

2020 Edition

UNFPA Evaluation Office

Evaluation Quality at UNFPA

Principles and their application



December 2020

Copyright © UNFPA 2020, all rights reserved.

This is an independent document by the UNFPA Evaluation Office.

Any enquiries about this document should be addressed to Mr Messay Tassew, tassew@unfpa.org

Table of Contents

- 1. Introduction 3
- 2. Background: Evaluation at UNFPA..... 3
- 3. The concept of evaluation quality and UNFPA quality assurance and assessment system..... 4
- 4. Translating the principles: Quality assurance and assessment in practice 6

1. Introduction

The UNFPA Evaluation Quality Assurance and Assessment (EQAA) system aims to facilitate the quality of evaluations at UNFPA through two processes: quality assurance and quality assessment. Quality assurance occurs *throughout* an evaluation, beginning with the evaluation's terms of reference and ending with the draft final evaluation report. Quality assessment, on the other hand, takes place after an evaluation is completed (ex post), with the final evaluation report quality assessed by an external independent assessor.

The EQAA system is aligned with the Evaluation Policy (2019), the updated UN Norms and Standards for Evaluation (2016), Integrating Human Rights and Gender Equality in Evaluations (2014), and the United Nations Disability Inclusion Strategy (2019). The system covers two types of evaluations: centralized evaluations conducted by the independent Evaluation Office and decentralized programme level evaluations (namely country and regional programme evaluations) managed by decentralized business units. Evaluation quality assessments (EQAs) are posted publicly alongside the corresponding evaluation report (and management response) on the Evaluation Database.

This revised version includes disability inclusion dimensions in specific evaluation quality assessment criteria and sub-criteria, where appropriate, as well as other enhancements to the quality assessment grid based on lessons learned from the implementation of the system over the last three years.

The purpose of this guidance is to further flesh out the concept of (and principles underpinning) evaluation quality at UNFPA, and detail how this is applied/operationalized through assurance and assessment tools. The 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 and, by so doing, contribute to improved evaluation quality and use for programming and policy decisions.

2. Background: Evaluation at UNFPA

UNFPA adheres to the United Nations Evaluation Group (UNEG) norms and standards for evaluation¹ and its definition of evaluation: “An evaluation is an assessment, conducted as systematically and impartially as possible, of an activity, project, programme strategy, policy, topic, theme, sector, operational area or institutional performance. It analyses the level of achievement of both expected and unexpected results by examining the results chain, processes, contextual factors and causality using appropriate criteria such as relevance, effectiveness, efficiency impact and sustainability. An evaluation should provide credible, useful evidence-based information that enables the timely incorporation of its findings, recommendations and lessons into the decision- making process of organisations and stakeholders.”

¹ UNEG, Norms and Standards for Evaluation, June 2016

Evaluation at UNFPA serves three main purposes:²

- Evaluation as a means to demonstrate accountability to stakeholders on performance in achieving development results, and on invested resources.
- Evaluation as supporting evidence-based decision making in order to achieve sustainable development results.
- Evaluation as contributing lessons learned to the existing knowledge base on how to accelerate implementation of the Programme of Action of the International Conference on Population and Development and the achievement of the Sustainable Development Goals.

By supporting learning and evidence-based decision-making, and demonstrating institutional accountability, evaluation contributes to UNFPA's drive for results. The production of **high-quality evaluation reports** is central toward this end; as such the UNFPA evaluation policy and evaluation function place a strong emphasis on **ensuring the quality** of the evaluation exercise and resulting reports, as well as related processes and products.

3. The concept of evaluation quality and UNFPA quality assurance and assessment system

How, then, to understand the concept of high-quality evaluations? What are the framing principles and underlying dimensions of evaluation quality?

Overall, UNFPA Evaluation Policy adheres to the foundation documents of UNEG and thereby to the norms of utility, credibility, and independence, as well as impartiality, ethics, transparency, human rights and gender equality.³ In 2016, the UNEG Norms and Standards were updated and the approach upgraded to give greater prominence to ethics, transparency and human rights and gender equality.

The UNFPA evaluation policy underscores the importance of quality evaluations to, inter alia, facilitate use.⁴ Accordingly, the Evaluation Office has made efforts in recent years to strengthen evaluation quality, including by developing additional tools and templates to support assurance efforts, and expanding the EQAA system to cover both programme-level evaluations as well evaluations conducted by the Evaluation Office (i.e. centralized evaluations), which had previously been excluded.

In setting up the evaluation quality assurance and assessment system, three overall and internationally agreed evaluation principles were used:

- **Independence and Impartiality**
- **Credibility**
- **Utility**

² UNFPA Evaluation Policy 2019

³ UNEG, Norms and Standards for Evaluation, June 2016

⁴ UNFPA Evaluation Policy 2019

Independence and Impartiality

Independence and impartiality are fundamental to the credibility and utility of evaluation. Independence of evaluation entails the ability of evaluators to work without the undue influence of any party. It includes, for instance, freedom for evaluators to choose the design of the evaluation and to select suitable methods and tools for data collection and analysis. Independence also requires that evaluators must not have been directly responsible for the policy setting, design or management of the subject being evaluated. Independence should lead to impartiality, key elements of which are objectivity, professional integrity and absence of bias.

Both independence and impartiality require transparency and should be ensured by the evaluation manager throughout all phases or stages of an evaluation process. Once the evaluation is completed, to uphold these principles, all UNFPA evaluation reports are publicly available on a database maintained by the Evaluation Office.⁵

Credibility

Credibility requires that both independence and impartiality are upheld, but must also be demonstrated and ensured through the competence of the evaluator(s) and the evaluation manager.

The quality of evaluation depends on the design and methodological rigor of the exercise, a transparent and inclusive evaluation process, the reliable use of data, and the ability to perform systematic, logical and sound analysis, reported in a structured way. This, therefore, requires that evaluators hold technical skills on evaluation design, methodology, data collection techniques and a high degree of analytical abilities. A commitment to reflective practice and communication skills are also necessary. Evaluation managers must have the skills to effectively manage the process, including the necessary technical knowledge.

Utility

Utility requires that evaluations be understood as a tool to influence change and stimulate learning.

For an evaluation to be used, it must be designed from the onset with careful consideration to how the evaluation – in both process and in its resulting products – will impact use. To facilitate this, evaluation managers ought to be skilled in communication, ensuring that reports are concise and easy to read and understand.

To further strengthen utility, an evaluation must be planned/executed with the intention that its results be used. It should be undertaken at a time when its results can meaningfully inform decision-making. It should also be relevant to the needs of intended users, which requires the participation of intended users throughout the evaluation process, including through stakeholder consultations (an important upshot of which is improved ownership).

The principles of utility and use are thus closely linked,⁶ with utilization-focused evaluations in particular enhancing the potential use of evaluations.⁷ The UNFPA evaluation policy reflects this principle, stating

⁵ The UNFPA Evaluation website can be found here: <https://www.unfpa.org/evaluation> and the database of evaluation reports is available here: <https://web2.unfpa.org/public/about/oversight/evaluations/>

⁶ Utilization-Focused Evaluation, 4th Edition, Michael Quinn Patton, 2008

⁷ UNFPA Evaluation Policy, 2013

that evaluations should be “...planned and conducted ensuring national ownership and leadership of evaluation processes by rights holders and duty bearers.”⁸

In addition to these principles, human rights and gender equality, disability inclusion, and ethical dimensions were incorporated into the evaluation quality assurance and assessment system.

A high-quality evaluation should integrate gender equality throughout the evaluation process - from determining the evaluation scope, criteria and questions through to using a gender responsive methodology, methods and data analysis techniques. The findings, conclusions and recommendations of an evaluation should also reflect a gender analysis.⁹ Both UNEG and UNFPA are committed to ensuring evaluations are gender and human rights responsive. UNEG has, for example, developed guidance on integrating gender and human rights in evaluation,¹⁰ and UNFPA has directly integrated the UN-SWAP on Gender Equality and the Empowerment of Women Evaluation Performance Indicator criteria among the seven assessment criteria used to assess the quality of evaluation reports, ensuring that gender equality is seen as a critical dimension of evaluation quality.

Similarly, a high-quality evaluation should also mainstream disability inclusion into its evaluation processes and products. This may include considering disability inclusion in specific evaluation questions or criteria, developing a methodology that integrates disability inclusions as a cross-cutting concern, or reflecting disability inclusion in the findings, conclusions and recommendations of an evaluation report. For this purpose, this guidance aligns with the United Nations Disability Inclusion Strategy and the related technical notes to operationalize the strategy.¹¹

In addition, UNFPA places great importance on ethical principles to be upheld when conducting evaluations, in line with UNEG Ethical Guidelines for Evaluation and the UNEG Code of Conduct for Evaluation.¹²

4. Translating the principles: Quality assurance and assessment in practice

The evaluation principles set forth in this guidance form the basis of the quality assurance and assessment system at UNFPA, and ground the evaluation function’s definition of high-quality evaluations.

The evaluation principles are translated directly into the tools, templates, and support provided by UNFPA to help facilitate and ensure the quality of evaluation reports. On quality assurance, for example, the Handbook on How to Design and Conduct a Country Programme Evaluation at UNFPA has been revised and updated, templates have been developed to further strengthen the quality of the terms of reference and the design/inception reports, and evaluation teams are vetted to ensure capacity to undertake an

⁸ UNFPA Evaluation Policy, 2013

⁹ UNEG Guidance Document – UN-SWAP Evaluation Performance Indicator Technical Note, August 2014

¹⁰ UNEG Guidance Document – Integrating Human Rights and Gender Equality in Evaluations, August 2014

¹¹ [UN Disability Inclusion Strategy, 2019](#)

[Technical Notes UN Disability Inclusion Strategy Entity Accountability Framework, 2019](#)

¹² UNEG Ethical Guidelines for Evaluation and UNEG Code of Conduct for Evaluation in the UN system, June 2008

evaluation. These quality assurance measures reflect the evaluation function's commitment to independence/impartiality, credibility and utility.

The principles are further reflected in UNFPA's approach to quality assessment. The evaluation quality assessment (EQA) grid details seven quality criteria against which UNFPA evaluation reports are assessed. The criteria are further unpacked into sub-criteria, which themselves, if met, signal confidence in an evaluation's independence/impartiality, credibility and utility. Sub-criteria are weighted differently based on the criteria's overall importance to the quality of the evaluation report. The approach to weighting sub-criteria is presented in the EQA grid that follows this section.

Approaching the sub-criteria of the EQA grid

In 2020, adjustments were made to the template to respond to issues raised in the application of the criteria as well as to incorporate disability inclusion dimensions across all criteria.

Under each quality criterion, the relative importance (and weight) of the sub-criteria varies, as detailed in the EQA grid presented below. For the quality assessment of quality criteria 1 to 6, core sub-criteria have been identified in **bold** text. These sub-criteria bear disproportionately on the quality of the evaluation report and, therefore, on the rating of the overall criteria. The absence of any of the core sub-criteria identified for a given quality criterion will result in an "unsatisfactory" rating for the overall quality criterion.

As an example, for quality criterion 1 (*Structure and clarity of reporting*), two core sub-criteria have been identified: sub-criteria 1.1 (*Is the report structured in logical way and easy to read and understand?*), and sub-criteria 1.4 (*Is an executive summary written as stand-alone section etc.?*). If any of these core sub-criteria is assessed as "No" (i.e. sub-criterion not met), the overall rating of quality criterion 1 will be "unsatisfactory", regardless of the assessment given to the other sub-criteria. If all identified core sub-criteria are either fully or partially met, the overall rating of the quality criterion will range from "fair" to "very good", based on the combination of values of the other sub-criteria.

A note on the "very good" rating: a "very good" rating should be used to highlight an exemplary level of quality, a good practice and/or a model to follow. As such, all sub-criteria should be met for a quality criterion to be rated as "very good".

Quality assessment of criterion 7 (*gender equality and the empowerment of women*) directly mirrors the UN-SWAP Evaluation Performance Indication Scoring Tool and is, as such, subject to its associated rating scale.

The EQA grid below outlines the specific combinations of values under each quality criterion and the resulting overall rating.

| Quality Assessment Criteria <i>(Annotations in italics)</i> | Insert <i>assessment level</i> followed by main <u>comments</u> | |
|---|---|---|
| <p>(1) Structure and clarity of reporting</p> <p><i>To ensure the report is comprehensive and user-friendly</i></p> <p>1. Is the report structured in a logical way? Is the report easy to read and understand (i.e. written in an accessible language appropriate for the intended audience) with minimal grammatical, spelling or punctuation errors? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?</p> <p>2. Is the report of a reasonable length? (maximum pages for the main report, excluding annexes: 60 for institutional evaluations; 70 for CPEs; 80 for thematic evaluations)</p> <p><i>*If rated YES, the report is within the indicated maximum page length. If rated PARTIAL, the report exceeds the maximum page limit by 1- 5 pages. If rated NO, the report exceeds the maximum page limit beyond 5 pages.</i></p> <p>3. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological and data collection tools used (e.g. interview guides; focus group notes, outline of surveys)?</p> <p><i>*If rated YES, the report contains all the annexes indicated. If rated PARTIAL, the report is missing the ToRs or the bibliography in the annexes. If rated NO, the report is missing any of the following annexes: a list of interviewees; the evaluation matrix; methodological and data collection tools used (e.g. interview guides; focus group notes, outline of surveys).</i></p> | Yes No Partial | <p>Assessment Level:</p> <p>A rating of “Very Good” will result if:</p> <p>1,2,3,4,5,6,7: Yes</p> <p>A rating of “Unsatisfactory” will result if:</p> <p>1,4: No (For any of the two)</p> <p>2,3,5: Yes/No/Partial</p> <p>A rating of “Good” will result if:</p> <p>1,4 Yes</p> <p>2,3, 5: Yes/Partial</p> <p>A rating of “Fair” will result if:</p> <p>Any combination different from the above.</p> |

| | | |
|---|-------------------------------|--|
| <p><i>Executive summary</i></p> <p>4. Is an executive summary written as a stand-alone section, presenting the i) Purpose; ii) Objectives, scope and brief description of interventions; iii) Intended audience; iv) Methodology; iv) Main results; Vi) Conclusions and recommendations</p> <p>5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?</p> <p><i>* If rated YES, the executive summary is within the indicated maximum page limit. If rated PARTIAL, the executive summary exceeds the maximum page limit by 1 to 2 pages. If rated NO, the executive summary exceeds the maximum page limit beyond 2 pages.</i></p> | | |
| <p>(2) Design and methodology</p> <p><i>To ensure that the evaluation is put within its context</i></p> <p>1. Is the development and institutional context of the evaluation clearly described and constraints explained?</p> <p>2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?</p> <p><i>To ensure a rigorous design and methodology</i></p> <p>3. Is the evaluation framework clearly described in the text and in the evaluation matrix? Does the evaluation matrix establish the evaluation questions, assumptions, indicators, data sources and methods for data collection?</p> <p>4. Are the tools for data collection described and their choice justified?</p> <p><i>*All aspects of this sub-criterion should be addressed in the rating: 1) are data collection tools described (i.e. documentary review, interviews, focus group discussions etc.) and 2) is the rationale for their selection detailed.</i></p> | <p>Yes No Partial</p> | <p>Assessment Level:</p> |
| | | <p>A rating of “Very Good” will result if: 1-10: Yes</p> <p>A rating of “Unsatisfactory” will result if: 2,3: No (for any of the two) 1,4,5,6,7,8,9,10: Yes/No/Partial</p> <p>A rating of “Good” will result if: 2,3,4,5 Yes 1, 6,7,8,9,10: Yes/Partial</p> <p>A rating of “Fair” will result if: Any combination different from the above.</p> |

| | | |
|---|--|--|
| <p>5. Is there a comprehensive stakeholder map? Is the stakeholder consultation process clearly described (in particular, does it include the consultation of key stakeholders on draft recommendations)?</p> <p><i>*All aspects of this sub- criterion should be addressed in the rating: 1) is a comprehensive stakeholder map included (in either the report itself or the annexes) 2) Is the overall stakeholder consultation process described and 3) within the consultation process were key stakeholders consulted on the recommendations specifically? 4) does the evaluation stakeholder mapping and data collection methods involve vulnerable and marginalized groups, including persons with disabilities and their representative organizations?</i></p> <p>6. Are the methods for analysis clearly described for all types of data?</p> <p><i>*Data analysis methods used may include contribution analysis, qualitative comparative analysis, or descriptive statistics. Triangulation is not a method of analysis; it is a validation technique.</i></p> <p>7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?</p> <p>8. Is the sampling strategy described?</p> <p><i>*This sub- criterion assesses whether the methodological approach to determining the sample of stakeholders consulted and the sample of site visits is described. It also examines whether the evaluation report includes information on how the universe was determined; the sampling approach used (i.e. purposive); the indicators used to develop the sample to be consulted (or visited); the resulting sample; and importantly limitations to the approach (including any potential resulting bias).</i></p> | | |
|---|--|--|

| | | |
|---|-------------------------------|--|
| <p>9. Does the methodology enable the collection and analysis of disaggregated data?</p> <p><i>*At a minimum, the methodology enables data to be disaggregated by sex. Whenever possible, this sub-criterion also assesses if a systematic disaggregation of data related to population groups (e.g. persons with disability) where there are implications related to UNFPA’s portfolio/interventions for these population groups.</i></p> <p>10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?</p> <p><i>*This sub-criterion is asking about the evaluation methodology itself – specifically does the evaluation’s design/methodology ensure that the evaluation is able to assess the extent to which the country programme integrates crosscutting issues across its portfolio of work. Therefore, we’re looking to see whether, for example, evaluation questions or indicators assess/capture the extent to which a human rights based approach to the development and implementation of the country programme was used (i.e. whether the evaluation queries/assesses whether beneficiaries/partners were consulted and through design process of the country programme); or whether the evaluation’s data collection methods capture the voices/perspectives of a range of stakeholders include beneficiaries/vulnerable/marginalized groups.</i></p> | | |
| <p>(3) Reliability of data</p> <p><i>To ensure quality of data and robust data collection processes</i></p> <p>1. Did the evaluation triangulate data collected as appropriate?</p> | <p>Yes No Partial</p> | <p>Assessment Level:</p> |
| | | <p>A rating of “Very Good” will result if: 1-3: Yes</p> |

| | | |
|---|-------------------------------|--|
| <p>2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?</p> <p><i>* This sub-criterion should address both elements, namely do the evaluators identify the sources of the qualitative and quantitative data they used and do they discuss the reliability (or lack thereof) of both?</i></p> <p>3. Is there evidence that data has been collected with sensitivity to issues of discrimination and other ethical considerations?</p> <p><i>*This sub-criterion is concerned with whether there is evidence in the report that evaluators' approach to data collection was sensitive to ethical considerations (i.e. consent, confidentiality, etc.) and were not discriminatory against particular groups' participation (i.e. were interviews or focus groups held in a location, at a time, in a setting, using language/translation, that is appropriate and respectful; and facilitates the participation of a full range of stakeholders, including persons with disability). It is also interesting if evaluators noted limitations in this regard.</i></p> <p><i>Note that mentioning/referencing UNEG standards in the report does not amount to evidence that the data was actually collected with a sensitivity to ethics and discrimination; the reviewer should assess whether there is evidence in the report of the UNEG standards actually being implemented. If the UNEG documents/standards are referenced in the text, but the evaluators do not explain how/show evidence of the data being collected with a sensitivity to issues of discrimination, etc., this should be a "partial".</i></p> | | <p>A rating of "Unsatisfactory" will result if: 1,2: No (For any of the two) 3: Yes/No/Partial</p> <p>A rating of "Good" will result if: 1,2: Yes 3: Yes/Partial</p> <p>A rating of "Fair" will result if: Any combination different from the above.</p> |
| <p>(4) Analysis and findings</p> <p><i>To ensure sound analysis and credible findings</i></p> <p>1. Are the findings substantiated by evidence?</p> | <p>Yes No Partial</p> | <p>Assessment Level:</p> <p>A rating of "Very Good" will result if: 1-7: Yes</p> |

| | | |
|--|-------------------------------|---|
| <p>2. Is the basis for interpretations carefully described?</p> <p>3. Is the analysis presented against the evaluation questions?</p> <p>4. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?</p> <p><i>*Both parts of the sub-criterion should be addressed in the rating: 1) are the cause/effect links (between UNFPA contribution and outputs/contribution to results/outcomes) explained as well as the results and 2) are unintended outcomes discussed. On the latter, it should be noted in the comments whether evaluators considered/looked for unintended outcomes and noted whether there were (or were not) any; or whether the report does not mention unintended outcomes.</i></p> <p>5. Does the analysis show different outcomes for different target groups, as relevant?</p> <p>6. Is the analysis presented against contextual factors?</p> <p>7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?</p> | | <p>A rating of “Unsatisfactory” will result if: 1,4: No (For any of the two) 2,3,5,6,7: Yes/No/Partial</p> <p>A rating of “Good” will result if: 1,4: Yes 2,3,5,6,7: Yes/Partial</p> <p>A rating of “Fair” will result if: Any combination different from the above.</p> |
| <p>(5) Conclusions</p> <p><i>To assess the validity of conclusions</i></p> <p>1. Do the conclusions flow clearly from the findings?</p> <p>2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?</p> | <p>Yes No Partial</p> | <p>Assessment Level:</p> <p>A rating of “Very Good” will result if: 1,2,3: Yes</p> <p>A rating of “Unsatisfactory” will result if: 2: No 1,3: Yes/No/Partial</p> <p>A rating of “Good” will result if:</p> |

| | | |
|---|-------------------------------|---|
| <p>3. Do the conclusions appear to convey the evaluators' unbiased judgement?</p> | | <p>2: Yes 1,3: Yes/Partial</p> <p>A rating of "Fair" will result if: Any combination different from the above.</p> |
| <p>(6) Recommendations</p> <p><i>To ensure the usefulness and clarity of recommendations</i></p> <p>1. Do recommendations flow logically from conclusions?</p> <p>2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?</p> <p>3. Do recommendations appear balanced and impartial and address, as relevant, key cross cutting issues such as equity and vulnerability, disability-inclusion, gender equality and human rights?</p> <p>4. Are the recommendations prioritised?</p> | <p>Yes No Partial</p> | <p>Assessment Level:</p> |
| | | <p>A rating of "Very Good" will result if: 1-5: Yes</p> <p>A rating of "Unsatisfactory" will result if: 2: No (For any of the two) 1,3,4: Yes/No/Partial</p> <p>A rating of "Good" will result if: 2: Yes 1,3, 4: Yes/Partial</p> <p>Fair Any combination different from the above.</p> |
| <p>(7) Gender</p> <p><i>To assess the integration of Gender Equality and Empowerment of Women (GEEW)</i></p> <p>1. Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected?</p> <p><i>* This sub-criterion assesses whether the evaluation itself integrates GEEW – in its scope of analysis (i.e. in the objectives for example)</i></p> | <p>0-3</p> | <p>Assessment Level:</p> |
| | | <p>See footnote¹³</p> |

¹³ This assessment criterion is fully based on the UN-SWAP Evaluation Performance Indicator Scoring Tool. Each sub-criterion shall be equally weighted in correlation with the calculation in the tool. The EPI Scoring Tool can be found here: <http://www.unevaluation.org/document/detail/1452>

| | | |
|--|--|--|
| <p><i>or the indicators the evaluation selects against which data will be collected so that the evaluation is able to assess whether the country programme is gender responsive.</i></p> <p><i>A general note on UNFPA programming: While there may be evidence of gender being referred to as a cornerstone of UNFPA programming - in the sense that most UNFPA programmes target women and girls - this does not necessarily mean that UNFPA’s work is gender/human rights responsive. GEEW is about power and shifting resources, social norms, attitudes, laws and policies. One could work on comprehensive sexuality education, for example, in a way that further entrenches gendered norms or power dynamics (i.e. pathologizing LGBTQ communities; or reifying gender binaries by assuming heteronormativity); this would not be GEEW sensitive. Another example: one could deliver sexual and reproductive health care that fails to adequately address the diverse health needs of women (i.e. women who are disabled, older women, LGBTQI women; conceptualizes women as mothers alone; and/or holding biases against contraceptive options); again, this would not be GEEW sensitive.</i></p> <p>2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?</p> <p><i>* This sub-criterion assesses whether the evaluation criteria and evaluation questions (i.e. the evaluation itself) are gender responsive; in other words, are the criteria interpreted/operationalized and evaluations questions developed in a way that is able to capture whether (or not) gender equality/human rights/the empowerment of women has been integrated into UNFPA’s country programme/support (in the design/planning, implementation and results)?</i></p> <p>3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</p> | | |
|--|--|--|

For more information, refer to the [UNFPA Evaluation Handbook](#) on how to design and conduct a country programme evaluation and the [Guidance on disability inclusion in UNFPA evaluations](#).

All UNFPA programme-level and corporate evaluations are published on the [Evaluation Database](#), together with the Evaluation Quality Assessment and a Management Response.



**United Nations Population Fund
Evaluation Office**

605 Third Avenue
New York, NY 10158 USA

✉ evaluation.office@unfpa.org

🌐 unfpa.org/evaluation

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

▶ [UNFPA Evaluation Office](#)