

Explanatory notes for UNFPA Evaluation Quality Assessment Grid



I. Explanations regarding the Quality Assessment criteria

I. Structure and Clarity of Reporting

Does the report clearly describe the evaluation, how it was conducted, the findings of the evaluation, and their analysis and subsequent recommendations?

Is the structure *logical*? Is the report *comprehensive*?

Can the information provided be *easily understood*?

2. Executive Summary

Does it read as a stand-alone section, and is a *useful* resource in its own right?

Is it brief yet *sufficiently detailed*, presenting the main results of the evaluation, and including *key elements* such as methodology and conclusions and recommendations?

3. Design and Methodology

Is the *methodology* used for the evaluation clearly described and is the rationale for the methodological choice justified?

Have cross-cutting issues (vulnerable groups, youth and gender equality) been paid specific attention (when relevant) in the design of the evaluation?

Are key processes (tools used, triangulation, consultation with stakeholders) discussed in sufficient detail? Are *constraints* and *limitations* made explicit (including limitations applying to interpretations and extrapolations; robustness of data sources, etc.) and discussed?

4. Reliability of Data

Are *sources* of data clearly stated for both primary and secondary data?

Is it clear why case studies were selected and what purpose they serve?

Are all relevant materials related to case studies, interviews (list of interviewees, questionnaires) etc. annexed to the report?

Are the limitations, and methods to address them, discussed?

What other *data gaps* are there and how have these been addressed?

5. Findings and Analysis

Findings

Is there a *clear pathway* from data to findings, so that all findings are *evidence-based*?

Are *biases* stated and discussed?

Are *unintended* findings reported and discussed?

Analysis

Are *interpretations* of the findings understandable? Are *assumptions* clearly stated and extrapolations well explained?

Are their *limitations* (or drawbacks) discussed?

Does the analysis respond to all evaluation questions?

If not, are *omissions* (of both evaluation criteria and questions) recognized and explained?

Has the analysis examined *cause and effect* links between an intervention and its end results?

Are *contextual factors* identified and their influence discussed?

6. Conclusions

Are the conclusions organized in priority order?

Do the conclusions amount to a reasonable *judgment* of the findings and are their links to evidence made clear?

Are there any limitations and are these made clear?

Do they present an *unbiased* judgment by the evaluators of the intervention or have they been influenced by preconceptions or assumptions that have not been discussed?

7. Recommendations

Is there a *logical flow* from the conclusions to recommendations?

Are they strategic and clearly presented in a priority order which is consistent with the *prioritization* of conclusions? Are they *useful* – sufficiently detailed, targeted and likely to be implemented and lead to *further action*?

How have the recommendations *incorporated* stakeholders' views and has this affected their *impartiality*?

8. Meeting Needs

Does the report adequately address the information needs and responds to the *requirements stated in the ToRs*?

In particular does the report respond to the evaluation questions, issues or criteria identified in ToR?

2. Explanations regarding scoring and weighing

The scoring system that is attached to the grid (and which has been thoroughly thought through in order to leave as little space as possible to subjectivity) is not meant to grade or compare evaluation reports. Instead the EQA should be seen as a practical tool for the benefit of all those involved in the evaluation process with a view to improving the quality of the evaluation products. In this sense, the EQA is also and primarily meant to be used as a self-assessment tool and a means of communication between all actors involved in evaluation (CO M&E officers, consultants/experts, partners, RO M&E officers etc.) in the field.

a. Why and how to score the quality of evaluation reports?

The scoring of EQAs serves two main purposes:

- to express an objective judgment both on the overall quality of an evaluation report as well as on each evaluation criterion used in the quality assessment (synchronic approach) ;
- to assess the progress (or lack thereof) over time, either in the overall quality of UNFPA funded evaluation reports or for each specific quality criterion (diachronic approach).

As indicated in the EQA grid, the scoring scale comprises four levels: (1) unsatisfactory, (2) poor, (3) good, (4) very good.

b. Why and how to weigh the different criteria of the EQA grid?

Each EQA criterion has been associated with a weight (or a multiplying factor) which is proportionate to, and illustrates its relative importance as regards the overall quality of the report.

As you will see (Table below) the **criteria 5 (Findings and analysis)** is the most prominent of all 8 criteria as a good analysis and credible findings are considered the backbone of a good quality report.

In fact, a report containing sound analysis and credible findings is useful even if the conclusions and recommendations are poorly formulated, as sound analysis and credible findings provide the reader with accurate information on the evaluated programme as well as potentially useful “lessons learned.”

In contrast, conclusions that appear convincing or recommendations that seem well-articulated cannot and should not be used when they are not grounded in sound analysis and related robust findings.

As a result: fulfillment of ***criteria 5 is indispensable to the production of a good quality report***, and, for this reason, ***it is associated with a weight accounting for half of the total quality score***.

c. The detailed weighing scale for EQA criteria

Quality assessment criteria	Multiplying factor
5. Findings and analysis	50
6. Conclusions	12
7. Recommendations	12
8. Meeting needs	12
3. Design and methodology	5
4. Reliability of data	5
1. Structure and clarity of reporting	2
2. Executive summary	2
TOTAL	100

d. Guidance on how to compile the Scoring Grid

Insert the multiplying factor associated with the criteria in the corresponding column e.g. - if “Finding and Analysis” has been assessed as “good”, please enter the number 50 into the “Good” column. The Assessment level scoring the higher number of points will determine the overall quality of the Report

Quality assessment criteria (and Multiplying factor *)	Assessment Levels (*)			
	Unsatisfactory	Poor	Good	Very good
1. Structure and clarity of reporting (2)			2	
2. Executive summary (2)	2			
3. Design and methodology (5)			5	
4. Reliability of data (5)	5			
5. Findings and analysis (50)			50	
6. Conclusions (12)		12		
7. Recommendations (12)		12		
8. Meeting needs (12)	12			
TOTAL	19	24	57	

Therefore, in this example, as the highest score is 57 in the Good column, the overall Assessment Level is Good.