SECTION 6

EVALUATION QUALITY ASSESSMENT

CONTENTS

6. QUALITY ASSESSMENT OF DECENTRALIZED EVALUATIONS[1]

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.[2] 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.

---

[1] Section 6 was comprehensively updated in January 2024 to reflect changes to the quality assessment process which would come into effect for all evaluations implemented in 2024 and beyond. Any evaluations implemented before this will be assessed in line with the 2021 quality assessment guidelines.

[2] UNDP Evaluation Policy, http://web.undp.org/evaluation/policy.shtml

- Supporting bureau oversight functions by providing concurrent feedback through detailed analysis of the quality of evaluation reports, with recommendations for their improvement.
- 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.[3]

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)** [4]
   - 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.[5]

---

[3] 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
[4] Access at: http://erc.undp.org
[5] 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.
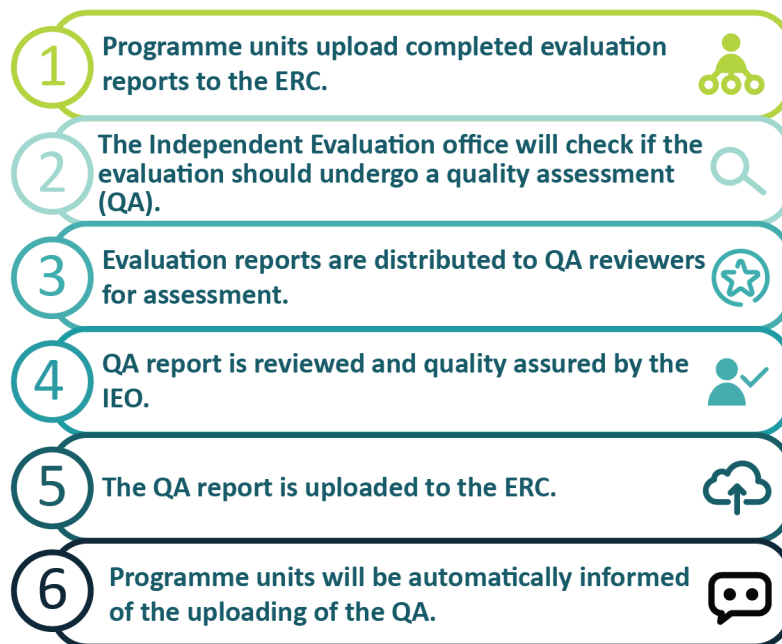
3. **Quality assessment**
   o The IEO sends the evaluation report to a contracted quality assessment reviewer to conduct a quality review.
   o The quality assessment rating is made available on the ERC typically within two weeks of completion and submission of the quality assessment report.
4. **Feedback**
   o Upon receiving the quality assessment report from the reviewer, the IEO reviews the report and then makes it available to the respective programme unit though the ERC.

Figure 1. Quality assessment process

1	Programme units upload completed evaluation reports to the ERC.
2	The Independent Evaluation office will check if the evaluation should undergo a quality assessment (QA).
3	Evaluation reports are distributed to QA reviewers for assessment.
4	QA report is reviewed and quality assured by the IEO.
5	The QA report is uploaded to the ERC.
6	Programme units will be automatically informed of the uploading of the QA.

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

The Quality Assessment process is undertaken at the end of each quarter, April for Q1, July for Q2, October for Q3, and January for Q4. In high completion periods, normally November through to January, the quality assessment process is undertaken more frequently. A quality assessment report for an individual evaluation is 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 and design:** Five questions weighted 15 percent.
  - Do the terms of reference appropriately and clearly outline the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?
- **Report structure, methodology and data sources:** 8 questions weighted 30 percent.
  - Does the evaluation report appropriately and clearly address the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?
- **Cross-cutting issues:** Five questions weighted 15 percent.
  - Does the evaluation report address LNOB and other cross cutting issues where relevant?
- **Findings, conclusions and recommendations:** Eleven questions weighted 40 percent.
  - Does the report clearly and concisely outline and support its findings, conclusions and recommendations?
- **Management response:** two questions not included in the overall QA scoring.
  - Is the management response clear, realistic and actionable?

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.

Figure 3. Quality assessment rating scale

Code	Rubric for assigning rating		Value
HS	**Highly satisfactory**	All parameters were fully met and there were no shortcomings in the evaluation report	6
S	**Satisfactory**	All parameters were fully met with minor shortcomings in the evaluation report	5
MS	**Mostly satisfactory**	The parameters were partially met with some shortcomings in the evaluation report	4
MU	**Mostly unsatisfactory**	More than one parameter was unmet with significant shortcomings in the evaluation report	3
U	**Unsatisfactory**	Most parameters were not met and there were major shortcomings in the evaluation report	2
HU	**Highly unsatisfactory**	None of the parameters were met and there were severe shortcomings in the evaluation report	1
N/A	**Not Applicable**	Not Applicable	unscored

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 satisfactory, if all UNDP evaluation requirements are met, an evaluation should receive at a minimum rating of 5 (satisfactory), which is the benchmark for a good evaluation.

6.8   Quality assessment tool

The quality assessment tool is accessible from the ERC website (http://erc.undp.org). 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, 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 Terms of Reference and Design

Do the terms of reference appropriately and clearly outline the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines? *(Section 4.3.2)*	
1.1	**Does the TOR clearly outline the purpose, scope, and objectives of the evaluation in a logical and realistic manner?** *(Section 4, page 14-15-16 and Annex 1, page 39-40)* ▪ Does it follow the proposed structure detailed in the UNDP evaluation guidelines? ▪ Does it provide details on evaluation context? ▪ Does it include information regarding the results framework in the main text or annexes?
1.2	**Does the TOR clearly detail timescales and allocation of days for the evaluation?** *(Section 4, page 15, Annex 1, page 49)* ▪ Is the allocation of days detailed for each step of the evaluation process and adequate given the scope of the evaluation? ▪ Does it specify details (including the composition) of the evaluation team and roles and responsibilities of individual team members?
1.3	**Does the TOR outline the involvement of evaluation partners, stakeholders and feedback mechanism?** (S*ection 4, page 15, Annex 1, page 47)* ▪ Does the ToR identify which partners will be involved in the evaluation and how? ▪ Is the audit trail-feedback process, corresponding/reporting parties, and timeline outlined?
1.4	**Is there an outline of the evaluation approach and methodology in the TOR?** *(Section 4, page 14-15 and Annex 1, page 45)* ▪ Are the number of key evaluation questions appropriate given the scope of the evaluation? ▪ Is there guidance for a general methodological approach? ▪ Are available data sources as well as data requirements of the evaluation explained?
1.5	**Does the TOR include a requirement for a "Leave no One Behind", gender responsive, and disability inclusive evaluation?** *(Section 4, page 16-17, and page 20, Annex 1, page 44)* ▪ Does the ToR detail LNOB, gender, and disability inclusive specific questions, supplementing or integrated under the key evaluation questions? ▪ Does the TOR require the evaluation to explain how the tools, methodologies, and data analysis adopted will capture the LNOB, gender and other cross cutting issues?

6.10.2 Report structure, methodology and data sources

	Does the evaluation report appropriately and clearly address the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?	
STRUCTURE		
2.1	**Is the evaluation report well balanced and structured?** ▪ Does it follow the proposed evaluation report structure detailed in the UNDP Evaluation guidelines (section 4, page 29-30 and Annex 4) ▪ Does it cover all evaluation criteria as requested in the ToR? ▪ Is the length of the main report reasonable for the requirements of the evaluation? ▪ Does the report include all annexes/appendices? (TOR for the evaluation, Evaluation matrix and data collection instruments, List of individuals or groups interviewed or consulted, and sites visited, List of supporting documents reviewed)	
2.2	**Does the report have a high-quality Executive summary that captures the main elements of the report? (Section 4, Annex 4, page 57)** ▪ Does it include a brief overview of the purpose and objective- why the evaluation is being undertaken? ▪ Does it include a summary of the evaluation scope and main areas of inquiry - what is being evaluated? ▪ Does it include a summary of key evaluation findings? ▪ Does it include a summary of concluding statements and proposed recommendations?	
2.3	**Is the nature and extent of partners' and stakeholder's role and involvement in project implementation explained adequately?**	
METHODOLOGY		
2.4	**Is there an examination of intervention`s ToC/ results pathway or an attempt to reconstruct the ToC/ results pathway?**	
2.5	**Is the evaluation's methodological approach clearly outlined?** (Section 4 Annex 4, page 58-59) ▪ Did the evaluator explain the general evaluation approach/framework used to answer Evaluation Questions? (i.e., if impact/attribution is to be uncovered, impact assessment methodology; if contribution is to be uncovered other frameworks such as outcome harvesting, contribution analysis, process tracing etc...) ▪ Did the evaluator make a distinction between Methodology adopted for the evaluation and data collection tools (which is covered under question 5)? ▪ Are any changes from the approach outlined in the TOR detailed?	
DATA COLLECTION and ANALYSIS		
2.6	**Is the data collection and analysis strategy clearly defined?** *(Section 4 Annex 4, page 59)* ▪ Are data sources being clearly detailed? ▪ Are data analysis approaches (including triangulation methods) explained? ▪ Are data collection methods (including sampling strategy) detailed?	

	Note: The evaluator should explain if he/she relied on primary data – their collection or secondary data (the sources from the project documents) whether he/she did surveys, case studies, interviews, focus groups, group discussions, employed geospatial data collection methods etc.
2.7	**Is the data collection and analysis adequate and sufficiently diverse for the scope of the evaluation?** ▪ Are a comprehensive set of data sources included where appropriate? ▪ Is there a clear presentation of data analysis and references to data sources within the report? ▪ Did the evaluator explain how project monitoring data was used and verified?
2.8	**Are limitations in implementation of the evaluation clearly explained as well as mitigation measures?** ▪ Are constraints in access to data (including interviewees and quantitative data sources) explained? ▪ Are explanations provided as to how constraints in access to data were addressed?

6.10.3 Cross-cutting issues

	Does the evaluation report address LNOB and other cross cutting issues where relevant?

3.1	**If requested in the ToR, does the evaluation adequately include and analyse cross cutting issues such as:**
	  ■ Leave no one behind (LNOB)
	https://erc.undp.org/methods-center/methods/assessing-crossing-cutting-themes/assessing-leaving-no-one-behind
	  ■ Social and Environmental standards (this is compulsory for GEF evaluations)
	https://erc.undp.org/methods-center/methods/assessing-crossing-cutting-themes/ses

3.2	**Are gender equality and empowerment of women (GEEW) integrated in the evaluation scope of analysis and evaluation criteria and questions designed in a way that ensures GEEW related data will be collected?**
	  ■ Does the evaluation include an objective specifically addressing gender equality and/or human rights issues and/or was gender mainstreamed in other objectives?
	  ■ Was a stand-alone evaluation criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?
	  ■ Were one or several dedicated gender equality and empowerment of women evaluation questions integrated into the evaluation?
	
	Section 6.10.3, page 11: https://erc.undp.org/pdf/UNDP_Evaluation_Guidelines.pdf
	UN SWAP Eval Performance Indicator Scorecard:
	http://www.unevaluation.org/document/download/2149

3.3	**Were gender-responsive methodologies, methods and tools, and data analysis techniques selected?**
	  ■ Does the evaluation specify 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?
	  ■ Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating gender equality and empowerment of women considerations?
	  ■ Were diverse range of data sources and processes employed (i.e., triangulation, validation) to guarantee inclusion, accuracy and credibility?
	  ■ Do the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?
	
	Section 6.10.3, page 11: https://erc.undp.org/pdf/UNDP_Evaluation_Guidelines.pdf
	UN SWAP Eval Performance Indicator Scorecard:
	http://www.unevaluation.org/document/download/2149

3.4	**Do the evaluation findings, conclusions and recommendation reflect a gender analysis?**
	  ■ Does the evaluation have 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?
	  ■ Do 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?

	- Are unanticipated effects of the intervention on gender equality and human rights are described? - Does the evaluation report provide 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? *Section 6.10.3, page 11:* https://erc.undp.org/pdf/UNDP_Evaluation_Guidelines.pdf UN SWAP Eval Performance Indicator Scorecard: http://www.unevaluation.org/document/download/2149
3.5	**Does the evaluation include considerations of disability inclusion in the intervention?** - Do the evaluation questions cover different aspects of disability inclusion? - Do evaluation findings and analysis provide data and evidence on disability inclusion? - Do the conclusions and/ or recommendations of the evaluation reflect their findings on disability inclusion? **Note,** *this is a new requirement from 2019. Only review this IF the TOR requests inclusion or the report includes an analysis of disability inclusion issues. If there is no request in the TOR mark as N/A* *Section 6.10.3, page 11:* https://erc.undp.org/pdf/UNDP_Evaluation_Guidelines.pdf The UN disability inclusion strategy requires the inclusion of an analysis of the inclusion of disability in evaluations. The strategy and guidance can be found here https://www.un.org/en/content/disabilitystrategy/

6.10.4 Findings, conclusions and recommendations

Does the report clearly and concisely outline and support its findings, conclusions and recommendations?	
FINDINGS AND CONCLUSIONS	
4.1	**Does the evaluation report contain a concise and logically articulated set of findings?** *(Section 4, 4.4.5, p.30; Annex 4, Evaluation template and quality standards, pp.59-60; Section 6, 6.10.4, p.12)* ▪ Do the findings address the evaluation criteria and evaluation questions? ▪ Are the findings detailed and supported by evidence?
4.2	**Does the Evaluation assess the RELEVANCE of the intervention, if required by the TOR?** *(As a separate Relevance section or as a discussion within the findings and conclusions)* *Including, but not limited to:* ▪ Government priorities and strategies ▪ UNDP Country Programme Document outcomes and outputs
4.3	**Does the evaluation assess the COHERENCE of the intervention, if required by the TOR?** *(As a separate Coherence section or as a discussion within the findings and conclusions)* *Including, but not limited to:* ▪ *Internal coherence: the synergies and interlinkages between the intervention and other interventions carried out by UNDP.* ▪ *External coherence: the consistency of the intervention with other actors' interventions in the same context. (Complementarity, harmonization and co-ordination with others, and the extent to which the intervention is adding value while avoiding duplication of effort).*
4.4	**Does the Evaluation assess the EFFECTIVENESS of the intervention, if required by the TOR?** *(As a separate Effectiveness section or as a discussion within the findings and conclusions)*
4.5	**Does the Evaluation clearly assess the EFFICIENCY of the intervention, if required by the TOR?** *(As a separate Efficiency section or as a discussion within the findings and conclusions)* *Including but not limited to:* ▪ How well did the intervention use and allocate its human and financial resources? ▪ Are variances between planned and actual expenditures across results/outputs assessed and explained?
4.6	**Does the Evaluation clearly assess SUSTAINABILITY of the intervention, if required by the TOR?** *(As a separate Sustainability section or as a discussion within the findings and conclusions)*
4.7	**Does the evaluation include an assessment of the results framework, its reporting and overall quality?** *(Section 6, p.10)* ▪ Is there an assessment of the quality of results framework, its indicators and level of disaggregation by sex and other vulnerable groups? ▪ Are all indicators in the Results Framework assessed and analysed by the evaluator with final achievements noted?
4.8	**Does the evaluation include an assessment of the monitoring system of the intervention and its overall quality?** *(Section 6, p.10)*
4.9	▪ Is there a discussion and analysis of risk management and the quality of adaption/mitigation measures in the evaluation report including Social and Environmental Safeguards?

4.10	**Does the evaluation report contain a concise and logically articulated set of conclusions which are standalone in nature?** *(Section 4, 4.4.5, p.30; Annex 4, UNDP evaluation report template and quality standards, p.60; Section 6, 6.10.4, p.12)* ▪ Do conclusions build on the findings? ▪ Are they clear, concise and well substantiated? ▪ Do they address the objectives of the evaluation?

RECOMMENDATIONS	
4.11	**Are the evaluation recommendations clear, concise, targeted (at UNDP and/or implementing partners), realistic and actionable given the size and scope of the intervention?** (Section 4, 4.4.5, p.30; Annex 4, UNDP evaluation report template and quality standards, pp.60; Section 6, 6.10.4, p.12)

6.10.5 Management Response

MANAGEMENT RESPONSE **Is the management response clear, realistic and actionable?**	
5.1	**Is the management response adequately formulated and does it respond to the spirit of the recommendations (Where accepted)? Where rejected is the reason for the rejection clearly articulated?** *(Section 4, Step 4.5, page 34)* *Only to be done if available at the time of the quality assessment*
5.2	**Are the key actions SMART (specific, measurable, actionable, relevant, time bound) (Where accepted or partially accepted)?** *(Section 4, Step 4.5, page 35)* *Only to be done if available at the time of the quality assessment*

6.10.6 General findings and lessons learned.

Most evaluations should identify several 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.

GENERAL FINDINGS AND LESSONS LEARNED	
To be completed in narrative format and not graded. This area is for the reviewer to give further thoughts and considerations of the report.	
6.1	**Lessons for Evaluations and Evaluators:** Overall thoughts on the quality of the evaluation report

6.2	Were there any innovative aspects of the evaluation report that you can highlight? Yes/No
	▪ If it is a yes, please explain in the comments.

6.10.7 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							

Overall project performance

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. | | | | | | |

6.10.8 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	
1. Evaluation structure and design							
Do the terms of reference appropriately and clearly outline the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?							
2. Evaluation report and methodology							
Does the evaluation report appropriately and clearly address the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?							
3. Cross-cutting and gender issues							
Does the evaluation report address LNOB and other cross cutting issues where relevant?							
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).[6]

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?

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.

---

[6] Access at: http://www.unevaluation.org/document/detail/1452

- **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 [7] 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 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)
- Haïti, 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

---

[7] Access at: http://www.unevaluation.org/document/detail/1452

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.[8]

---

[8] The 2020 report can be accessed at https://documents-dds-ny.un.org/doc/UNDOC/GEN/N20/092/53/pdf/N2009253.pdf?OpenElement