROS system to enhance the coherence and collaboration between the two agencies. Notably, UNICEF GEROS system underwent a thorough revision in recent years and introduced a more rigorous assessment framework, which is consistent with UNFPA Evaluation Policy 2024 that calls for a higher rigor and threshold for ex-post quality assessments.

To conduct an evaluation quality assessment, the external reviewers will be expected to:

- Thoroughly review the evaluation report and all its annexes in the framework of the EQA grid
- Review all additional documents that may be relevant to the evaluation report, in the framework of the EQA grid, e.g. the terms of reference of the evaluation, the inception/design report, country case study notes or reports
- Assign a rating for each indicator and include a detailed justification for the rating which includes specific examples from the report. The justification should go beyond the indicators (i.e. does not repeat the information from the indicators in the grid) and add value to understand the reviewer's assessment
- Ensure each section score is coherent and consistent with the questions and respective indicator ratings. Re-examine and adjust as necessary. Do the same for the overall report rating
- Write a top-line summary of the assessment that highlights the strengths and weaknesses of each section, constructive feedback for the sections that need improvement, and any good practices or innovations
- Elaborate and explain their assessments to the IEO, and if deemed necessary, conduct another review of specific sections if the assigned ratings and written justifications are inconsistent, inaccurate or incomplete
- Complete and finalize the evaluation quality assessment grid.

Once the evaluation quality assessment is completed, the final assessment will be provided to the evaluation manager and relevant business units involved.

4. The EQA grid

The main tool used to assess the quality of evaluation reports is the EQA grid, available [here](#). The EQA grid uses a rating system and applies specific weighting to nine sections representing different components of the evaluation report. While all components of an evaluation report are important, the EQA grid applies more weight to the methodology, findings, conclusions, and recommendation sections. Taken together, these four sections constitute nearly 70 per cent of the total rating score.

Each section contains questions that elucidate specific aspects that need to be addressed. The questions have indicators to help determine the extent to which the aspects adequately are reflected in the evaluation report.

Although the EQA is devised to accommodate all types of evaluations in UNFPA, some sub-criteria might not be fully applicable to certain types of evaluations. For instance, developmental evaluation is distinct from more traditional formative and summative evaluations in terms of evaluation practices, standards, and approach to measurement and methods. For this reason, a distinct evaluation quality

assessment grid for developmental evaluations at UNFPA was created and will be applied to these types of evaluations¹⁴.

5. The EQA rating scale

Overall rating

The overall EQA rating uses the following five-scale rating system¹⁵.

Overall Rating		Implication
*****	Excellent	The report is an exemplary level of quality ¹⁶ and exceeds all UNFPA/UNEG standards for evaluation reports. Decision makers can use the evaluation with a high degree of confidence.
****_	Highly Satisfactory	The report fully meets all UNFPA/UNEG standards for evaluation reports, with minor shortcomings in certain indicators. Decision makers may use the evaluation with a high degree of confidence.
***_ _	Satisfactory	The report meets UNFPA/UNEG standards for evaluation reports, but some indicators are inadequately addressed or missing. Decision makers may use the evaluation with some confidence.
**_ _ _	Fair	The report meets UNFPA/UNEG standards for evaluation reports in some regards, but not all. Substantive improvements in some areas are needed. Decision makers may use the evaluation with caution.
*_ _ _ _	Unsatisfactory	The report does not sufficiently meet the UNFPA/UNEG standards for evaluation reports, and there are significant gaps and shortcomings. The report may offer some learning, but decision makers should use it with high caution.

The overall rating is based on an aggregation of the weighted values and assigned ratings for the indicators under the questions for each section of the evaluation report.

¹⁴ [Assessing the quality of developmental evaluations at UNFPA.](#)

¹⁵ The overall aim of the overall assessment is not to grade or compare evaluation reports rather evaluations are generally quality assessed against specific criteria to determine the overall quality of the evaluation report to build confidence in the findings and understanding of their limitations in decision-making.

¹⁶ The top-rated evaluations will be promoted by the IEO organization-wide to push the bar for high quality evaluations in UNFPA.

The weighting of each section is based on its relative importance to the overall quality of the evaluation report.

The weighting for each section is as follows:

Section	Weighting (%)
Executive Summary	5
Background	5
Purpose, Objectives and Scope	5
Design and Methodology	20
Findings	25
Conclusions	10
Recommendations	15
Structure and Presentation	5
Evaluation Principles/cross-cutting issues	10
TOTAL	100

Indicator ratings

The EQA grid includes 20 questions with multiple indicators for each question. Each indicator is assigned a rating and an explanatory comment to justify the rating.

Indicators are assigned ratings as follows:

Yes	The indicator is fully met
Partially	Some aspects of the indicator have been met, but not all
No	There is no or only minimal evidence of the indicator being met
Not applicable	The indicator is not applicable to the evaluation report ¹⁷

In the grid, each indicator rating corresponds to a numerical value: Yes (1), Partially (0.5), No (0) or Not applicable (-). The indicators are equally weighted and are auto calculated to provide the section score.

UN-SWAP indicator ratings

All UN entities are required to report on their evaluation performance to meet the requirements of the UN-SWAP. In this effort, the UNFPA EQA grid includes the three required evaluation performance indicators (EPI) in Question 20.

¹⁷ Questions may be left unrated by reviewers where they find them not applicable due to the direction of the ToR or the context of the intervention under evaluation, and the justification should be explained in the comments portion of the EQA grid.

The indicators relating to the UN-SWAP (Question 20) reflect a different rating scale and numerical value which is as follows:

0 = Not at all integrated. Applies when none of the elements under an indicator are met.

1 = Partially integrated. Applies when some minimal elements are met, but further progress is needed and remedial action to meet the indicator is required.

2 = Satisfactorily integrated. Applies when a satisfactory level has been reached and many of the elements are met, but still improvement could be done.

3 = Fully integrated. Applies when all of the elements under an indicator are met, used and fully integrated in the evaluation and no remedial action is required.



Driving evidence-based actions
Delivering rights and choices for all

**United Nations Population Fund
Independent Evaluation Office**

605 Third Avenue
New York, NY 10158 USA

 unfpa.org/evaluation

 evaluation.office@unfpa.org

 [@unfpa_eval](https://twitter.com/unfpa_eval)

 [@UNFPA_EvaluationOffice](https://www.youtube.com/@UNFPA_EvaluationOffice)