

## SECTION 6

# EVALUATION QUALITY ASSESSMENT



# CONTENTS

<b>6. QUALITY ASSESSMENT OF DECENTRALIZED EVALUATIONS .....</b>	<b>1</b>
6.1 Purpose and scope .....	1
6.2 Quality assessment process .....	2
6.3 Roles and responsibilities.....	3
6.4 Quality assessment review pool .....	4
6.5 Quality assessment reporting .....	4
6.6 Quality assessment sections and weighting .....	4
6.7 Quality assessment question ratings .....	5
6.8 Quality assessment tool.....	6
6.9 Supporting documentation.....	6
6.10 Quality assessment questions.....	7
6.10.1 Evaluation TOR, evaluation design (GEF and UNDP) .....	7
6.10.2 Evaluation report structure, methodology and data sources.....	8
6.10.3 Cross-cutting issues.....	10
6.10.4 Evaluation findings, conclusions and recommendations.....	11
6.10.5 GEF Independent Evaluation Office validation of terminal evaluation results.....	12
6.10.6 General findings and lessons learned .....	14
6.10.7 Summary quality assessment result .....	15
6.11 UN-SWAP evaluation performance indicator and assessment.....	16
6.11.1 What is the UN-SWAP evaluation performance indicator?.....	16
6.11.2 The evaluation performance indicator criteria and scorecard .....	17
6.11.3 The assessment process.....	17
6.11.4 Reporting.....	18
Figure 1. Quality assessment process .....	3
Figure 2. ERC quality assessment summary report by region .....	4
Figure 3. Quality assessment rating scale.....	5

## 6. QUALITY ASSESSMENT OF DECENTRALIZED EVALUATIONS

Section 6 describes the UNDP Independent Evaluation Office (IEO) system for quality assessment of decentralized evaluations, including the purpose, roles and responsibilities, processes and tools. In addition, the section explains the Evaluation Performance Indicator for the United Nations System-wide Action Plan (SWAP) on Gender Equality and the Empowerment of Women, and how independent and decentralized evaluations are assessed to provide UNDP data for this indicator.

High quality evaluations are critical for results-based management, knowledge generation, and accountability to programme partners. One of the requirements of the UNDP Evaluation Policy is that programme units—headquarters bureaux, regional bureaux and country offices—ensure that evaluations inform programme management and contribute to development results.<sup>1</sup> There is therefore increased emphasis to strengthen support for decentralized evaluations (those carried out by programme units) in order to improve their compliance with the Evaluation Policy, improve the quality of evaluations and increase the use of evaluations by policymakers and stakeholders.

The IEO annually assesses the quality of decentralized evaluations and reports on the results to the UNDP Executive Board. The quality assessment process supports the improvement of the quality of evaluative evidence including findings, coverage and scope, as well as recommendations, through the independent analysis of evaluations undertaken by programme units. The quality assessment process also supports management of evaluations and implementation of the evaluation plan by programme units, as well as oversight by regional bureaux, the Bureau for Policy and Programme Support (BPPS) and IEO. This quality assessment system for decentralized evaluation reports facilitates uniformity and consistency of the quality assessment process and reporting.

### 6.1 Purpose and scope

Using a set of parameters, a rating system and weightings, the quality assessment of an evaluation report provides an assessment of an evaluation's design, the quality of its findings and evaluative evidence and the robustness of its conclusions and recommendations. For Global Environment Facility (GEF) evaluations, the assessment also includes the extent to which project outputs and/or programme outcomes were achieved (or are expected to be achieved).

The purposes of a quality assessment of an evaluation report include:

- Improving the quality of evaluative evidence to better manage contributions to development results.
- Supporting accountability by providing an independent assessment of the quality of decentralized evaluation reports to the UNDP Executive Board and management.
- Strengthening consistency in evaluation reporting and quality across projects.
- Supporting bureau oversight functions by providing concurrent feedback through detailed analysis of the quality of evaluation reports, with recommendations for their improvement.

---

<sup>1</sup> UNDP Evaluation Policy, <http://web.undp.org/evaluation/policy.shtml>

- Contributing to corporate lessons learned by drawing from good evaluations in the annual report on evaluation.

These guidelines enhance the quality standards of decentralized evaluations such as utility, clarity of objectives to all stakeholders, credibility, accuracy and reliability of the evaluability evidence, transparency of the judgements, and depth and clarity of reporting.

Quality assessments are carried out for all decentralized evaluations conducted by UNDP, as well as the United Nations Capital Development Fund and United Nations Volunteers programme, outcome, project and programme evaluations and thematic evaluations. Feedback from IEO can be used by programme units and country offices to make adjustments that will strengthen areas of the evaluative evidence and the report, as well as adjust the management and implementation of evaluations to ensure usable findings and recommendations and the overall utility of decentralized evaluation reports. The quality assessment questions are in line with and reflect the UNDP quality standards for programming.<sup>2</sup>

The scope of analysis of GEF evaluation reports is broader than for other UNDP evaluation reports. GEF analysis includes an assessment of project documentation (e.g. project objectives, project or programme planning and implementation) and an analysis of the validity of an evaluation's findings and conclusions.

## 6.2 Quality assessment process

The key steps of the quality assessment process are as follows (see also figure 1):

- 1. Posting evaluations to the Evaluation Resource Center (ERC)<sup>3</sup>**
  - The programme unit posts an electronic and printable copy of the terms of reference (TOR) for an evaluation and the final evaluation report on the ERC within two weeks of completion.
  - Only final documents should be uploaded. Drafts should not be uploaded as the ERC is a public website.
  - The management response and key actions should be uploaded within six weeks of completion of the report.
- 2. Verification**
  - The IEO will verify if a report posted on the ERC is part of the programme unit evaluation plan and whether it is the final document.
  - If a TOR or evaluation appear to be in draft and not final versions, or if supporting annexes are not uploaded, IEO will contact the country office and regional office to ensure that the correct documents are uploaded.<sup>4</sup>

<sup>2</sup> Access at:

[https://popp.undp.org/\\_layouts/15/WopiFrame.aspx?sourcedoc=/UNDP\\_POPP\\_DOCUMENT\\_LIBRARY/Public/PPM\\_Programming%20Standards\\_Quality%20Standards%20for%20Programming.docx&action=default](https://popp.undp.org/_layouts/15/WopiFrame.aspx?sourcedoc=/UNDP_POPP_DOCUMENT_LIBRARY/Public/PPM_Programming%20Standards_Quality%20Standards%20for%20Programming.docx&action=default)

<sup>3</sup> Access at: <http://erc.undp.org>

<sup>4</sup> The ERC is a public website and therefore all documents should be final and of high quality. The quality assessment ratings are available only to UNDP.

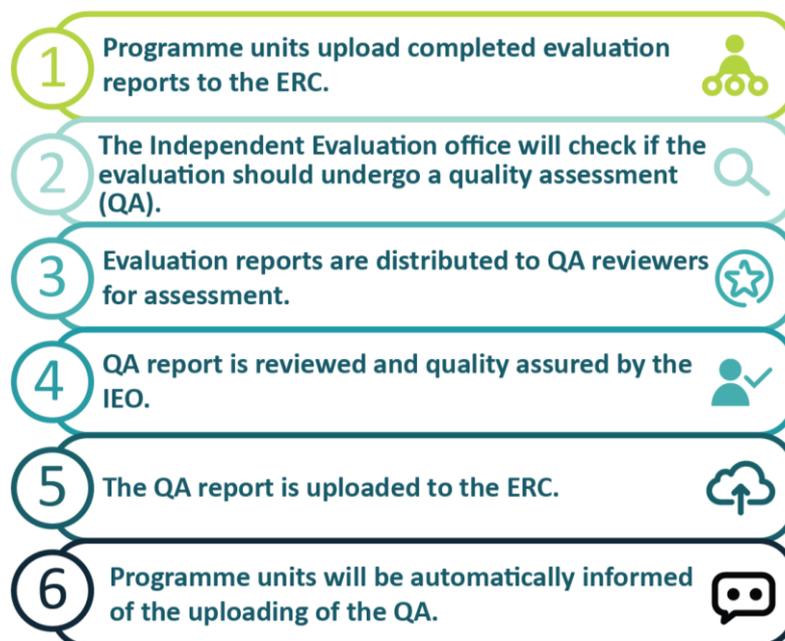
### 3. Quality assessment

- The IEO sends the evaluation report to a contracted quality assessment reviewer to conduct a quality review.
- The quality assessment rating is made available on the ERC typically within two weeks of completion and submission of the quality assessment report.

### 4. Feedback

- Upon receiving the quality assessment report from the reviewer, the IEO reviews the report and then makes it available to the respective programme unit through the ERC.

Figure 1. Quality assessment process



## 6.3 Roles and responsibilities

IEO has the overall responsibility for evaluation quality assessment and reporting and providing timely feedback to programme units.

Regional bureaux should oversee the quality assessment process and use it to highlight weaknesses and challenges in the implementation of evaluations across their regions and within specific country programmes. The ERC offers an overview tool to show the quality of evaluations at regional and country office levels. In cases where evaluations are consistently below a satisfactory level, regional evaluation focal points should work closely with country offices to address implementation issues and ensure that programme units understand the issues in the evaluation process highlighted and detailed in the quality assessment process.

Equally, BPPS and IEO support regions to address the issues in evaluation implementation highlighted through the quality assessment process and support bureaux to address issues consistently highlighted.

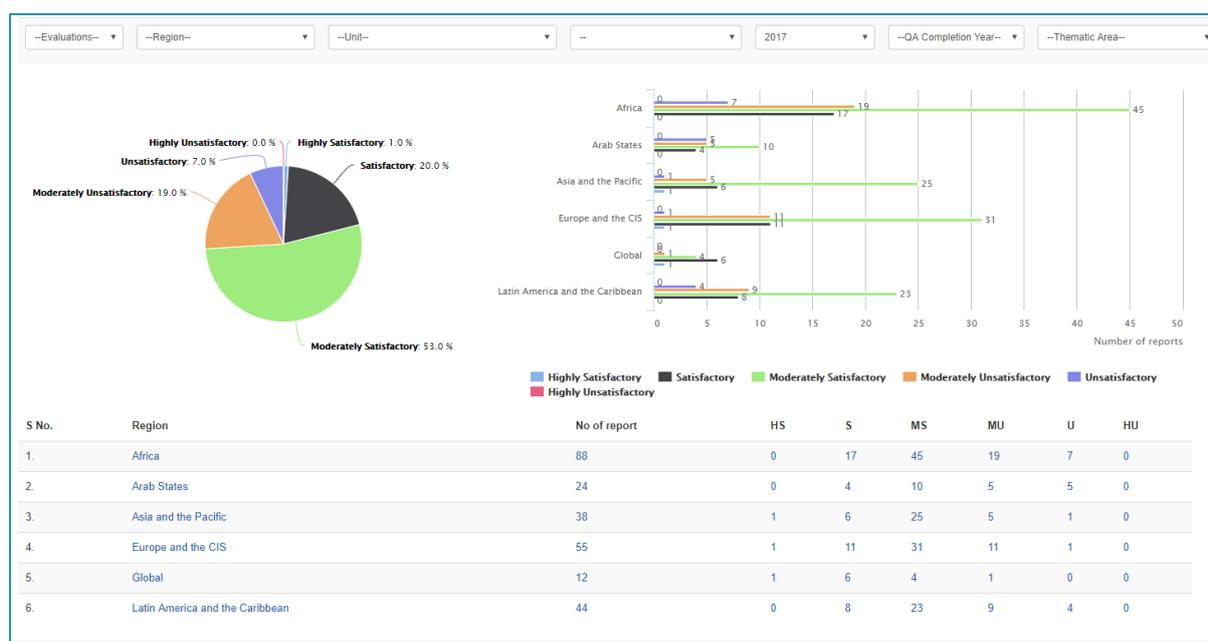
## 6.4 Quality assessment review pool

In order to ensure the quality and consistency of evaluation report assessments, the IEO retains a pool of expert quality assessment reviewers, who are experienced evaluators with a detailed knowledge of UNDP thematic areas and evaluation approaches as well as global, regional and country knowledge and experience. To ensure the uniformity and consistency of evaluation quality assessments, the reviewers are oriented in the application of the quality assessment tools and the IEO periodically verifies the quality assessment process to ensure reliability.

## 6.5 Quality assessment reporting

A quality assessment report for an individual evaluation will be made available as soon as the IEO performs quality assurance checks on the assessment (normally **within two weeks of completion** and submission of the quality assessment report). Results at the global, regional and country office levels are available through the ERC.

Figure 2. ERC quality assessment summary report by region



Annually, IEO will report on the results of the quality assessment process through its annual report on evaluation, along with a more detailed annual quality assessment report, which is distributed to headquarters and regional bureaux for distribution and follow-up with country offices.

## 6.6 Quality assessment sections and weighting

The key parameters of a quality assessment draw on the basic quality requirements for acceptable evaluation reports as outlined in the Evaluation Guidelines. Overall, the quality assessment process

includes four weighted sections and 39 questions. Questions may be left unrated by reviewers where they find them not relevant due to the direction of the TOR or the context of the intervention under evaluation.

Quality assessment sections include:

- **Terms of reference:** Five questions weighted 15 percent
  - Do the TOR appropriately and clearly outline the purpose, objectives, criteria, and key questions for the evaluation?
- **Evaluation structure, methodology and data sources:** 16 questions weighted 30 percent
  - Is the evaluation well structured, with a clearly articulated set of objectives, criteria and methodology that are fully described and appropriate?
- **Cross-cutting issues:** Eight questions weighted 15 percent
  - Does the evaluation adequately review and address cross-cutting issues such as gender, human rights, disabilities and vulnerable groups?
- **Findings, conclusions and recommendations:** Nine questions weighted 40 percent
  - Are **findings** appropriate and based on the evaluation criteria (e.g. relevance, coherence, effectiveness, efficiency, sustainability and impact), and do they respond directly to the evaluation questions?
  - Do the **conclusions** go beyond findings and identify underlying priority issues? Do the conclusions present logical judgements based on findings that are substantiated by evidence?
  - Are the **recommendations** relevant to the subject and purposes of the evaluation, and are they supported by evaluation evidence?

Quality assessments of GEF terminal evaluations include an additional section in which the quality assessment reviewer validates the evaluation’s ratings or recommends adjustment. GEF midterm reviews are currently not quality assessed though they are included in the evaluation plan.

## 6.7 Quality assessment question ratings

Quality assessment questions under each section are scored using a six-point rating system ranging from highly satisfactory (6) to highly unsatisfactory (1) or not applicable (0) (see figure 3). The rating scale assesses whether an evaluation has met expectations, norms and criteria. While ratings of 4, 5 and 6 could be considered satisfactory, if all UNDP evaluation requirements are met then an evaluation should receive at a minimum rating of 5 (satisfactory), which is the benchmark for a good evaluation.

Figure 3. Quality assessment rating scale

Code	Rubric for assigning rating		Value
HS	<b>Highly satisfactory</b>	All parameters were fully met and there were no shortcomings in the evaluation report	6
S	<b>Satisfactory</b>	All parameters were fully met with minor shortcomings in the evaluation report	5

<b>MS</b>	<b>Mostly satisfactory</b>	The parameters were partially met with some shortcomings in the evaluation report	<b>4</b>
<b>MU</b>	<b>Mostly unsatisfactory</b>	More than one parameter was unmet with significant shortcomings in the evaluation report	<b>3</b>
<b>U</b>	<b>Unsatisfactory</b>	Most parameters were not met and there were major shortcomings in the evaluation report	<b>2</b>
<b>HU</b>	<b>Highly unsatisfactory</b>	None of the parameters were met and there were severe shortcomings in the evaluation report	<b>1</b>
<b>N/A</b>	<b>Not Applicable</b>	Not Applicable	<b>unscored</b>

## 6.8 Quality assessment tool

The quality assessment tool is accessible from the ERC website (<http://erc.undp.org>). Login is restricted to registered monitoring and evaluation (M&E) focal points. M&E focal points should share the results of evaluation quality assessments with evaluation commissioners and managers.

Quality assessment reviewers use drop-down menus to assign content ratings and detailed comments supporting their ratings. Overall scores, using the weightings above, are assigned automatically through the ERC. Scoring and comments can be found under each evaluation when a quality assessment is completed.

## 6.9 Supporting documentation

All supporting documentation for evaluations being assessed is made available via the ERC and to the quality assessment reviewer.

**For UNDP projects** the documentation includes:

- The TOR for the evaluation (key document for the quality assessment).
- Final evaluation report and annexes (key document for the quality assessment).
- Project/ evaluation information (project details, evaluation budget and time frame).
- Evaluation lessons and findings.
- Evaluation recommendations.
- Management response and key actions.

For the purposes of the quality assessment, the **TOR and final evaluation report** are the key documents, including all annexes.

The ERC will contain the same information for **GEF project terminal evaluations**. However, in order to further validate the terminal evaluation ratings for project implementation, GEF will provide additional information to quality assessment reviewers via the IEO. These documents are not available on the ERC at present. Additional documentation includes:

- The project concept note and identification forms (PIF/Pdf A &B).

- Project document (ProDoc), including results framework.
- Project implementation reviews (APR/PIR).
- Tracking tools (as available).
- Midterm evaluation, if carried out.
- Project implementation action plan.

## 6.10 Quality assessment questions

### 6.10.1 Evaluation TOR, evaluation design (GEF and UNDP)

Does the TOR appropriately and clearly outline the purpose, objectives, criteria and key questions for the evaluation and give adequate time and resources? (Section 4.3.2)	
1.1	<p><b>Does the TOR clearly outline the focus for the evaluation in a logical and realistic manner?</b></p> <ul style="list-style-type: none"> <li>▪ Follows the proposed structure detailed in the UNDP evaluation guidelines</li> <li>▪ Includes the evaluation purpose, scope, and objectives</li> <li>▪ Includes outputs and/or outcomes to be evaluated</li> <li>▪ Provides evaluation context and detail</li> <li>▪ Includes information regarding the results framework and the theory of change in the main text or annexes</li> <li>▪ Includes information about the project / programme beneficiaries (type, sex, number)</li> </ul>
1.2	<p><b>Does the TOR clearly detail timescales and allocation of days for the evaluation?</b></p> <ul style="list-style-type: none"> <li>▪ There is a timescale for the scope and focus of the evaluation</li> <li>▪ The allocation of days across the evaluation is detailed and appropriate given the scope of the evaluation</li> <li>▪ There is an outline for the evaluation team size which recognizes the needs and scope of the evaluation</li> <li>▪ Roles and responsibilities of team members (where a team is called for) are delineated</li> </ul>
1.3	<p><b>Does the TOR clearly outline the evaluation implementation and management arrangements?</b></p> <ul style="list-style-type: none"> <li>▪ A clear role for evaluation partners is outlined</li> <li>▪ A feedback mechanism is clearly outlined</li> </ul>
1.4	<p><b>Is the proposed outline of the evaluation's approach and methodology clearly detailed in the TOR?</b></p> <ul style="list-style-type: none"> <li>▪ The number of evaluation questions seems appropriate given the scope of the evaluation</li> <li>▪ General methodological approach is outlined</li> <li>▪ Data required, sources and analysis approaches are outlined</li> <li>▪ Funding analysis requirements and funding data are outlined</li> </ul>
1.5	<p><b>Do the TOR include a detailed request to the evaluator to include gender, vulnerable groups, disability issues, and/or human rights in the evaluation? (non-GEF evaluations)</b></p>

	<ul style="list-style-type: none"> <li>▪ Details for gender, vulnerable groups, disability issues and/or human rights specific questions are requested in the TOR</li> <li>▪ The TOR outline proposed tools, methodologies, and data analysis to meet this requirement</li> </ul>
--	---

## 6.10.2 Evaluation report structure, methodology and data sources

**Are the evaluation objectives, criteria, methodology and data sources fully described and are they appropriate given the subject being evaluated and the reasons for carrying out the evaluation?**

### STRUCTURE

	<p><b>Is the evaluation report well balanced and structured?</b></p> <ul style="list-style-type: none"> <li>- Follows the proposed evaluation report structure detailed in the UNDP Evaluation guidelines (section 4, 4.4.5 and annex 4)</li> </ul>
2.1	<p>If not followed, does the report structure used allow for a well-balanced report?</p> <ul style="list-style-type: none"> <li>- The report includes sufficient and comprehensible background information</li> <li>- The report is a reasonable length</li> <li>- The required annexes are provided</li> </ul>
2.2	<p><b>Does the evaluation report clearly address the objectives of the evaluation as outlined in the TOR?</b></p>

### METHODOLOGY

2.3	<p><b>Is the evaluation methodological approach clearly outlined?</b></p> <ul style="list-style-type: none"> <li>- Any changes from the proposed approach are detailed with reasons why</li> </ul>
2.4	<p><b>Is the nature and extent of stakeholder roles and involvement explained adequately?</b></p>
2.5	<p><b>Does the evaluation clearly assess the project/ programme's level of relevance/ coherence?</b></p>
2.6	<p><b>Does the evaluation clearly assess the project/ programme's level of effectiveness?</b></p>
2.7	<p><b>Does the evaluation clearly assess the project/ programme's level of efficiency?</b></p>
2.8	<p><b>Does the evaluation clearly assess the project/ programme's level of sustainability?</b></p>

### DATA COLLECTION

2.9	<p><b>Are data collection methods and analysis clearly outlined?</b></p> <ul style="list-style-type: none"> <li>- Data sources are clearly outlined (including triangulation methods)</li> <li>- Data analysis approaches are detailed</li> </ul>
-----	---

	<ul style="list-style-type: none"> <li>- Data collection methods and tools are explained</li> </ul>
2.10	<p><b>Is the data collection approach and analysis adequate for the scope of the evaluation?</b></p> <ul style="list-style-type: none"> <li>- A comprehensive set of data sources (especially for triangulation) is included where appropriate</li> <li>- A comprehensive set of quantitative and qualitative surveys, and analysis approaches is included where appropriate</li> <li>- Clear presentation of data analysis and citation within the report</li> <li>- Meetings and surveys with stakeholders and beneficiary groups are documented, where appropriate</li> </ul>
2.11	<p><b>Are any changes to the evaluation approach or limitations in implementation clearly explained?</b></p> <ul style="list-style-type: none"> <li>- Issues with access to data or verification of data sources</li> <li>- Issues in the availability of interviewees</li> <li>- Outline of how these constraints were addressed</li> </ul>

## REPORT CONTENT

2.12	<p><b>Does the evaluation draw linkages to the UNDP country programme strategy and/ or UNDAF/ UNSDCF?</b></p> <ul style="list-style-type: none"> <li>- It evaluates the programme/ project theory of change and its relevance</li> <li>- It analyses the linkage of the project/ programme being evaluated to the UNDP country programme strategy</li> <li>- It makes linkages to the United Nations Sustainable Development Cooperation Framework (UNSDCF)</li> </ul>
2.13	<p><b>Does the evaluation draw linkages to related national government strategies and plans in the sector/area of support?</b></p> <ul style="list-style-type: none"> <li>- The evaluation discusses how capacity development, or the strengthening of national capacities, can be addressed</li> </ul>
2.14	<p><b>Does the evaluation detail project funding and provide funding data (especially for GEF)?</b></p> <ul style="list-style-type: none"> <li>- Variances between planned and actual expenditures are assessed and explained</li> <li>- Observations from financial audits completed for the project are considered</li> </ul>
2.15	<p><b>Does the evaluation include an assessment of the project/ programme's initial results framework, M&amp;E design, implementation, and its overall quality?</b></p> <ul style="list-style-type: none"> <li>- Monitoring data presented and sufficiently detailed to enable analysis for the evaluation</li> <li>- Data was disaggregated by sex and vulnerable groups</li> </ul>
2.16	<p><b>Does the evaluation identify ways in which the programme/ project has produced a catalytic role and demonstrated: the production of a public good; demonstration; replication; and/or scaling up? (GEF ONLY)</b></p>
2.17	<p><b>Are all indicators in the logical framework assessed individually, with final achievements noted?</b></p>

### 6.10.3 Cross-cutting issues

<b>Does the evaluation report address gender and other key cross-cutting issues?</b>	
<b>3.1</b>	<b>Where relevant, does the evaluation adequately include and analyse the intervention’s impact on gender, human rights, disabilities and vulnerable groups?</b>
<b>3.2</b>	<b>Does the report analyse the poverty and environment nexus or sustainable livelihood issues, as relevant?</b>
<b>3.3</b>	<b>Does the report discuss disaster risk reduction and climate change mitigation and adaptation issues where relevant?</b>
<b>3.4</b>	<b>Does the report discuss crisis prevention and recovery issues, as relevant?</b>
<b>3.5</b>	<p><b>Are gender equality and empowerment of women integrated in the evaluation scope, and are the evaluation criteria and questions designed in a way that ensures data related to gender equality and empowerment of women will be collected?</b></p> <ul style="list-style-type: none"> <li>- The evaluation includes an objective specifically addressing gender equality and/or human rights issues and/or gender was mainstreamed in other objectives</li> <li>- A stand-alone evaluation criterion on gender and/or human rights was included in the evaluation framework or mainstreamed into other evaluation criteria</li> <li>- One or several dedicated gender equality and empowerment of women evaluation questions were integrated into the evaluation</li> </ul>
<b>3.6</b>	<p><b>Were gender-responsive methodology, methods and tools, and data analysis techniques selected?</b></p> <ul style="list-style-type: none"> <li>- The evaluation specifies how gender issues are addressed in the methodology, including how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex</li> <li>- The evaluation methodology employs a mixed-methods approach, appropriate to evaluating gender equality and empowerment of women considerations</li> <li>- A diverse range of data sources and processes are employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility</li> <li>- The evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate</li> </ul>
<b>3.7</b>	<p><b>Do the evaluation findings, conclusions and recommendation reflect a gender analysis?</b></p> <ul style="list-style-type: none"> <li>- The evaluation has a background section that includes analysis of specific social groups affected and/ or spelling out the relevant instruments or policies related to gender equality and human rights</li> <li>- The findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/ or disaggregates quantitative data by sex, where applicable</li> </ul>

	<ul style="list-style-type: none"> <li>- Unanticipated effects of the intervention on gender equality and human rights are described</li> <li>- The evaluation report provides specific recommendations addressing issues of gender equality and empowerment of women, and priorities for action to improve gender equality and empowerment of women or the intervention or future initiatives in this area</li> </ul>
<b>3.8</b>	<b>Does the evaluation consider disability issues?</b> <ul style="list-style-type: none"> <li>- Evaluation questions cover different aspects of disability inclusion</li> <li>- Evaluation findings and analysis provide data and evidence on disability inclusion</li> <li>- Evaluation conclusions and/ or recommendations reflect the findings on disability inclusion</li> </ul>
<b>3.9</b>	<b>Does the evaluation draw linkages to the Sustainable Development Goals (SDGs) and relevant targets and indicators for the area being evaluated?</b>
<b>3.10</b>	<b>Does the terminal evaluation adequately address social and environmental safeguards, as relevant? (GEF ONLY)</b>

#### 6.10.4 Evaluation findings, conclusions and recommendations

This section details all the evaluation results, findings, conclusions and recommendations. Both GEF and UNDP projects use the same questions for quality assessment.

**Does the report clearly and concisely outline and support its findings, conclusions and recommendations?**

#### **FINDINGS AND CONCLUSIONS**

<b>4.1</b>	<b>Does the evaluation report contain a concise and logically articulated set of findings?</b> <ul style="list-style-type: none"> <li>- The findings are structured around the evaluation criteria and evaluation questions</li> <li>- The findings are detailed and supported by evidence</li> <li>- The findings go beyond an analysis of activity implementation</li> </ul>
<b>4.2</b>	<b>Does the evaluation report contain a concise and logically articulated set of conclusions which are stand-alone in nature?</b>
<b>4.3</b>	<b>Does the evaluation report contain a concise and logically articulated set of lessons learned?</b> <ul style="list-style-type: none"> <li>- The lessons learned are substantive</li> <li>- The lessons learned are appropriately targeted at different implementation and organizational levels</li> </ul>
<b>4.4</b>	<b>Do the findings and conclusions relate directly to the objectives of the project /programme and the evaluation?</b> <ul style="list-style-type: none"> <li>- They relate directly to the objectives of the project/ programme</li> <li>- They relate to the objectives of the evaluation as outlined in the TOR for the evaluation</li> </ul>

4.5	<b>Are the findings and conclusions supported with data and interview sources?</b> <ul style="list-style-type: none"> <li>- Constraints in access to data and interview sources are detailed</li> </ul>
4.6	<b>Do the conclusions build on the findings of the evaluation?</b> <ul style="list-style-type: none"> <li>- The conclusions go beyond the findings and present a balanced picture of the strengths and limitations of the intervention</li> </ul>
4.7	<b>Are risks discussed in the evaluation report?</b>
<b>RECOMMENDATIONS</b>	
4.8	<b>Are the evaluation recommendations clear, concise, realistic and actionable?</b> <ul style="list-style-type: none"> <li>- They are reasonable given the size and scope of the project/ programme</li> </ul>
4.9	<b>Are recommendations linked to country programme outcomes and strategies and actionable by the country office?</b> <ul style="list-style-type: none"> <li>- Guidance is given for implementation of the recommendations</li> <li>- Recommendations identify implementing roles (UNDP, government, programme, stakeholder, other)</li> </ul>

### 6.10.5. Validation of the ratings given by GEF terminal evaluations

This section is used only for GEF evaluations to validate the project ratings identified during the initial terminal evaluations. In order to undertake the quality assessment of GEF terminal evaluations and to validate the rating of project implementation identified by the initial evaluator, additional documentation will be provided to quality assessment reviewers. This will include:

- The project concept note, and identification forms (PIF/Pdf A &B), and project document (ProDoc) including results framework.
- Project implementation reviews (APR/PIR).
- Tracking tools (as available).
- Midterm evaluation, if carried out.
- Project implementation action plan.

**GEF Evaluation Ratings Validation Table**

	UNDP IEO quality assessment rating		GEF terminal evaluation rating		Comments and/or justification for rating/ score adjustment	Suggestions for improvement
	Rating	Score	Rating	Score		

### Assessment of outcomes

Project focus	Indicate what the terminal evaluation has rated for project effectiveness, efficiency and relevance, and based on the available documentation, indicate and justify your rating. Provide your rating also in cases where the terminal evaluation has not included one.						
1	Effectiveness						
2	Efficiency						
3	Relevance						
4	Overall project outcome						

### Sustainability

Project focus	Indicate what the terminal evaluation has rated for sustainability and based on the available documentation indicate and justify your rating. Provide your rating also in cases where the terminal evaluation has not included one.						
5	Financial sustainability						
6	Socio-political sustainability						
7	Institutional framework and governance sustainability						
8	Environmental sustainability						
9	Overall likelihood of sustainability						

### Monitoring and evaluation

Project focus	Indicate what the terminal evaluation has rated for M&E quality and based on the available documentation indicate and justify your rating. Provide your rating also in cases where the terminal evaluation has not included one.						
10	M&E design at entry						
11	M&E plan and implementation						
12	Overall quality of M&E						

### Implementation and execution

Project focus	Indicate what the terminal evaluation has rated for the performance of UNDP as the project implementing agency and based on the available documentation indicate and justify your rating. Provide your rating also in cases where the terminal evaluation has not included one.						
13	Quality of UNDP implementation/ oversight						
14	Quality of Implementing partner execution						
15	Overall quality of implementation and execution						
<b>Overall project performance</b>							
Project focus	Does the terminal evaluation include a summary assessment and overall rating of the project results? Indicate the terminal evaluation rating and then indicate whether, based on the available documentation, you think a different rating of overall project results would be more appropriate.						
16	Provide justification for any agreement or adjustment to ratings.						

#### 6.10.4 General findings and lessons learned

Most evaluations should identify a number of lessons learned from project implementation. This section is not scored in the overall quality assessment but gives the reviewer an opportunity to identify the key lessons that could be drawn out of an evaluation and that should be shared more widely within a country office, regionally or globally.

##### Quality Assessment reviewer general evaluation findings

<b>1</b>	<p><b>Overall thoughts and lessons from the evaluation report and for future evaluations.</b></p> <ul style="list-style-type: none"> <li>• Detail positive and innovative aspects of the evaluation report</li> <li>• Lessons for other evaluators</li> <li>• Does the final score adequately reflect the quality of the evaluation?</li> <li>• What could have been done differently to strengthen the evaluation report (if not covered in the main assessment).</li> </ul> <p><i>Note: this area is free for the reviewer to give further thoughts and considerations of the report unrated. This should be constructive, for the organization to gather lessons learned in both project implementation and evaluation implementation.</i></p>
----------	---

LL 1
LL 2
LL 3

### Quality Assessment reviewer general programme/ project findings

2	<p><b>Overall thoughts and lessons from the evaluation report for future projects and programmes.</b></p> <ul style="list-style-type: none"> <li>• Detail positive and innovative aspects of the project or programme if any</li> <li>• Detail lessons for other projects or programmes</li> <li>• What could have been done differently to strengthen the project or programme?</li> <li>• What key project or programme lessons can be drawn from the report?</li> </ul> <p><i>Note: this area is free for the reviewer to give further thoughts and considerations of the report unrated. This should be constructive, for the organization to gather lessons learned in both project implementation as well as evaluation implementation.</i></p>
LL1	
LL2	

### 6.10.5 Summary quality assessment result

The overall quality assessment will automatically be summarized in the ERC and will be available for the reviewer to consider before submitting to IEO for approval and finalization.

QUALITY ASSESSMENT CRITERIA	Rating						Weighted score
	HS	S	MS	MU	U	HU	
<b>1. Evaluation structure and design</b>							
Do the TOR appropriately and clearly outline the purpose, objectives, criteria and key questions for the evaluation and allow adequate time and resources?							

## 2. Evaluation report and methodology

Are the evaluation report objectives, criteria, methodology and data sources fully described, and are they appropriate given the subject being evaluated and the reasons for carrying out the evaluation?

## 3. Cross-cutting and gender issues

Does the evaluation report address gender and other key cross-cutting issues?

## 4. Evaluation results, findings, conclusions and recommendations

Does the report clearly and concisely outline and support its findings, conclusion and recommendations?

### 6.11 UN-SWAP evaluation performance indicator and assessment

The United Nations System-wide Policy on Gender Equality and the Empowerment of Women (GEEW) was endorsed by the Chief Executive's Board for Coordination in October 2006 as a means of furthering the goal of gender equality and women's empowerment in the policies and programmes of the United Nations system. In 2012, the United Nations agreed on the System-wide Action Plan (UN-SWAP) on Gender Equality and the Empowerment of Women to implement the aforementioned gender policy. The UN-SWAP on Gender Equality and the Empowerment of Women assigned common performance standards for the gender-related work of all United Nations entities, including evaluation and includes an evaluation performance indicator (EPI). In 2020, the UN-SWAP was updated. The UN-SWAP EPI reporting follows the [UN-SWAP EPI Technical Notes](#) published by the United Nations Evaluation Group (UNEG).<sup>5</sup>

UNDP is required to report against the EPI annually, assessing both independent evaluations and decentralized evaluations. Detailed information on the EPI is available [here](#). This chapter summarizes key elements of the EPI and explains the UNDP assessment process.

#### 6.11.1 What is the UN-SWAP evaluation performance indicator?

<sup>5</sup> Access at: <http://www.unevaluation.org/document/detail/1452>

The EPI assesses the extent to which the evaluation reports of an entity meet the gender-related UNEG Norms and Standards for Evaluation and demonstrate effective use of the [UNEG Guidance on integrating human rights and gender equality](#) during all phases of the evaluation.

### 6.11.2 The evaluation performance indicator criteria and scorecard

A scorecard is used to assess evaluation reports against three criteria (a fourth criterion applies at the agency level). The first two criteria look at whether gender equality concerns were integrated in the evaluation scope of analysis and methods and tools for data collection and analysis.

1. Gender equality and the empowerment of women are integrated in the evaluation scope of analysis, and evaluation criteria and questions are designed in a way that ensures that relevant data will be collected.
2. Gender-responsive methodology, methods, tools and data analysis techniques are selected.

The third criterion is focused on whether the evaluation report reflects a gender analysis captured in various ways throughout the evaluation report.

3. The evaluation findings, conclusions and recommendations reflect a gender analysis.

The fourth criterion is focused on whether the entity – in the present case UNDP – has commissioned:

4. At least one evaluation to assess corporate performance on gender mainstreaming or equivalent every five to eight years.

Each evaluation report is assessed against the first three criteria using a four-point scale (0-3):

- **0 = Not at all integrated.** Applies when none of the elements under a criterion are met.
- **1 = Partially integrated.** Applies when some minimal elements have been met but further progress is needed, and remedial action is required to meet the standard.
- **2 = Satisfactorily integrated.** Applies when a satisfactory level has been reached and many of the elements have been met but improvement could still be made.
- **3 = Fully integrated.** Applies when all the elements under a criterion have been met, used and fully integrated in the evaluation and no remedial action is required.

The annex to the [UN-SWAP EPI technical note](#)<sup>6</sup> sets out guiding questions for the assessment against each criterion. After reviewing the individual evaluation report for each criterion, a score is assigned to the report as follows:

0-3 points = Misses requirement  
4-6 points = Approaches requirement  
7-9 points = Meets requirement

### 6.11.3 The assessment process

The UNDP IEO is the focal point for the EPI. Before 2020, the IEO contracts an external expert to conduct the assessment of a set of evaluations, including all independent evaluations and a sample of

<sup>6</sup> Access at: <http://www.unevaluation.org/document/detail/1452>

decentralized evaluations that were finalized in the period being reported (January-December of each year).

Since the SWAP EPI was integrated into the IEO online quality assessment system in 2020, all quality assessed decentralized evaluations have also been assessed against it by reviewers engaged by IEO. A reviewer also assesses SWAP EPI for all independent evaluations. Scores for all evaluations, independent and decentralized, are aggregated into a final score for UNDP as a whole. In 2020, the UNDP aggregate score was “exceeds requirements” for the first time.

#### Box 1: Sample evaluations that have met EPI requirements

- Bangladesh, 2020, [Final Evaluation of Partnership for a Tolerant, Inclusive Bangladesh \(PTIB\) project](#)
- Nepal, 2020, [Final Evaluation of Resilient Reconstruction and Recovery of Vulnerable Communities Severely Affected by 2015 Earthquake \(EU II\)](#)
- Haiti, 2020, [Évaluation finale du projet Promotion de la Cohésion sociale à Jérémie](#)
- South Sudan, 2020, [Final evaluation Peace and Community Cohesion project](#)

*Note: the EPI assesses the extent to which the **evaluation** integrates gender equality and women’s empowerment. A report may score well against the EPI even if the findings of the evaluation as to the integration of gender in the programme/ project being evaluated are negative.*

#### 6.11.4 Reporting

The IEO prepares a final synthesis report, which is uploaded to the UN-SWAP on Gender Equality and the Empowerment of Women reporting portal. UN-Women analyses all UN-SWAP performance indicators, including for evaluation, and an aggregated report is presented every year through the report of the Secretary-General to the Economic and Social Council on mainstreaming a gender perspective into all policies and programmes in the United Nations system.<sup>7</sup>

---

<sup>7</sup> The 2020 report can be accessed at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N20/092/53/pdf/N2009253.pdf?OpenElement>