Independent Evaluation Office
United Nations Development Programme

UNDP EVALUATION GUIDELINES

Revised 2026

UNDP EVALUATION GUIDELINES ©UNDP Independent Evaluation Office, September 2026

Permission to reproduce any part of this publication is required. Please contact: ieo.cd@undp.org.

Suggested citation: UNDP Evaluation Guidelines, UNDP Independent Evaluation Office,
New York, September 2026.

Contents

Preface from the Director **04**

Note from the Administrator **05**

Foundations **06**

Step 1. **Plan** **17**

Step 2. **Prepare** **27**

Step 3. **Inception** **39**

Step 4. **Data collection** **46**

Step 5. **Data analysis and reporting** **52**

Step 6. **Finalize and respond** **61**

Step 7. **Quality assessment** **69**

Preface from the Director

Evaluation is a cornerstone of UNDP's results-based management system. It reflects the organization's commitment to evidence-informed decision-making, to transparency, and to learning from both successes and challenges.

Evaluation asks us to be open to critical inquiry – and that openness is precisely what makes it such a valuable investment. Evaluation provides information, helping us understand the difference UNDP's work is making, how change happens, and where to adjust course. It generates insights, allowing us to learn from experience and continually improve the way UNDP works towards its development goals. And it creates a space to examine data in context, so that complexity is better understood and differences in perspective are worked through, rather than left unspoken.

For evaluation insights to inform decisions and action, people must trust the evidence behind them. That is why these revised UNDP Evaluation Guidelines place quality at the centre of the entire evaluation process, from the earliest stages of planning to the communication of evaluation findings and lessons learned.

The Guidelines also introduce important innovations designed to strengthen evaluation quality and use across UNDP. Among them are the Artificial Intelligence for Development Analytics (AIDA) platform, which opens access to insights from more than 7,000 UNDP evaluation resources, and the AI Quality Assessment Tool, which supports more consistent and systematic review of evaluation reports. These innovations complement professional judgement, enhance learning, and make evaluation resources more accessible and actionable for evaluators and users alike - within UNDP and with our partners around the world.

I encourage everyone involved in evaluation - as commissioners, managers, evaluators, partners, or users of evidence - to make full use of these Guidelines. Together, we can ensure that evaluation continues to serve learning, accountability, and better development outcomes.

Isabelle Mercier

Director, Independent Evaluation Office

United Nations Development Programme

Note from the Administrator

Development only matters if it changes lives.

At the United Nations Development Programme (UNDP), that principle guides every project, policy and partnership we undertake. But achieving meaningful results across more than 170 countries - amid poverty, conflict and climate disruption - requires a constant commitment to learning.

Six decades of global experience give UNDP a unique advantage. Yet experience creates value only when we examine it critically, learn from it and apply those lessons to future decisions. What works in one community may not work in another. Understanding that difference is essential to delivering effective development outcomes.

That is why evaluation matters. Evaluation turns experience into evidence, evidence into insight, and insight into better decisions. It helps us understand what worked, what did not, and what we should do differently next time. Our commitment to listen, partner and deliver depends on this discipline.

This discipline matters more than ever as the development landscape evolves. Resources are tightening and risks are becoming more complex. Evaluation must evolve with them. It must help us learn faster, adapt more quickly and direct scarce resources where they have the greatest impact. By feeding evidence back into planning and delivery, evaluation strengthens accountability, reinforces a culture of learning and helps build the trust that underpins effective development.

These revised Evaluation Guidelines provide a practical framework for planning, conducting and using evaluations that are credible, useful and focused on improving performance. They reflect UNDP's commitment to evidence-based decision-making and to a culture of continuous learning. I thank everyone who contributed to their preparation.

Evaluation is not an end in itself. It is a means to better decisions, better programmes and better results for the people we serve.

Alexander De Croo

Administrator, United Nations Development Programme

Foundations

Why have these guidelines been revised?

UNDP has developed these revised Evaluation Guidelines to align with the UNDP Evaluation Policy for 2025-2030. These guidelines translate evaluation policy requirements into practical, actionable guidance so that UNDP has access to ready-to-use evaluations that are credible, ethical, and useful for learning, accountability, and decision-making.

These guidelines were endorsed by the UNDP Organizational Performance Group in April 2026.

Who are these guidelines for?

The revised UNDP Evaluation Guidelines are designed primarily for the implementation of decentralized evaluations. However, all those involved in planning, managing, conducting, and using evaluations across UNDP and its associated entities will find these guidelines helpful. They aim to promote a shared understanding of the evaluation principles, standards and processes, and responsibilities that support consistent and high-quality evaluation practice across the organization.

Evaluation touches all areas of UNDP. Moreover, it supports the foundations for sound decisions at the policy, programming, and project levels. Each group listed below has a role in the use, promotion, and carrying out of evaluation:

Audience group	Role in evaluation
UNDP Country Offices	Commission evaluations to meet information needs and accountability requirements.
UNDP Regional Offices	Ensure quality and alignment of evaluations with evaluation plans and commission regionally relevant evaluations.
UNDP Corporate Offices (principally, the Bureau of Policy and Programme Support)	Ensure compliance with the UNDP Evaluation Policy and accountability to the Executive Board on behalf of the UNDP Administrator. Establish mandatory requirements for decentralized evaluations and ensure that evaluations remain a key accountability mechanism within the organization through the monitoring of key performance indicators (KPIs) for evaluation. Commission decentralized evaluations, as needed by the organization.
UNDP Hosted Agencies including UN Capital Development Fund, UN Volunteers, and the UN Office of South-South Cooperation	Commission decentralized evaluations to make informed decisions, and establish costed evaluation plans.
Independent Evaluation Office	Provide UNDP with independent, credible, and useful evidence and technical evaluation advice to make informed decisions.
UNDP Executive Board	Review and provide guidance for multi-year evaluation workplans and approve corporate and policy-level evaluations and their associated management responses and action plans.

For all involved, evaluation is an essential tool to ensure transparency and accountability within the organization. Evaluation helps promote shared learning for better development outcomes when used both within and outside of UNDP.

How are they structured?

The UNDP Evaluation Guidelines are designed as a step-by-step companion for planning, managing, conducting, and using evaluations. They follow the life cycle of an evaluation, structured into the following seven main sections that mirror the evaluation process:

1. **Plan:** Develop the annual country-level evaluation plan, managing it against programme cycles, information needs, resource allocation, and organizational requirements.

2. **Prepare:** Design the evaluation, determining evaluability and terms of reference (TOR), as well as identifying evaluators and feedback systems throughout the evaluation.

3. **Inception:** Kickstart the evaluation, introducing it and the team, creating (and agreeing on) the inception report to plan the work, and identifying how evaluation questions will be answered and what approach will be used to pull the information together into a cohesive report.

4. **Data collection:** Gather the evidence base for the evaluation, collecting data across primary and secondary sources including field mission set up, documentation review, and storing and protecting data to ensure that privacy and ethical standards are met.

5. **Data analysis and reporting:** Interpret and synthesize data, including reporting against findings, assessing the quality of the report's structure, analysis, assessments, and recommendations, safeguarding the independence of the assessment, and mitigating disputes related to the report.

6. **Finalize and respond:** Formally endorse evaluation recommendations and commitment to follow-up actions including preparation of management responses, publication of final reports, and tracking implementation of agreed key actions and commitments.

7. **Quality assessment:** UNDP quality assessment of decentralized evaluations, including quality rating of evaluation reports and verifying compliance with the established criteria, processes, and tools.

Policy, purpose and principles



UNDP's evaluation function is anchored in the UNDP Evaluation Policy for 2025-2030. This policy establishes the purpose, principles, and institutional architecture for evaluation across UNDP and its associated entities, namely the United Nations Capital Development Fund (UNCDF), United Nations Volunteers (UNV), and the United Nations Office for South-South Cooperation (UNOSSC).

> **"The evaluation function supports a results-driven, learning-oriented, and accountable organization, where all evaluations are aligned with the UNDP Results-Based Management (RBM) and Accountability Framework."**
>
> ***- UNDP Evaluation Policy for 2025-2030, page 2.***

The Policy defines the framework within which all evaluations are planned, implemented, and used, and underpins these Evaluation Guidelines.

Purpose of evaluation

Evaluation in UNDP is defined as an assessment, as systematic and impartial as possible, of an activity, project, programme, portfolio, strategy, policy, topic, theme, sector, operational area, or institutional performance, in line with UNEG Norms and Standards (2016).

Evaluations analyse both expected and achieved accomplishments, the causal pathways leading to them, and contextual factors that influence the achievement of results. They can help to determine the relevance, coherence, efficiency, effectiveness, impact, and sustainability of UNDP's work, and support evidence-based adjustments to improve its organizational and system-wide contributions to development.

Evaluation at UNDP serves multiple, interrelated purposes, as follows:

- **Accountability:** Identify the extent to which decisions matched stated planned objectives, in line with how and why programming and/or policy decisions were made.
- **Learning and use:** Clarify what works and what does not (and the reasons why), and when they can be timed most effectively. Moreover, learning from evaluation supports change and adaptation.
- **Evidence-based and strategic decision-making:** Evaluation fills knowledge gaps, guides course corrections, and enhances the quality and effectiveness of UNDP programming.

Evaluations are distinct from other forms of oversight at UNDP - including reviews, appraisals, audits, inspection and investigation, and monitoring - because they apply explicit evaluative judgement against agreed criteria.

Furthermore, evaluations triangulate multiple lines of evidence, assess causal pathways, and generate conclusions and recommendations that require formal management responses and follow-up actions. This makes evaluation a unique mechanism through which to generate systematic, credible evidence to inform decision-making, ensure accountability, and foster organizational learning.

Type	Definition
Appraisal	A critical assessment of the potential value of a proposed undertaking before a decision is made to implement it.
Audit	An independent assessment of the adequacy of management controls in order to ensure: • Economical and efficient use of resources; • Safeguarding of assets; • Reliability of financial and other information; • Compliance with regulations, rules, and established policies; • Effectiveness of risk management; and • Adequacy of organizational structures, systems, and processes.
Investigation	An impartial examination of evidence to determine wrong-doing and/or the violation of UN rules, regulations, or standards.
Monitoring	The continuous collection and review of information to track progress achieved against planned targets during the implementation of activities.
Research	An objective and systematic examination of a topic designed to develop or contribute to knowledge.
Results-based management (RBM)	A management strategy that focuses on the achievement of results across planning, implementation, and evaluation phases of a project or programme. It entails the systematic use of results frameworks, indicators, baselines, targets, monitoring, learning, and reporting.
Review	The periodic or ad hoc (often rapid) assessment of performance that does not apply the due process of evaluation. Reviews tend to focus primarily on operational issues.

Guiding principles

All UNDP evaluations are conducted in accordance with the principles articulated in the 2025-2030 Policy and UNEG Norms and Standards, as outlined below:

- High ethical standards

Ethical evaluation practice in UNDP is guided by the UNEG Ethical Guidelines for Evaluation (2020), which help evaluators to do their work responsibly and fairly. These guidelines set clear expectations to be honest and independent, to treat people with respect, and to handle information carefully. In practice, this means seeking informed consent, protecting confidentiality, and using data in a responsible and transparent way. Evaluators are also expected to be sensitive to context

and culture, to avoid causing harm, and to clearly explain how evaluation findings are produced and used. Together, these principles help protect the rights and dignity of everyone involved in an evaluation and build confidence in the credibility of evaluation results.

Artificial intelligence (AI) is increasingly shaping how information is collected, analysed, and used in evaluations. This makes it especially important to think carefully about ethics when using AI, particularly with regard to how new technologies affect people's rights, dignity, and privacy. UNDP follows the *UNEG Ethical Principles for Harnessing AI in United Nations Evaluations (2025)* which provide guidance on using AI responsibly, being transparent about its use, and protecting confidential information. Safeguarding privacy is therefore a fundamental ethical responsibility, rather than a mere technical step. It helps ensure that AI is used in ways that adhere to UNEG principles, maintain trust, and protect everyone involved in the evaluation process.

- Independence, credibility, and impartiality

Independence must be safeguarded so that evaluations remain free from undue influence while also integrating stakeholder consultation to enhance credibility and use. The UNDP Evaluation Policy for 2025-2030 clearly outlines the responsibility of UNDP management to take "all necessary actions to ensure the objectivity and impartiality of the process and persons hired." (page 10)

- Integration of cross-cutting dimensions

All UNDP evaluations assess the gender equality, human rights, disability inclusion, climate and environmental sustainability, resilience, South-South cooperation, and "leave no one behind" (LNOB) principles.

- Rule-bound planning and implementation

All evaluations must follow UNEG Norms and Standards, apply transparent and consultative processes, ensure quality assurance and independent review, align with approved results frameworks, and safeguard impartiality by avoiding conflicts of interest.

- Rigor and technical competence

Evaluations should employ robust designs, clear theories of change, and appropriate qualitative and quantitative methods, leveraging innovation and new technologies where relevant.

- Transparency and stakeholder engagement

Evaluations are participatory and publicly accessible through the Evaluation Resource Centre (ERC).

Complementary normative references

The Evaluation Guidelines are also informed by the broader normative and policy framework that shapes evaluation practice across UNDP, including:

- UNEG Norms and Standards for Evaluation in the UN System (2016);
- UNEG Ethical Principles for Harnessing AI in United Nations Evaluations (2025);

- UNEG Ethical Guidelines for Evaluation (2020);
- UNEG Guidance on Integrating Human Rights and Gender Equality in Evaluations (2024);
- OECD/DAC Evaluation Criteria (2019);
- UNDP Social and Environmental Standards (2021);
- UNDP Gender Equality Strategy (2022–2025); and
- UN Disability Inclusion Strategy (2019).

UNDP evaluation resources and tools

Evaluation in UNDP is supported by a comprehensive set of resources and tools designed to make evaluations consistent, credible, and user-friendly. Together, these form an integrated evaluation support package that enables UNDP staff to plan, manage, conduct, and use evaluations effectively across all contexts.

At the core of this package are the following key components:

The UNDP ERC: This is the main platform for finding information about evaluation in UNDP. It is an organizational online repository where all final evaluation reports, management responses, and key actions are uploaded and made publicly accessible. The internal ERC also hosts evaluation TORs and quality assessment results. In addition, it houses the following key tools:

- **UNDP Evaluation Guidelines:** The central reference for UNDP decentralized evaluations, which clarifies the mandatory requirements for evaluation across the organization. The guidelines also offer step-by-step guidance through every evaluation phase - plan, prepare, inception, data collection, data analysis and reporting, finalize and respond, and quality assessment - helping users navigate the process from start to finish. The ERC also hosts a series of guidelines on evaluative methods, including on portfolios and impact evaluation.
- **Templates, quality checklists, flowcharts, and other tools:** Practical, ready-to-use templates for evaluation plans, TORs, inception reports and evaluation matrices, evaluation reports, and management responses.
- **Evaluation Quality Assessment System:** An independent mechanism managed by the UNDP Independent Evaluation Office (IEO) to quality review all decentralized evaluation reports. Using a standardized rating framework, the system provides feedback through the ERC, supports reporting to the UNDP Executive Board, and promotes continuous improvement in evaluation quality, accountability, and learning across UNDP.
- **The UNDP Evaluation Methodology Center:** A dynamic online hub offering methodological guidance, toolkits, and examples of evaluation approaches, methods, and techniques. It provides how-to guides, notes, and links to emerging innovations.

- **Artificial Intelligence for Development Analytics (AIDA):** An AI-powered platform developed by the IEO to provide easy access to evaluative evidence from thousands of UNDP evaluation reports housed in the ERC. It is an intelligent search and knowledge extraction system that helps users mine insights from UNDP's vast evaluation archive.

- **UNDP Policies and Procedures Portal (POPP):** The key reference for policies, which provides the operational standards and procedural guidance on UNDP's core business processes.

Types of evaluations undertaken by UNDP

In line with the UNDP Evaluation Policy for 2025-2030, evaluations fall into two categories:

- Independent evaluations, carried out by the IEO; and

- Decentralized evaluations, commissioned by UNDP country and regional offices and carried out by third-party evaluators to maintain their independence.

Under each category, there are different types of evaluations. Certain types of evaluations can fall under both categories.

Independent evaluations

Independent evaluations are commissioned and managed by the IEO, which is a UNDP division not subject to the organization's normal accountability hierarchies to ensure independence, credibility, and quality. These evaluations are planned through extensive consultation with UNDP's senior management and approved by the UNDP Executive Board to inform global, regional, and country-level programming.

The IEO carries out the following types of evaluations:

Independent Country Programme Evaluations (ICPEs): Assessments of UNDP's performance and contributions at the country level. They cover one or more country programme cycles and inform new Country Programme Documents (CPDs) and UN Sustainable Development Country Frameworks (UNSDCFs). ICPEs accompany new CPDs when being considered for approval by the UNDP Executive Board.

Thematic evaluations: Analyses of UNDP's global performance on thematic areas or organizational strategies. They draw lessons across countries and regions, and provide evidence to inform strategic decision-making at the highest level.

Evidence syntheses: Aggregations of findings from multiple evaluations into meta-syntheses to distill lessons, identify trends and gaps, and assess the quality of UNDP's evaluation system.

All IEO evaluations can be found at the ERC.

Decentralized evaluations

Decentralized evaluations are commissioned and managed by UNDP country and regional offices. They meet the policy and programming information needs of country offices, regional bureaux, and headquarters divisions for accountability and learning. In alignment with the Evaluation Policy, they are carried out by external evaluators and managed by staff outside of the programme/project to ensure independence.

> **”UNDP business units (programme and policy units) commission decentralized evaluations according to evaluation plans that coincide with relevant programmes (central, regional, and country) and global projects. These evaluations are to be carried out by independent external consultants, and UNDP management shall take all necessary actions to ensure the objectivity and impartiality of the process and persons hired.”**
>
> ***- UNDP Evaluation Policy, 2025-2030, Article 64, page 10***

They are not presented to the Executive Board for approval.

Types of decentralized evaluations include:

Project evaluations: Assessing the performance of projects. They are mandatory for projects above defined thresholds as discussed in the “Plan” section of these guidelines. They include evaluations of Global Environment Facility (GEF), Green Climate Fund (GCF), and Adaptation Fund (AF) projects.

Strategic evaluations: An umbrella term used in practice to refer to evaluations of clusters of projects, which can include country programme, outcome, thematic, and portfolio evaluations. These assessments look at the big picture - examining how well a strategy, country programme, theme, outcome area, or portfolio is working overall, rather than evaluating a single project. They focus on system-level learning, coherence, and long-term results.

Portfolio evaluations: Assessing the performance, coherence, and system-level contributions of a UNDP portfolio, which is an integrated set of interventions designed to address complex development challenges. According to the UNDP Portfolio Policy and Portfolio Evaluation Guidance, portfolio evaluations examine how a portfolio’s theory of change, results framework, and adaptive management approach contribute to shifts in systems, portfolio momentum, and learning. They provide evidence for strategic decision-making at the CPD (or Regional Programme Document (RPD)) or strategic plan level, and are especially relevant where UNDP is applying a systems or portfolio approach.

Outcome evaluations: Assessing UNDP's contribution to one or more country programme outcomes. These evaluations assess a country programme's efforts to achieve the outcomes found within the CPD. They can focus on one or more outcomes at a time.

Thematic evaluations: Focusing on cross-cutting issues (e.g., governance, gender equality, and environment) at country office, regional bureau, or HQ levels. They are aligned to IEO corporate/thematic evaluations and provide context-specific insights.

Country programme evaluations: Assessing country programme achievements against planned outcomes. These evaluations also look at the relevance, efficiency, sustainability, and coherence of a country programme. These differ from ICPEs in that they are commissioned and carried out at the country office level rather than through the IEO.

The Types of Decentralized Evaluations in UNDP Tool (Tool 0.2) clarifies the specific functions of the different types of decentralized evaluations, including: what they assess; how they are used; what questions they typically answer; and any specific requirements.

Other evaluations

Impact evaluations: Rigorous assessments that aim to determine the causal effects of UNDP interventions on development outcomes. Unlike other types of evaluations that focus primarily on contribution, impact evaluations establish a credible counterfactual to measure what changes can be directly attributed to an intervention. In line with the UNDP Evaluation Policy for 2025-2030, impact evaluations are gaining greater prominence as a strategic tool. Their use should be prioritized where the demand for causal evidence is high, the methodological conditions are feasible, and the resources are available to ensure sufficient quality.

Donor-commissioned evaluations: Governed by the terms of individual partnership agreements, they may follow donor-specific rules and procedures, and are managed either by UNDP or directly by the donor. These include all Vertical Fund (VF) evaluations related to GEF, GCF, and AF projects.

Joint evaluations: Undertaken with other UN agencies or partners, often for jointly implemented programmes. They provide a holistic assessment of collective contributions and foster system-wide learning. While optional, joint evaluations are encouraged as good practice. UNDP participates in them as a key partner.

Roles and responsibilities

The Evaluation Policy for 2025-2030 delineates clear roles and responsibilities across the system, as follows:

- **The Executive Board** is the custodian of the Evaluation Policy. It approves the IEO multi-year evaluation workplan and budget, and reviews and approves evaluation reports carried out by the IEO and their corresponding management responses and action plans.

- **The UNDP Administrator** safeguards the independence of the evaluation function, ensures adequate resourcing for evaluation, and appoints the IEO Director in consultation with the Executive Board and the Audit and Evaluation Advisory Committee.

- **The Audit and Evaluation Advisory Committee** advises the Administrator in fulfilling their responsibilities regarding, among others, the evaluation function, and commissions independent reviews of the evaluation function.

- **UNDP Country and Regional Offices** *"commission decentralized evaluations according to evaluation plans. These evaluations must be carried out by independent external consultants, with UNDP management taking all necessary actions to ensure the objectivity and impartiality of the process and persons hired."* (Article 64, page 10)

- **The Bureau for Policy and Programme Support (BPPS)** is delegated by the UNDP Administrator to serve as the organizational focal point for the decentralized evaluation function. The BPPS coordinates communication between UNDP management and the IEO, and advises regional bureaux on the decentralized evaluation function for UNDP. Moreover, the BPPS works with the monitoring and evaluation staff in country and regional offices to ensure that evaluation plans are implemented, and that lessons from evaluations inform decision-making. Together with the IEO, the BPPS provides guidance to UNDP on the use of evaluation findings and lessons learned to improve organizational decision-making and accountability. In addition, the BPPS monitors the implementation of management responses to the evaluations.

- **The IEO** is a functionally independent headquarters unit reporting directly to the Executive Board. It conducts independent country programme, organizational and thematic evaluations and evidence syntheses, manages a quality assessment system for decentralized evaluations, develops standards and methodologies, and promotes the use of evaluation findings across UNDP and the UN system.

To complement this division of accountability at system level, these revised guidelines clarify operational roles and responsibilities at each stage of an individual evaluation. Across the sections, each phase includes a dedicated table outlining the specific roles of key actors (e.g., evaluation commissioner, evaluation manager, M&E focal point, and programme teams).

Step 1.

Plan



Developing an evaluation plan is the first step in ensuring that UNDP's evaluations are strategic, timely, and useful.

The evaluation plan outlines all evaluations to be conducted during the programme cycle, aligning them with priorities, resources, and decision-making needs.

The plan must be costed, approved, and uploaded to the ERC, and then reviewed to remain relevant and balanced.

Deliverable: *Evaluation plan*

Tools:

- *Tool 1.1 How to decide which evaluations to include in the UNDP Evaluation Plan*
- *Tool 1.2 Evaluation calendar template*
- *Tool 1.3 Budget considerations and calculation for evaluations*
- *Tool 1.4 Evaluation plan template*
- *Tool 1.5 Evaluation plan checklist*

1.1 Key UNDP Roles Implicated in Step 1. Plan

Level	Who	Main Roles
Country	Resident Representative / Senior Management	Evaluation plan owner - internal review & plan endorsement.
Country	M&E Focal Point	Decentralized evaluation manager - planning, coordination, stakeholder engagement.
Country	Programme and Project Teams	Stakeholders - identifying information needs and supporting requirements for evaluation.
Regional	Regional Evaluation Focal Point	Plan approver and quality advisor – providing capacity support and ensuring country office compliance.
Global	Bureau for Policy and Programme Support	Quality assurance – plan compliance, coverage, and feasibility.
Global	Independent Evaluation Office	Repository holder of global data from ERC for Executive Board reporting. Technical advisor (when requested).

1.2 Evaluation planning

An evaluation plan is a mandatory strategic tool that outlines all evaluations conducted during a programme cycle. It supports accountability, learning, and evidence-based decision-making by ensuring that the evaluation work reflects UNDP's programmatic priorities and informs ongoing and future interventions.

The evaluation plan is annexed to each CPD or RPD submitted to the Executive Board for approval and subsequently uploaded to the ERC.

Evaluations can take between 8-10 months to complete from evaluability assessment to final management response and action plans. The evaluation plan must be practical and ensure that it is fully costed, and is regularly updated to remain relevant throughout the programme cycle. While developing the evaluation plan business units should consider the timing of independent evaluation to avoid duplicating evaluation efforts.

1.3 Why plan evaluations?

Evaluation planning is an integral part of programme design. As UNDP country offices plan activities over a strategic period (e.g. the CPD cycle), they must also plan how progress and results will be assessed and actioned through evaluation. An effective evaluation plan enables UNDP to:

- Map planned results of projects;
- Capture results and lessons learned, through project or strategic evaluations;
- Support course correction when evaluations reveal areas needing improvement; and
- Inform strategic and programmatic decision-making at country, regional, and global levels.

The evaluation plan should be used to monitor the implementation of evaluation commitments, facilitate knowledge sharing, and strengthen management decision-making.

UNDP country offices and regional bureaux may review and update their evaluation plan annually and conduct a formal midterm review to adjust for changes in context, priorities, or resources in accordance with UNDP Portfolio and Project Management (PPM) policies and procedures on managing change.

1.4 Developing the evaluation plan

The Country or Regional M&E Focal Point coordinates the evaluation planning process under the supervision of senior management. They prepare a workplan that defines milestones, along with the timeline for developing the evaluation plan in alignment with the CPD formulation schedule.

To guide the selection of evaluations, Tool 1.1 provides a flowchart illustrating how to decide which evaluations should be included in the evaluation plan, based on mandatory requirements, strategic priorities, and evaluability criteria.

Step 1: Identifying potential evaluations

As per the UNDP Evaluation Policy for 2025-2030: "*All development programmes, projects, and portfolios are subject to a decentralized evaluation in line with strategic priorities and needs.*" (Article 65, page 10)

Programme and project teams, in consultation with partners and donors, propose evaluations. The M&E Focal Point then compiles these proposals into an initial long list.

All mandatory evaluations must be included in the evaluation plan. The following table outlines the *minimum* thresholds. Country and regional offices are encouraged to use these mandatory requirements to ensure a strong evidence base, and to plan for evaluations according to their individual information needs and accountability requirements.

When considering the minimum thresholds, the following principles apply:

1. High value and high visibility require accountability and learning.

2. Investing in a new phase or the scaling-up of programming requires an understanding of what was learned, while adaptation and change can be added to new phases to better reach outcomes.

3. Clusters of projects that seek to achieve the same outcome can support strategic planning and decision-making.

4. As identified in the Evaluation Policy, short-term delivery instruments including project initiation plans and development services "are not subject to evaluation unless deemed strategic necessary to measure development impact, support the scale up of initiatives, mobilize resources, or meet funding partner requests."

Category	Mandatory evaluations
Project-level	Projects ≥ $10 million: 2 evaluations (midterm and final)*. Projects no matter their value, but lasting more than 5 years: 1 evaluation (midterm or final). Projects entering a second or subsequent phase: 1 evaluation before new phase**. Initiatives proposed for scaling up: 1 evaluation before expansion.
Partnership and donor requirements	GEF, AF, and GCF projects follow respective requirements: GEF medium-sized projects (MSPs) with GEF grants over $2 million and all full-sized projects (FSPs) require a midterm review; All GEF MSPs, FSPs, and programmes require a terminal evaluation; All GCF-funded activities require interim and final evaluations ; AF projects with 4 or more years of implementation require a midterm review; All AF projects require a final evaluation; Donor and cost-sharing agreement with required evaluations.
Outcome and country-level	At least one strategic evaluation during the CPD/RPD cycle***.

*Projects and programmes of high financial value should always be evaluated. If projects or programmes are shorter than five years but high in value and meet the minimum threshold of $10m, country offices may choose to do only one evaluation.

Projects and programmes that are high value and meet the minimum threshold and are considered for additional phases do not need to be evaluated more than once. These projects meet at least one of the minimum thresholds.

**Projects that will carry on from one phase to the next should be evaluated to understand design and implementation needs and are in line with UNDP's foundations as an evidence-based and learning organization. Evaluations should be done in the penultimate year of programming to ensure sufficient time to implement recommendations.

***Strategic evaluations may help country offices meet their evaluation thresholds. Projects considered for outcome-level, thematic or portfolio evaluations may be counted in the number of project evaluations planned for the year.

Evaluation plans should include both project and strategic evaluations. Where multiple eligible evaluations are identified, priority should be given to those that:

- Meet minimum evaluation thresholds;
- Inform key strategic or policy decisions (e.g. CPD formulation and programme scaling, redesign, or exit);
- Address emerging or innovative areas where evidence is limited or uncertainty is high;
- Offer strong potential for learning or replication; and
- Integrate cross-cutting priorities.

These minimum evaluation thresholds amount to a framework. Country offices should judge for themselves when deciding if evaluations falling outside of the minimum would be beneficial to programming, and plan these evaluations accordingly.

Step 2: Developing an evaluation calendar

The M&E Focal Point prepares a multi-year calendar outlining the sequence of all planned evaluations. Evaluation calendars can be found in the ERC. The calendar should:

- Avoid "bunching" evaluations at year-end or the end of the programme cycle;
- Reflect realistic completion and reporting deadlines;
- Consider upcoming ICPE and thematic evaluations planned in the Executive Board-approved multi-year evaluation workplan;
- Take into account contextual factors such as elections, holidays, or seasonal extreme weather that may affect fieldwork;
- Bear in mind the evaluation submission requirements for GEF, GCF, and AF projects, as some of these (such as the GCF evaluation dates) are legally binding; and
- Ensure sufficient time between evaluation completion and implementation of recommendations.

To guide the process, Tool 1.2 provides a template to check the scope and balance of the evaluation plan.

All evaluations should be completed and uploaded to the ERC within six weeks of their completion.

Step 3: Costing and identifying funding sources

Each evaluation must be fully costed and resourced. Costs should reflect the scope, complexity, and duration of each evaluation, and should include the following broad categories:

- Professional costs:
  Time and fees for evaluators or thematic experts, including travel, daily subsistence allowance (DSA), and translation and interpretation needs.

- Data collection costs:
  Expenses related to interviews, focus groups, and field visits, such as travel, venue hire, participants' transport, hospitality, and any required materials or logistics.

- Communication and dissemination costs:
  Editing, layout, translation, publication, stakeholder validation workshops, and any dissemination activities (e.g. public events and social media campaigns).

- Contingencies:
  Reasonable provisions for unforeseen costs that may arise.

Sources of funds may include:

- Individual project budget lines (for project evaluations and exclusive monitoring activities);
- Country programme budget lines (for CPE or outcome evaluations);
- Cost-sharing across related projects (for thematic or portfolio evaluations); and
- Fees received by UNDP (for some GCF and AF projects).

To support this step, Tool 1.3 provides a template for estimating costs and identifying funding sources. This tool can be adapted by country and regional offices to calculate evaluation budgets in a consistent and transparent way.

For reference, the ERC's planned evaluation overview report gives average budgets for different evaluation approaches at the country, regional, and global levels.

Monitoring and evaluation budgets must be clearly distinguished in programme and project documents. Evaluation budgets should be commensurate with the size of project/outcome/portfolio under scrutiny.

Step 4: Drafting the plan and rationale

The M&E Focal Point prepares:

- A draft evaluation plan, including timing, coverage, costs, and funding sources; and
- A 300-word evaluation rationale for internal Programme Appraisal Committee (PAC) review, explaining how the proposed evaluations contribute to learning, accountability, and balanced coverage.

Tool 1.4 provides an evaluation plan template.

Step 5: Review and quality assurance

The plan and rationale undergo pre-PAC, regional, and HQ PAC review, and then BPPS quality assurance. The Regional Evaluation Focal Point provides quality assurance on scope, feasibility, and mandatory coverage, and may suggest adjustments to ensure that the evaluation is balanced and realistic. Reviewers use a checklist to confirm that the evaluation plan meets all organizational requirements regarding content, coverage, feasibility, and quality.

Tool 1.5 presents the evaluation plan checklist.

Step 6: Approval and submission

The country office senior management team reviews and endorses the final plan before submission together with the CPD to the Executive Board. Once approved, the plan becomes a formal organizational commitment.

Step 7: Uploading to the ERC

Following Executive Board approval:

- The M&E Focal Point uploads the approved plan under the "Plan Details" section in the ERC; and
- Each planned evaluation is entered as a separate record with key metadata (title, type, year, budget, and responsible unit) within the ERC no more than six weeks after the start of the new CPD.

1.5 Making changes to the evaluation plan

The evaluation plan may require adjustment as circumstances change. Country offices should use their discretion when making changes to their evaluation plan, keeping in mind organizational requirements and contextual realities that emerge over the course of the programming cycle.

- **Annual review:** Conducted during country office business planning to check progress, update timelines, and confirm new or completed evaluations.
- **Midterm review:** Recommended at the programme midpoint to ensure continued relevance, balance, and feasibility.

Both reviews should confirm that the evaluations and management responses are up to date and uploaded to the ERC. Adjustments may include:

- Extending completion dates;
- Changing the scope or purpose due to contextual shifts (e.g. crisis settings);
- Adding new evaluations for new projects; or
- Deleting evaluations (exceptionally, as discussed in the "Prepare" section) for the following reasons:
  - Entry in error – duplicate or misclassified;
  - Planned completion is outside of current CPD period;
  - Reorganization of evaluation into another evaluative exercise such as a change from project to incorporation into a strategic evaluation;
  - Evaluation resources are not adequate to produce usable or credible reports; or
  - Changes to operating context that impact evaluability (e.g. security, political, and/or environmental factors).

According to the UNDP POPP, the country office evaluation plan remains part of the Executive Board-approved document and, therefore, changes should follow the steps suggested in the POPP. All changes must include a clear rationale, discussed with and approved by the programme/project boards, based on which the changes are endorsed by the Regional Evaluation Focal Point and updated on the ERC.

> **"Any changes approved by the programme board to the programme document, results framework or evaluation plan must be recorded in a revised document and filed with the board's meeting minutes. Revisions to the results framework or evaluation plan must be updated in organizational systems, including the Evaluation Resource Centre, to facilitate monitoring, reporting and evaluation."**
>
> ***UNDP POPP, Manage Change***

All changes must be reviewed by the country office senior management alongside the M&E Focal Point and the project manager, and then approved by the country office senior management. The submission request should be submitted to the ERC along with clear justification. This justification will be reviewed and approved by the Regional Evaluation Focal Point. All changes to a country programme evaluation plan outside of the annual planning process, will be recorded on the ERC.

1.6 Overseeing compliance

The M&E Focal Point, with the Regional Evaluation Focal Point, oversees implementation through the ERC to ensure:

- Evaluations are completed as planned;
- Reports and management responses are uploaded; and
- Key actions from management responses are implemented and updated in the ERC.

The BPPS monitors overall compliance with evaluation plans and the implementation of management responses and key actions, and follows up with regional bureaux to ensure timely implementation and reporting.

The IEO consolidates global reporting for the Executive Board on:

- Planned vs. completed evaluations;
- Plan changes and corresponding reasons; and
- Status of management responses and key actions.

Step 2.

Prepare



The preparation stage begins with the initiation of an evaluation, as scheduled in the approved evaluation plan - which can be up to six months before the planned commencement of the evaluation.

In this stage, the country office agrees on the management arrangements (who commissions, who manages, and who quality-assures), confirms evaluability, develops and approves the TOR, establishes an evaluation reference group (where relevant), and selects the evaluation team through a competitive and transparent process.

Deliverable: *Evaluation TOR*

Indicative Timeframe: *8-12 weeks*

Tools:

- *Tool 2.1 Evaluability checklist*
- *Tool 2.2 Evaluation TOR template*
- *Tool 2.3 TOR quality checklist*
- *Tool 2.4 Evaluation information sources checklist*
- *Tool 2.5 UNEG Pledge of Ethical Conduct*

Ethics in evaluation: preparation stage

All evaluation commissioners, managers, and evaluators contracted by UNDP should follow the UNEG Ethical Guidelines, which include a pledge of ethical conduct in evaluation (Tool 2.5) and a checklist of ethical issues that should be considered at each phase of the process.

As AI continues to support processes and efficiency, the TOR should provide explicit details on the acceptable ways in which it may be used in the evaluation process. In all cases, the TOR will require explicit reference to the disclosure requirements on the use of AI. This information can be found in the evaluation TOR template.

2.1 Key UNDP staff involved in Step 2. Prepare

Who	Main Roles
Country Office M&E Focal Point	Decentralized evaluation manager - evaluability assessment, coordination of reference group (if used), TOR review and quality assurance, and sharing in the evaluation team selection.
Resident Representative / Senior Management	Decentralized evaluation commissioner - ensuring independence of the evaluation.
Programme and Project Teams	TOR development, participating in evaluation reference group, document access, and sharing in the evaluation team selection.
Reference Group (if used)	Stakeholders - TOR review.

2.2 Evaluation management

All evaluations must have a clearly defined management and oversight structure that specifies who commissions the evaluation, who manages it, and who quality-assures each deliverable throughout the evaluation process. Before initiating any evaluation, the country office must finalize these arrangements.

- The evaluation commissioner, usually the Resident Representative (RR) or Deputy Resident Representative (DRR), is the senior official who is accountable for commissioning the evaluation and ensuring that it is conducted independently and used.

- The evaluation manager is appointed by the evaluation commissioner and is responsible for the day-to-day oversight of the entire evaluation process. They must be a team member independent from the design, implementation, and supervision of the project, programme, or outcome being evaluated (the “evaluand”).

In most cases, the M&E Focal Point should serve as the evaluation manager as they are best positioned to ensure methodological integrity and independence.

In smaller offices where an M&E specialist is not available, the senior management team may designate another staff member with no direct or supervisory responsibility for the evaluand.

To safeguard impartiality, the evaluation manager cannot supervise, manage, or be managed by the programme or project under evaluation and must operate free from undue influence.

2.3 Preparing for an evaluation

Step 1: Conduct the evaluability assessment

An evaluability assessment is a short, structured review undertaken before the evaluation begins, normally up to six months in advance, to determine whether the subject of the evaluation, called the *evaluand* (project, programme, outcome, portfolio, or thematic area), is sufficiently clear, measurable, and resourced to allow a credible, useful, and timely evaluation.

Who	Main Role
Evaluation Manager / M&E Focal Point	Oversees evaluability assessment including recommendations on evaluation proceedings.
Programme / Project Manager	Supports access to project and programme documents Identifies data gaps and mitigations.
Evaluation Commissioner	Reviews results of evaluability assessment. Approves decisions to proceed, postpone, or delete evaluations in the evaluation plan.
Regional Evaluation Focal Point	Reviews justification for postponement/deletion. Approves ERC update.

The evaluability checklist (Tool 2.1, below) guides this process and helps the evaluation manager to document the readiness and recommended course of action for the evaluation commissioner's approval.

Evaluability checklist	
The subject of the evaluation has a clearly defined theory of change and shared understanding of what will be evaluated.	☐
There is a well-defined results framework with SMART indicators linking outputs, outcomes, and goals.	☐
Sufficient data and documentation are available (baseline, monitoring, progress reports, and prior evaluations).	☐
The planned evaluation is still relevant to the evolving context and organizational priorities.	☐
The political, social, and economic context allows effective evaluation implementation and use of results.	☐
Adequate human and financial resources are available for the conducting of the evaluation.	☐

Together, these conditions determine whether the evaluation can proceed as planned, if it requires corrective actions to strengthen its readiness, or should be postponed.

The evaluation manager leads the evaluability assessment, coordinating inputs from programme or project managers and other relevant staff. Key stakeholders and national counterparts should be consulted to confirm the relevance of the evaluation and to validate data availability, access constraints, and contextual risks.

Special considerations for evaluability in crisis settings

Evaluability assessments in crisis or high-risk contexts must include the following questions:

- Is an evaluation required for this project or programme? What are the risks of carrying out this evaluation vs. not carrying out this evaluation?
- Does the project or programme have an up-to-date conflict analysis? If so, what are the key findings from this analysis and how do the findings affect the carrying out of the evaluation?
- Can conflict-sensitive approaches related to the "do no harm" principle be identified for this evaluation? If so, what are these approaches?
- What plan is already in place or still needed to prioritize the safety of evaluators, staff, and participants?
- Is the evaluation design flexible and is adaptation possible? (as an example, this could mean identifying alternative data collection options (e.g. remote interview or use of secondary data)). What are the alternative options?
- What are the key ways in which this evaluation could impose operational and programmatic burdens on the country office and national stakeholders? What mitigatory measures could be taken to lessen these burdens?

Delaying an evaluation

If the assessment concludes that a project or programme is not ready for an evaluation, country offices may revise the evaluation plan and submit the updated completion date and corresponding justification through the ERC for review and approval by the Regional Evaluation Focal Point.

Deleting an evaluation

According to the UNDP POPP Manage Change document, evaluations may be deleted when:

- The entry was added in error (duplicate or misclassified);
- The planned completion date falls outside the current CPD period (and is thus to be moved to the next plan);
- The evaluation has been merged into another exercise (e.g., portfolio, outcome, thematic, or regional evaluation);
- Funding is insufficient and the evaluation was not properly costed; or
- Security, political, environmental, or health conditions make it impossible to conduct the evaluation safely or effectively.

If an evaluation is deemed no longer relevant or is not expected to meet evaluability requirements, then UNDP senior management and the programme team should review and approve deletion with the M&E Focal Point or specialist and project manager, ensuring that the programme or project board has approved the deletion.

The deletion request should be submitted to the ERC, along with clear and detailed justification, for review and approval by the Regional Evaluation Focal Point. All changes to the evaluation plan will be recorded on the ERC to support and strengthen oversight of the plan's implementation.

*Optional: establish the evaluation reference group**

**Evaluation reference groups are not always necessary but should be considered when there are high levels of visibility and resource requirements associated with a programme/project.*

An evaluation reference group may be established to support transparency, inclusiveness, and quality in the evaluation process. They are particularly useful when the evaluand is complex or where sensitivities exist in which the information needs demand additional subject-matter expertise to help guide the evaluation.

The evaluation commissioner and evaluation manager should consider forming an evaluation reference group composed of key partners and stakeholders who can provide technical input, contextual knowledge, and feedback at critical stages of the evaluation. A reference group is a valuable addition to the preparation of strategic evaluations and, in particular, portfolio evaluations.

Who	Role
Evaluation Manager	Coordinates reference group.
Evaluation Reference Group Members	Provides technical and contextual input on draft TOR, inception report, and draft evaluation report.

The evaluation reference group helps to:

- Ensure that different perspectives - including those of government counterparts, implementing partners, donors, and civil society - are reflected throughout the evaluation process;
- Review and comment on key deliverables such as the draft TOR, inception report, and draft evaluation report; and
- Strengthen the credibility, ownership, and use of the evaluation findings and recommendations.

The evaluation reference group is advisory in nature and does not interfere with the independence of the evaluation team or the evaluation manager. Its members are selected by the evaluation manager, in consultation with the evaluation commissioner, and their roles and responsibilities should be clearly defined from the beginning.

Even without an evaluation reference group, it is useful to share key evaluation deliverables with relevant stakeholders to gain their feedback. Consultation on key deliverables supports greater alignment with information needs, corporate requirements, and UN values, allowing for a diversity of viewpoints to be included at each step of the process.

Step 2: Draft the evaluation TOR

Who	Role
Programme Team	Prepares draft TOR using evaluation TOR template.
Evaluation Manager	Reviews evaluation TOR for compliance for quality and with corporate standards and System-wide Action Plan Evaluation Performance Indicator (SWAP EPI) requirements. Coordinates with evaluation reference group on feedback and adjustments.

The TOR is the foundation of every evaluation. It defines its scope, purpose, and expectations, and guides the evaluation team and management throughout the process. A clear, well-structured TOR ensures that the evaluation is relevant, feasible, credible, and useful.

Purpose and importance

High-quality TORs provide a clear mandate to the evaluation team, stating what is being evaluated and why, who will use the results, and how the findings will inform decision-making. Because the TORs set the parameters for quality assurance and the use of results, adequate time for consultation and revision (normally one to two months before inception) must be allocated for their preparation, consultation, and approval.

Drafting the TOR

The programme or project team prepares the first draft using the evaluation TOR template (Tool 2.2). The TOR should be driven by the evaluation's purpose, scope, and objectives in light of the available resources, time, and information gaps. It is important that the document clearly expresses the information needs of the evaluation – with the purpose and objectives clearly addressing the following questions:

1. Does this evaluation meet accountability needs? And for whom and in what ways?
2. Does this evaluation seek to understand something? What do we want to learn from this evaluation?
3. What decisions or actions will this evaluation feed into?
4. Will the timeline of the evaluation need to be adapted to support evidence-to-action needs?

Evaluation TORs for interventions supported by GEF, GCF, AF, as well as portfolio and impact evaluations commissioned by UNDP, follow different formats. For GEF midterm and terminal project evaluations, the ERC Methods center provide separate TOR templates. For all VF evaluations, please see this guidance.

Each TOR should include the following sections (full details and guidance for each section are provided in the evaluation TOR template (Tool 2.2)):

TOR section	Instructions
Background and context	Describe the intervention, theory of change, main results, target groups (with special attention paid to cross-cutting issues such as gender, disability, and the LNOB principle), and relevant contextual factors.
Evaluation purpose, scope, and objectives	Explain why the evaluation is being undertaken, who will use its findings, and how results will inform decisions (use). Include gender equality, disability and inclusion, and LNOB considerations related to the scope and objectives.
Evaluation criteria and key questions	Structure questions around OECD/DAC criteria according to information or accountability needs. Ensure inclusion of cross-cutting issues including gender equality, disability inclusion, and the LNOB principle. Adhere to: • UNEG Guidance on Integrating Human Rights and Gender Equality in Evaluations (2024); and • UNEG Guidance on Integrating Disability Inclusion in Evaluations (2022). • UNEG Norms and Standards on Environmental and Social Impacts (2025). Consider the purpose and objectives of the evaluation when drafting questions. Use open-ended questions to allow for unintended findings to surface, to provide greater contextualization in the findings and allow for nuanced insights.
Methodology	Outline the proposed evaluation approach, methods, and techniques. The TOR should require a mixed-methods approach as a minimum. Refer to UNDP Methods Center for details on additional types of approaches that will support different evaluation needs. For example, portfolio evaluations should focus on mixed methods and a systems-based approach to ensure that the information collected in the evaluation meets its purpose and objectives. Allow the evaluation team to refine methods during inception. The overall methodology should ensure that it highlights how it will integrate cross-cutting issues into the data collection and analysis of the report. Cross-cutting issues related to gender equality, disability inclusion, and the LNOB principle should be clearly identified within the methodology.
Expected products and deliverables	List required outputs (inception report, draft report, audit trail, final report, evidence binder, debriefing, and knowledge products).

TOR section	Instructions
Team composition and required competencies	Specify qualifications, expertise, and ethical standards. Gender and disability inclusion expertise are required. Adhere to: • UNEG Evaluation Competency Framework (2016).
Ethical standards	Reference UNEG Ethical Guidelines for Evaluation and require all evaluators to sign the UNEG Pledge of Ethical Conduct (Tool 2.5). Make explicit reference to the disclosure of use of AI within the evaluation process. Reference the UNEG Ethical Principles for Harnessing AI in United Nations Evaluations.
Management and coordination arrangements	Clarify roles, responsibilities, and feedback mechanisms for each stage. This should identify what the programme team, the evaluation manager, and the evaluation team are each respectively responsible for throughout the evaluation.
Timeline and budget	Provide indicative schedule and number of workdays per evaluation team member, including deadlines for inception, data collection, and reporting.
Submission process and selection criteria	Outline how applications will be evaluated and selected.
Annexes	Attach or provide links to the project results framework, theory of change, evaluation matrix, stakeholder map, and relevant guidance and templates.

Consultation, quality assurance, finalization, and approval

The programme team shares the draft TOR with the evaluation manager for review. If the evaluation has a reference group, the evaluation manager facilitates their consultation on the draft as well.

During review and consultation, the TOR should be reviewed for quality, feasibility, and balance, and revised accordingly (if required). The evaluation manager is responsible for ensuring that the TOR meets UNDP quality standards, as indicated in the UNDP Evaluation Quality Assessment System. To facilitate internal quality assurance, the evaluation manager should complete the TOR quality assurance checklist (Tool 2.3), derived from the Evaluation Quality Assessment System, to confirm that all standards have been met prior to approval.

Once finalized, the TOR is submitted to the Regional Evaluation Focal Point for approval. After approval, the TOR must be uploaded to the ERC before the recruitment of evaluators begins.

Examples of high-quality TORs

When developing or reviewing an evaluation TOR, teams should consult the ERC for examples of high-quality TORs. These examples reflect good practice and alignment with UNDP and UNEG evaluation standards, and are intended to guide the clear definition of scope, evaluation questions, roles, and quality expectations from the outset.

Step 3: Select the evaluation team

Who	Role
Evaluation Manager	Defines evaluator qualifications and team composition. Oversees the selection of the evaluation team in coordination with operations teams.
Evaluation Commissioner	Approves final selection and contracting of evaluation team. Ensures independence and manages conflict-of-interests declarations.
Regional Evaluation Focal Point	Advises on selection (if requested).

The selection of evaluators is a critical factor in determining the credibility, quality, and utility of an evaluation. Evaluators must be selected through a fair and transparent process, in line with UNDP's procurement rules and regulations and guided by UNEG Ethical Guidelines for Evaluation.

Step-by-step process

1. Plan the selection early

The recruitment process should be planned early and be included in the evaluation timeline. Allowing sufficient time for the recruitment and contracting of evaluators is vital, as the process can take longer than anticipated and may require adding more time to the evaluation schedule.

Before launching the selection process, the evaluation manager should ensure that:

- The TOR is finalized;
- The evaluation budget is defined; and
- The evaluation arrangements are confirmed.

2. Determine the evaluation arrangements

Depending on the scope, complexity, and resources available, the evaluation may be conducted by:

- An individual evaluator (national or international);
- A team of evaluators, composed of both national and international experts; or
- A firm which provides one or several evaluators.

Even when contracting through a firm, the evaluation manager must review and endorse the CVs and composition of the proposed team to ensure they meet the TOR requirements.

3. Define required qualifications and expertise

Evaluators must demonstrate the following competencies in line with the UNEG Evaluation Competency Framework:

- Proven experience in conducting and managing evaluations;
- Strong analytical and report-writing skills, including experience in qualitative and quantitative analysis;
- Technical expertise relevant to UNDP and the evaluation's thematic area(s);
- Demonstrated knowledge of cross-cutting areas (human rights, gender equality, disability inclusion, and the rights-based approach); and
- Familiarity with national or regional contexts.

When forming teams, diversity and complementarity of skills (thematic, methodological, and contextual) as well as gender balance and geographical representation should be pursued.

Evaluation teams should comprise at least two evaluators to compare notes, verify the accuracy of information collected and recorded, allow for a division of efforts to interview more people, and share ideas.

4. Identify and shortlist candidates

The evaluation manager should draw from the following available rosters and databases to ensure a wide and qualified pool of candidates:

Source	Description	Notes
Regional Evaluator Rosters	Maintained by regional service centres; include vetted and frequently used evaluators.	Contact Regional Evaluation Focal Point for advice or access.
ERC	The ERC indicates previously used evaluators, who can be contacted by the evaluation manager for recruitment.	Search by name, thematic area, or region, and check quality assessment ratings.
Global Policy Network (GPN) / ExpRes Roster	UNDP's roster of technically vetted consultants across thematic and evaluation profiles. Managed by GPN / ExpRes.	Request via GPN / ExpRes deployment form. No financial proposals are required.
External Professional Networks	For example, UNEG, EvalPartners, and national and regional evaluation societies.	Use when internal rosters lack relevant expertise.

Safeguard independence and manage conflicts of interest

Independence and impartiality are fundamental to the credibility of UNDP evaluations. UNDP must ensure that all evaluators are free from conflicts of interest (both real and perceived).

Evaluators must not have worked on, or contributed to, the programme, project, or outcome under evaluation, in any capacity, at any time. They must not have any expectation of securing future contracts linked to the subject or unit being evaluated. In both cases, the impartiality of the analysis would be compromised.

As a matter of good practice, an evaluator who has conducted a midterm evaluation should not be contracted for the final evaluation of the same intervention.

Common sources of conflicts of interest include:

- **Past involvement:** Consultants or institutions previously engaged in the design, implementation, financing, or management of the project or programme being evaluated.
- **Institutional affiliation:** UNDP staff, government officials, NGO staff, or contractors directly or indirectly involved in the same programme or project previously.
- **Personal relationships:** Evaluators who have close family relationships with UNDP personnel directly involved in the same programme or project.
- **Dependence on future work:** Evaluators who are likely to seek future assignments with the same unit, or to design or implement follow-up phases of the evaluated intervention.
- **Multiple roles:** Consultants who have undertaken multiple advisory or management roles with UNDP in the same thematic area or country, creating perceived bias or over-familiarity.

UNDP country and regional offices should also consider whether a consultant's overall reliance on UNDP assignments could affect perceived independence. The ERC provides a record of an evaluator's past UNDP evaluations and their quality assessment ratings, which should be checked before selection.

Each evaluator is required to declare any potential or actual conflict of interests at the time of application and again upon contracting, by signing the UNEG Pledge of Ethical Conduct (Tool 2.5). No contract should be signed if there is a conflict of interests. If an unforeseen conflict of interests arises during the process, the evaluation manager must assess its implications and, if necessary, remove or replace the evaluator.

To strengthen transparency and prevent disputes:

- Share the CVs of shortlisted candidates with key stakeholders for feedback prior to final selection. This should be carefully managed to prevent undue influence over selection decisions. Evaluation managers should be mindful that stakeholder consultations may, in some contexts, lead to the exclusion of qualified candidates for non-technical or political reasons, or to the implicit preference for evaluators perceived as "aligned" with particular interests.
  The objective is to identify an evaluator who is acceptable to key stakeholders without being dependent on, or influenced by, any of them. Clearly setting expectations about the independent role of the evaluator is therefore essential.
- Conduct reference checks with colleagues or partners who have previously worked with the candidate.
- Use ERC quality assessment scores as an additional check on past performance.

By applying these safeguards, UNDP ensures that evaluations remain independent, objective, and credible, thereby upholding the principles of impartiality and integrity defined in the UNEG Norms and Standards for Evaluation in the UN System.

Finalize selection and contract

The evaluation manager, in consultation with the operations and procurement teams, conducts the technical and financial evaluation of candidates, documents the selection process, and finalizes the selection of the evaluation team. The final selection decisions are determined by the evaluation manager, who acts as the independent technical expert in the selection process.

Once the contract is issued:

- Ensure all evaluators sign the UNEG Pledge of Ethical Conduct (Tool 2.5) before the work begins.
- Share the inception report outline (Tool 3.1) and evaluation report template (Tool 5.1) with evaluators.
- Share the quality assessment questions checklist (Tool 7.1) with the evaluation team.

Step 4: Compile and share a document library

Who	Role
Programme / Project Manager	Facilitates access to key documents and fills data gaps. Uses the evaluation information sources template (Tool 2.4).
Evaluation Manager	Shares compiled information with evaluation team.
Evaluation Team	Reviews all documentation as part of desk review. Identifies information gaps.

Once the evaluation has been launched, the evaluation manager works with the programme or project team to compile and share a document library containing all relevant information that will enable the evaluation team to conduct a comprehensive desk review and prepare the inception report. The programme or project team identifies relevant documentation, providing missing information and facilitating access to internal systems or national data sources.

Content of the document library

The document library should include the following categories of information:

- Strategic and results frameworks;
- Monitoring and evaluation data;
- Donor, government, and contextual data;
- Financial and management information;
- Decision documents; and
- Reports.

The evaluation team should be provided with contact details of focal points for follow-up on specific datasets or restricted documents.

Step 3.

Inception



The inception stage begins once the evaluation team has been contracted, and concludes with the approval of the inception report.

This report ensures that all parties share a common understanding of the evaluation's focus, methods, and expected outputs before any data collection starts.

It is a workplan that serves as a bridge between planning and implementation, translating the TOR into an operational design.

Deliverable: *Inception report*

Tools:

- *Tool 3.1 Inception report template*

Indicative Timeframe: *4-6 weeks*

Ethics in evaluation: inception stage

The evaluation team must behave ethically in all interactions with stakeholders. In the inception phase, evaluators need to indicate adequate ethical safeguards in the inception report, which should also include an assessment of risks and mitigation measures. Evaluators must also avoid conflicts of interest and discretely report any discovery of apparent misconduct (in line with UNEG Ethical Guidelines for Evaluation).

Step 1: Brief the evaluation team

Who	Role
Evaluation Manager / M&E Focal Point	Kicks-off evaluation launch. Ensures a shared understanding of the evaluation purpose, scope, roles, timelines, and deliverables.
Evaluation Team	Confirms timelines, deliverables, and communication channels.
Programme / Project Team	Provides background context at kick-off and confirms focal points.

The inception phase starts with an orientation briefing and/or kick-off meeting involving the evaluation manager, the programme/project team, and the evaluation team.

Together, they ensure that all parties are aligned on the purpose, standards, and working arrangements before the evaluation design is finalized.

The orientation briefing and/or kick-off meeting are presented separately in these guidelines, however they may be carried out at the same time to support efficiency and evaluation timelines.

Orientation briefing

Held soon after the contract is signed, the orientation briefing is a focused technical meeting involving the evaluation manager and the evaluation team. Its purpose is to familiarize the hired evaluators with UNDP's evaluation requirements, procedures, and quality expectations before engaging with other stakeholders.

The meeting establishes a shared understanding of deliverables, roles, and working arrangements, and provides an opportunity to clarify any outstanding matters on the TOR. Its objectives are to:

- Confirm the evaluation scope, deliverables, and overall timeline;
- Review UNDP quality standards for evaluation reports, UNEG ethical obligations including disclosure on the use of AI, and independence provisions;
- Clarify roles, reporting lines, communication channels, and logistical arrangements; and
- Introduce key resources, including the inception report template, evaluation guidance materials, and document library.

Kick-off meeting

Following the orientation, the evaluation manager convenes a broader kick-off meeting to officially launch the evaluation with key stakeholders, such as programme management and the evaluation reference group, where relevant.

This meeting introduces the evaluation team, promotes transparency and ownership, and establishes cooperation modalities for the evaluation. Its objectives are to:

- Present the evaluation purpose, scope, and intended use of its findings;
- Confirm the roles of the evaluation manager, evaluation reference group, and programme or project manager;
- Clarify coordination and communication arrangements;
- Discuss logistical access to data, field sites, and stakeholders; and
- Reiterate principles of independence, impartiality, and ethical conduct.

The following table summarizes the typical agenda for the two meetings.

Meeting	Participants	Typical agenda
Orientation briefing	Evaluation manager (chair), evaluation team.	Welcome and introductions. Overview of how evaluation is done at UNDP, the Evaluation Policy, and UNEG Norms and Standards. Review of the TOR. Quality and ethical standards. Mission logistics and security, BSAFE security training, and travel and visa arrangements. Documentation and data access, an update on available materials, and the process for requesting missing information. Assignment of roles and confirmation of evaluation manager and operational focal points for documentation, logistics, and administration. Next steps and scheduling of the kick-off meeting.
Kick-off meeting	Evaluation manager (chair), evaluation team, programme or project manager, evaluation reference group (if established for the given evaluation).	Opening and introductions. Purpose and expectations: objectives of the evaluation and how findings will be used. Background and context: presentation by programme or project manager on implementation progress, strategic relevance, and positioning within the broader UNDP programme. Evaluation overview, presentation of the evaluation team's initial understanding of the TOR, planned approach, and next steps. Stakeholders and fieldwork planning, identification of key partners, and suggestions for field visits or site selection. Roles and facilitation including clarification of how programme management will support access to information, logistics, and contacts, and confirmation of focal points for documentation, visa issues, and logistics. Timeline and next deliverables – confirmation of milestones, submission of inception report, and review process. Q&A, closing summary, and next steps.

Step 2: Inception report

Who	Role
Evaluation Team	Drafts inception report using the inception report template (Tool 3.2).
Evaluation Manager / M&E Focal Point	Oversees the inception phase including approval of the inception report.
Regional Evaluation Focal Point	Offers on-demand guidance to the evaluation manager before finalizing the report guidance to the evaluation manager before finalizing the report.

The inception report is the first substantive deliverable of the evaluation and a key milestone in the evaluation process. Prepared by the evaluation team after reviewing background documentation and holding initial consultations, it translates the TOR into a detailed and operational plan for conducting the evaluation.

The report presents the team's understanding of the evaluand, the purpose and scope of the evaluation, and the conceptual and methodological framework that will guide subsequent phases. These components should demonstrate logical links between what the evaluation will study and why, and how it will study it in order to generate accurate and usable findings, conclusions, and recommendations.

Purpose of the inception report

The inception report ensures that all parties share a common understanding of the evaluation's focus, methods, and expected outputs before any data collection begins. It provides an opportunity to clarify the evaluation scope, questions, methodology, and resource requirements, and to address any ambiguities or constraints identified during the initial review.

Inception report content

The inception report should include:

- A short description of the evaluand's context, objectives, and theory of change;
- The purpose, objectives, and scope of the evaluation;
- A summary of findings from the initial desk review;
- A summary of the limitations identified in the evaluability assessment and mitigation measures for them;
- The evaluation criteria and evaluation questions;
- The evaluation approach, methods, and techniques to be used, including data sources, sampling, and analytical strategies;

- A preliminary evaluation matrix linking questions, indicators, data sources, and techniques;
- Consideration of cross-cutting issues such as gender equality, human rights, disability inclusion, and the LNOB principle;
- A detailed workplan, schedule, and resource requirements; and
- The proposed structure of the final report.

For detailed guidance, evaluators should refer to the inception report template (Tool 3.2), which provides a structure and descriptions for each section, aligned with UNDP evaluation quality standards.

Step 3: Review and approval

Depending on the requirements outlined in the contract with the evaluation team, the inception report should be submitted within two to four weeks of contract signature, in line with the timeline specified in the TOR.

The evaluation manager circulates the draft to the evaluation reference group (if applicable), and other relevant stakeholders for review and quality assurance, ensuring that it meets UNDP's methodological, ethical, and quality standards.

Feedback should focus on the adequacy of the evaluation design, the clarity of the scope and questions, and the feasibility and logic of addressing the evaluation questions using the proposed methods.

The inception report must be formally approved by the evaluation manager before the evaluation proceeds to data collection.

Step 4.

Data collection



Data collection is the stage in which evaluators gather the evidence required to answer the evaluation questions as indicated in the evaluation matrix.

It combines desk-based and field-based activities, with the field mission being the central component of qualitative methods.

During this phase, the evaluation team collects and synthesizes information from primary and secondary sources, ensuring evidence is credible, verifiable, and representative of diverse stakeholder perspectives.

Data collection begins after the inception report has been cleared and concludes with an end-of-mission debriefing before the drafting of the evaluation report.

Deliverable: *Mission debriefing*

Tools:

- *Tool 4.1 Debriefing presentation template*
- *Tool 4.2 Mission agenda template*
- *Tool 4.3 Consent statement for interviews with project participants, project implementers, community members and local experts/leaders*

Indicative Timeframe: *3-4 weeks*

Ethics in evaluation: data collection stage

The evaluation team must behave ethically in all interactions with stakeholders by protecting privacy, confidentiality, and anonymity, while ensuring informed consent and maintaining cultural sensitivity. They must follow the UNEG Ethical Guidelines for Evaluation and ensure that the participation of stakeholders in the evaluation does not harm them or their communities.

Informed consent is required for all interactions with stakeholders during the data collection stage. In particular, informed consent requires clear communication of the evaluation activity, how the information collected will be used and stored, and the extent to which stakeholders have the right to retract their consent.

Step 1: Prepare the data collection mission/build the agenda

Who	Role
Evaluation Commissioner	Authorizes data collection field missions.
Evaluation Manager	Oversees data collection planning, acting as interlocutor when needed.
Programme / Project Manager	Supports planning of meetings and site visits for the evaluation team.
Evaluation Team	Organizes the data collection mission in collaboration with the programme/project team, identifying site and group selection and providing justification for these choices. The evaluation team ensures that cross-cutting issues such as gender equality, disability inclusion, and the LNOB principle are balanced in site and group selection. Integration of these important cross-cutting issues is outlined in the UNEG Guidance on Integrating Human Rights and Gender Equality in Evaluations. Ensures compliance with ethical and data protection requirements and follows the "do no harm" principle as outlined in the UNEG Ethical Guidelines for Evaluation and UNDP's Personal Data Protection and Privacy Policy.
Regional Evaluation Focal Point	Provides technical advice to the evaluation manager on changes due to crisis or restricted-access contexts.

The data collection mission builds on the approved inception report, which specifies the resources, logistical support, and fieldwork arrangements needed to complete the evaluation.

To facilitate planning and coordination of fieldwork, the mission agenda template (Tool 4.2) provides a standardized format for detailing the evaluation itinerary, stakeholder meetings, security and logistical arrangements, and communication protocols.

The evaluation team makes the bulk of the decisions during mission planning. Site visits and interview selection decisions should be made by the evaluation team, in coordination with the programme/project team to ensure the operational requirements of the selection can be maintained. Justification for choices made should be well documented and any changes to these choices should be noted by the evaluation team within the itinerary. Considerations of gender equality and inclusion are required to ensure that selected sites and key informants reflect a diversity of perceptions and programming realities.

Overall, as an independent exercise, the evaluation team's choices should be *informed* by insights of the programme/project team but not *governed* by them.

The evaluation team finalizes interview protocols, focus group protocols, consent forms, and note-taking templates consistent with the approved methodology, UNEG Ethical Guidelines for Evaluation, and UNDP's Personal Data Protection and Privacy Policy.

Prior to the mission, the evaluation team ensures that all members are familiar with informed consent procedures and UNDP's requirements on personal data protection, anonymity, and the "do no harm" principle. All security protocols must be in place before the evaluation team begins the data collection mission.

Step 2: Conduct fieldwork and data collection

Who	Role
Evaluation Team	Conducts and documents highquality field data collection. Ensures ethical, safe, and inclusive evaluation practice while engaging with project and programme stakeholders.
Evaluation Manager / M&E Focal Point	Maintains oversight of field progress. Ensures compliance with ethical standards, including the "do no harm" principle, and that ethical safeguards are appropriately documented.
Programme / Project Manager	Facilitates access to project sites and data sources, and addresses coordination, access, or logistical issues. Ensures no undue influence on respondents or evaluation content.

The evaluation team leads all data collection activities independently, following the approved methodology, ethical standards, and logistical arrangements agreed in the inception report. Data collection combines individual and group interviews, focus group discussions, direct observation, and review of documentary and quantitative evidence. The team triangulates data from multiple sources to answer the evaluation questions and test the theory of change.

Field missions typically last one to three weeks, depending on the size of the programme/project, the scope of the assessment, and accessibility of project sites and stakeholders. If contextual factors (e.g., insecurity, travel restrictions, or extreme weather) hinder access, the team may adjust the methodology – for instance, by using remote interviews or alternative sites - while documenting and justifying any such changes in the evaluation report.

Conducting fieldwork

The evaluation team conducts meetings and site visits according to the agreed agenda. Day one of fieldwork usually begins with a security briefing delivered by the United Nations Department of Safety and Security (UNDSS), followed by interviews with UNDP staff to clarify programme context and strategic positioning. The next days are dedicated to interviews and focus group discussions with national counterparts, partners, beneficiaries, and other external stakeholders, as outlined in the approved stakeholder list.

Facilitating logistical and operational support

The programme/project team provides logistical and administrative support to the evaluation team, including assistance with visa issuance, local transportation, access to UN premises, and temporary office space to enable the team to conduct interviews and analysis securely and efficiently.

The evaluation manager oversees the data collection mission's progress but does not attend or participate in interviews or focus group discussions, in line with the principles of independence and confidentiality.

Ensuring independence and impartiality

The evaluation team carries out all interviews and field observations independently. UNDP staff and partners may facilitate introductions or provide background information but must not be present during interviews or focus group discussions with stakeholders or beneficiaries.

Adherence to ethical and data protection standards

Evaluators must uphold the highest ethical and data protection standards throughout data collection, following UNEG Ethical Guidelines for Evaluation.

Informed consent requires that participants clearly understand the purpose of the evaluation, what their participation involves, the voluntary nature of participation, as well as how their information will be used and with whom it may be shared. When collecting personal data, through interviews, surveys, or focus groups, evaluators must inform participants of their rights, including on access, deletion, and objection to use, and provide a contact point for any concerns regarding personal data (Tool 4.3).

Confidentiality must always be guaranteed: evaluators have to protect the identity of participants when reporting findings and ensure that only authorized team members have access to raw data. In some evaluations, anonymity cannot be guaranteed (e.g. where there are small groups or identifiable roles) and this must be explained to participants clearly. Evaluators should discuss whether interviews will be recorded or transcribed, obtain explicit consent for recording, and ensure that any audio files, transcripts, or notes are handled securely.

All personal data collected during the evaluation are handled in accordance with UNDP and UNEG Ethical Guidelines for Evaluation. Evaluators will apply appropriate measures to protect confidentiality and anonymity, which may include secure storage of files, restricted access to data, anonymization of transcripts and notes, separation of identifying information from evaluation data (codebook separation), and avoidance of unnecessary collection of personal identifiers. Throughout the mission, evaluators must apply the "do no harm" principle, respect cultural norms, and take measures to avoid risks to participants or communities. Data protection, management, retention, and disposal arrangements will be detailed in the inception report (Tool 3.1).

Step 3: End-of-mission debriefing

Who	Role
Evaluation Team	Prepares and presents preliminary observations and data patterns during debriefing presentation using the debriefing presentation template (Tool 4.1).
Evaluation Manager / M&E Focal Point	Coordinates the debriefing process, ensuring appropriate timing, participation, and format. Facilitates constructive dialogue, while safeguarding the independence of findings emerging from the evaluation.
Programme / Project Manager	Attends debriefing presentation. Provides logistical feedback and identifies follow-up data needs.
Evaluation Commissioner	Attends debriefing presentation and reinforces the importance of confidentiality and ethical conduct.

At the conclusion of the data collection mission, the evaluation team prepares and delivers a data collection debriefing to present initial observations and emerging patterns. The debriefing serves as an opportunity to highlight preliminary observations, validate information, clarify any misunderstandings, and ensure transparency before the team proceeds to the analysis and reporting phase.

The end-of-mission debriefing should not present findings, conclusions, or recommendations. It is meant to give the evaluation stakeholders an opportunity to hear initial insights into the programme or project, and raise any observed issues that require specific attention.

To support this process, the debriefing presentation template (Tool 4.1) is to be referred to.

Content of the debriefing presentation

The presentation should summarize:

- The evaluation objectives, scope, and methodology, including any changes made during fieldwork;
- The sites visited and stakeholders consulted;
- Preliminary observations and early patterns emerging from the data;
- Information gaps or issues that may require clarification; and
- Next steps and timelines for the draft report and quality assurance process.

Good practice

One day should be allowed between the end of data collection and the debriefing to enable the team to consolidate and make sense of observations. Presentations should be concise, use visual summaries where possible, and focus on observations rather than recommendations.

Step 5.

Data analysis and reporting



The data analysis and reporting stage transforms the evidence collected during the evaluation into findings, conclusions, and recommendations. It involves systematic analysis, triangulation, and synthesis of data to answer evaluation questions and assess performance against the evaluation criteria. The process culminates in the evaluation report.

Deliverable: *Evaluation report and annexes*

Tools:

- *Tool 5.1 Evaluation report template*
- *Tool 5.2 Evaluation report quality checklist*
- *Tool 5.3 Audit trail template*

Indicative Timeframe: *6-10 weeks*

Ethics in evaluation: data analysis stage

During data analysis and reporting, the evaluation team should indicate whether any risks identified in the inception report materialized and, if so, describe how they were addressed and what mitigation measures were applied.

If, during the evaluation, the team uncovers allegations of wrongdoing or misconduct related to the implementation of the programme/project - whether by UNDP personnel, implementing partners, contractors, or other stakeholders - the evaluation team must immediately report these allegations through UNDP's established disclosure channels.

Step 1: Data analysis

Who	Role
Evaluation Team	Conducts evaluative analysis aligned with the inception report methodology and maintains data privacy protocols in line with ethical guidelines.
Evaluation Manager / M&E Focal Point	Oversees consistency of analysis with the approved TOR, inception report, and quality standards, and facilitates access to data and clarifications.
Programme / Project Manager	Provides factual clarifications and supplementary data, when requested.

Data analysis should follow on from the approved methodology in the inception report, and be rigorous, transparent, and replicable. During this phase, evaluators:

- Systematically organize, code, and analyse qualitative and quantitative data to address each evaluation question;
- Use multiple data sources (e.g., documents, interviews, surveys, and observations) to validate findings (known as triangulation);
- Explicitly identify patterns, causal pathways, and contributing factors;
- Assess both intended and unintended results, as well as contextual and cross-cutting dimensions such as gender equality, human rights, and disability inclusion; and
- Clearly document analytical steps, data limitations, and mitigation measures to strengthen credibility and transparency.

Step 2: Draft report

Who	Role
Evaluation Team	Drafts report, following evaluation report template (Tool 5.1) and in line with UNDP quality standards and the timeline agreed in the contract. Responds to each comment in the audit trail (Tool 5.3 - Audit trail template).
Evaluation Manager / M&E Focal Point	Ensures accuracy and compliance with reporting structure and UNDP evaluation quality standards. Oversees the review process with UNDP stakeholders and the evaluation reference group, including the transparency and traceability of comments.
Evaluation Commissioner	Approves draft report for external stakeholder review.
Evaluation Reference Group (if applicable)	Provides technical and contextual comments on draft report through evaluation manager.
Programme / Project Manager	Reviews and comments on factual accuracy and completeness of programme data.

Drafting the report

The evaluation manager is ultimately responsible for the quality of the evaluation and should ensure that it meets the information needs of the organization. They should share the UNEG Norms and Standards and the UNDP evaluation report template with the evaluation team. The evaluation manager should also inform the evaluation team about the UNDP Evaluation Quality Assurance System. This information should be shared again even when it has been shared earlier in the evaluation process to reinforce adherence to the template.

Additional requirements for VF-specific evaluations can be found in the Vertical Funds Evaluation Guidelines.

The minimum contents of an evaluation report include:

1. Title and opening pages with details of the project/programme/outcome being evaluated (evaluand) and the evaluation team.
2. Project and evaluation details, including the project title, Quantum number, budgets, dates, and other key information.
3. Table of contents.
4. List of acronyms and abbreviations.
5. Executive summary (a stand-alone section of a maximum of four pages including the quality standards and assurance ratings).
6. Introduction and overview, explaining what is being evaluated and why.
7. Description of the intervention being evaluated, providing a basis for readers to understand the design, general logic, results framework (theory of change), and other information relevant to the initiative being evaluated.
8. Evaluation scope and objectives, providing a clear explanation of these and the main questions.
9. Evaluation approach and methods, describing in detail the selected methodological approaches and methods.
10. A disclosure of AI use during the evaluation process and as presented in the final evaluation report.
11. Data analysis, describing the procedures used to analyse the data collected to answer the evaluation questions.
12. Findings and conclusions, setting out the evaluation findings, based on analysis of the data collected, and the conclusions drawn from these findings.
13. Recommendations, of which there should be a reasonable number of practical, feasible ones directed to the intended users of the report about what actions to take or decisions to make.
14. As appropriate and when requested in the TOR, the report should include a discussion of lessons learned from the evaluation of the intervention.
15. All findings, conclusions, recommendations, and lessons learned need to consider gender equality and women's empowerment, disability, and other cross-cutting issues.

16. Annexes. At a minimum, these should include:

   a. TOR for the evaluation.

   b. Evaluation matrix and data collection instruments.

   c. List of groups interviewed or consulted, identified in aggregate and without identification of individual names, jobs, and contact information.

   d. List of sites visited.

   e. List of supporting documents reviewed.

Once the first draft of the evaluation report has been submitted, the evaluation manager and evaluation reference group should assure the quality of the report and provide comments. The evaluation report should be logically structured, contain evidence-based findings, conclusions, and lessons and recommendations, and must be presented in a way that makes the information accessible and comprehensible. A quality evaluation report should:

- Be easy to follow;
- Use appropriate and transparent methodology;
- Explain what is being evaluated and clearly reflect the theory of change for the programme/project;
- Cover the methods and findings addressing gender equality and LNOB considerations;
- Include findings acquired from credible evidence and in line with clear conclusions; and
- Provide realistic and actionable recommendations that are linked to the findings.

Ethics in evaluation: report drafting stage

AI is increasingly being used to support evaluations. While it is a useful tool and one that can help with many of the rote and systematic areas of work within the evaluation process, it is especially important that evaluation teams are clear within the report on how they have used and integrated work carried out with the help of AI. Details on disclosure of AI use in the evaluation can be found alongside other considerations around AI use in the UNEG Ethical Principles for Harnessing AI in United Nations Evaluations.

Confidentiality and anonymity are key ethical principles that require specific attention in the drafting of an evaluation report. Reports should be written in aggregate, meaning that findings and details within them should not be attributed to any one source. This ensures not only fully triangulated findings to support evaluation judgements, but it also ensures that stakeholder and informant information is protected. The identification of interviewed and consulted stakeholders should be recorded with care and kept internal to the organization. Published documents should only mention groups and their members in aggregate in the annexes. Interview coding systems and separation of information collected from interviewed sources is a way of ensuring confidentiality of information throughout the evaluation process.

Good practice when writing recommendations

While each evaluation is different, all evaluation teams should prioritize recommendations and provide an appropriate number that can be realistically undertaken by the programme/project. Recommendations should, to the greatest extent possible, be addressed to stakeholders within UNDP. If issues require changes from those outside UNDP, recommendations should indicate how UNDP can support these changes with or through the stakeholders.

Clarity of recommendations is key to appropriate action and change for better programming. The following elements can help to increase the use of recommendations:

- Limiting the number of recommendations to maintain focus;
- Clustering and/or sequencing recommendations (e.g. by time, priority, or focus);
- Having clear and specific recommended actions, including to whom they are addressed and who has responsibility for follow-up actions; and
- Specifying the rationale for each recommendation, including to which issues in the evaluation the recommendation is addressing.

Recommendations should be aligned with a realistic understanding of the operating context and situation. Where recommendations require budgetary or other resources to be carried out, the evaluation team should recalibrate them according to what is feasible.

Step 3: Draft evaluation report review

Upon submission of the first draft report, the evaluation manager coordinates the review process with relevant UNDP stakeholders, including the evaluation reference group (if applicable). Comments should focus on the clarity, accuracy, evidence base, and usability of the report. They should follow UNEG guidance, including relevant UNEG Norms and Standards and UNEG Ethical Guidelines for Evaluation, and UNEG Ethical Principles for Harnessing AI in United Nations Evaluations.

The evaluation manager should consolidate all feedback into an evaluation audit trail, using the audit trail template (Tool 5.3) to ensure transparency and accountability in the review process.

Steps to be taken:

- Gather comments from UNDP, the evaluation reference group (if applicable), and key stakeholders. If supporting data are required, UNDP staff should include access issues within their comments.
- Record all feedback in the audit trail template.
- Share the consolidated file with the evaluation team within the agreed timeline.
- The evaluation team responds to each comment, explaining revisions or justifying non-action.

- When there is a challenge to the findings in the report, the evaluation team may request additional data and information.
- If disagreements persist, these are noted in the audit trail; UNDP may then document its differing view in the management response, rather than by editing the report.

The audit trail must remain an internal document and not be uploaded to the ERC.

Ethics in evaluation: report review stage

Where there are discrepancies about the report's findings and/or factual inaccuracies in the data, the evaluation manager should ensure that the programme/project team provides the necessary additional evidence to make changes within the report. If evidence is not available to substantiate such changes, the evaluator's interpretation must stand. New evidence submitted after the report is written must be triangulated, and the usability of the additional data should be left at the evaluator's discretion.

Report quality assurance

Who	Role
Evaluation Manager / M&E Focal Point	Oversee the quality assurance process using UNDP QA checklists (Tool 5.2 – Evaluation report quality checklist) and confirms compliance with UNDP and UNEG norms and standards.
Regional Evaluation Focal Point	Ensures alignment with corporate quality standards and SWAP EPI requirements.
Evaluation Commissioner	Approves final report upon satisfactory quality assurance undertaken by the evaluation manager, and authorizes submission to the ERC.
IEO	Conducts quality assessment of all final decentralized evaluations uploaded to the ERC.

The evaluation manager is responsible for ensuring that the evaluation report meets UNDP quality standards, as indicated in the UNDP Evaluation Quality Assessment System.

To facilitate internal quality assurance, the evaluation manager should complete the evaluation report quality checklist (Tool 5.2) to confirm that all standards have been met prior to approval of the report.

Step 4: Dispute settlement and reporting wrongdoing

Who	Role
Evaluation Team	Records disagreements in the audit trail and provides judgement on the credibility of new evidence to support change requests.
Evaluation Manager	Oversees requests for changes to the evaluation's findings and conclusions, and facilitates the provision of new evidence to the evaluation when introduced. If requests for changes are unresolved, escalates the matter to the Regional Deputy Director with the IEO copied in (evaluation.office@undp.org).
Regional Evaluation Focal Point	Mediates disputes and provides written resolutions.
IEO	Provides independent technical opinion when sought by the regional office and monitors disputes for adherence to independence principles.
Office of Audit and Investigation (OAI)	Receives reports of wrongdoing via hotline, online form, or email and ensures confidentiality and due process.
Evaluation Commissioner	Ensures staff are aware of reporting channels and protected from retaliation.

Dispute settlement

Disputes between evaluators and those being evaluated are not uncommon. The audit trail provides an avenue through which to highlight issues within an evaluation, and offers opportunities for programme/project teams to submit new evidence to provide factual justification for changes. It also gives the evaluator the chance to provide further evidence to support their findings.

If there is a continued disagreement, then either the programme/project team or the evaluation team can address their concerns to the Deputy Director of the region where the evaluation is being undertaken, including the IEO in their correspondence (evaluation.office@undp.org). The Regional Deputy Director will ensure a timely response.

More details on individual contractor dispute settlement can be found in the UNDP POPP.

Reporting wrongdoing

UNDP takes all reports of alleged wrongdoing seriously. In accordance with the UNDP Legal Framework for Addressing Non-Compliance with UN Standards of Conduct, the Office of Audit and Investigation (OAI) is the principal channel for submitting allegations.

Anyone with information regarding alleged fraud or cases of sexual exploitation and abuse against UNDP programmes or involving UNDP staff is strongly encouraged to report this information through the Investigations Hotline (+1-844-595-5206).

People reporting wrongdoing to the Investigations Hotline have the option to leave relevant contact information or to remain anonymous. However, allegations of workplace harassment and abuse of authority cannot be reported anonymously.

The Investigations Hotline, managed by an independent service provider on behalf of UNDP to protect confidentiality, can be directly accessed worldwide and free of charge in the following different ways:

1. ONLINE REFERRAL FORM *(you will* be *redirected to an independent third-party site*)

2. PHONE - REVERSED CHARGES - Click here for worldwide numbers (interpreters available 24 hours/day) Call +1-844-595-5206 in the USA

3. EMAIL directly to OAI at: reportmisconduct@undp.org

4. REGULAR MAIL

   Deputy Director (Investigations)
   Office of Audit and Investigations
   United Nations Development Programme
   One UN Plaza, DC1, 4th Floor
   New York, NY 10017 USA

Step 6.

Finalize and respond



Once an evaluation is complete, the focus shifts to ensuring its findings are used, its lessons are shared, and follow-up actions are implemented and tracked.

Deliverables: *Management response; entry in the ERC*

Tool:

- *Tool 6.1 Management response template*

Indicative Timeframe: *6-12 weeks for initial entry into the ERC, then up to 5 years for follow-up and completion of the actions recommended in the evaluation report.*

Step 1: Prepare the management response

Who	Role
Evaluation Commissioner	Confirms and approves evaluation management response.
Programme / Project Manager	Drafts management response for recommendations addressed to their unit based on the management response template (Tool 6.1). Proposes realistic key actions, timelines, and responsible parties for all high-level and sub-level recommendations. Ensures actions are budgeted and integrated into workplans.
Evaluation Commissioner	Reviews and approves the final management response that fully addresses recommendations. Remains accountable for the implementation of the action plan associated with evaluation management responses.
Evaluation Manager / M&E Focal Point	Leads and coordinates the management response and oversees the process for every evaluation. Ensures evaluation is uploaded into the ERC within six weeks of finalization.
Regional Evaluation Focal Point	Oversees follow-up and tracking of agreed key actions in the ERC, ensuring periodic updates and provision of appropriate evidence of implementation where applicable.
IEO	Reports corporate-level compliance and implementation to Executive Board annually.
BPPS	Monitors completeness and timeliness of management responses in the ERC.

A management response is UNDP's formal response to each evaluation recommendation. This response is always approved by senior management, who is accountable for its implementation.

For every recommendation, senior management indicates whether it is accepted, partially accepted, or not accepted. Where a recommendation is accepted, UNDP management commits to a clear action plan outlining how the identified issue will be addressed.

Why management responses matter

All UNDP evaluations require a management response for each recommendation. The purpose of the management response is to:

- Ensure that recommendations are reviewed, understood, and acted upon by the responsible UNDP unit;
- Provide a formal commitment on how identified issues will be addressed, by whom, and within what timeline;
- Enable tracking of actions through the ERC, so that follow-up is visible and accountable; and

- Strengthen programming by ensuring that evaluation recommendations lead to specific improvements in the way UNDP designs and implements its work.

The management response is a practical mechanism that helps UNDP teams identify issues raised by the evaluation and specify how they will be corrected and monitored.

Content and structure

Each management response must:

1. State the management position on each recommendation (fully accepted, partially accepted, or rejected) with justification:
    - **Fully accepted:** Management agrees entirely and provides the rationale for its agreement. Management commits to action.
    - **Partially accepted:** Management agrees in part: for areas of agreement, it must provide reasoning; for areas of disagreement it must provide justification. Management commits to action that aligns with partial acceptance.
    - **Rejected:** Management provides a clear rationale for not accepting or pursuing the recommendation. Management does not commit to action.
2. For every fully or partially accepted recommendation, specify key follow-up actions, indicating:
    - The nature of the action and how it addresses the recommendation;
    - The responsible actor/group accountable for implementation;
    - A realistic timeline, ideally within 18 months of report completion;
    - Performance indicators or deliverables to facilitate tracking - indicators must be SMART (specific, measurable, achievable, relevant, and time-bound); and
    - Evidence for each key action completed or in progress.
3. For every rejected recommendation, the programme/project team provides a justification that needs to be reviewed and approved by the evaluation manager. Relevant stakeholders are informed of the reasons and justifications for the recommendation being rejected or deemed no longer applicable.

    In exceptional circumstances, and as needed, the evaluation manager can elevate this discussion to the Regional Evaluation Focal Point for their technical advice on how to proceed.

Good practice

A good management response and action plan should follow SMART principles and thus:

- Be specific, measurable, achievable, relevant, and time-bound, avoiding generic or open-ended commitments;

- Take into consideration the whole recommendation, not only the headline statement;
- Explain clearly, with explicit judgement and justification, why and which parts of the recommendation have been accepted, partially accepted, or rejected;
- Make resourcing assumptions for planned actions explicit, provide a clear commitment to address issues identified by the evaluation and integrate necessary adjustments, identify clear UNDP responsibility centres, and be addressed to the appropriate UNDP stakeholder(s);
- Be discussed in project or programme boards, ensuring alignment with broader UNDP and partner strategies; and
- Upload evidence to the ERC for each key action completed or in progress.

Joint evaluations

For joint or inter-agency evaluations, UNDP prepares a management response only for recommendations directed to UNDP. Coordination with partner agencies is encouraged to ensure consistency and collective learning, but accountability for follow-up remains within each agency's respective scope.

Monitor implementation of key actions

Country office M&E focal points are responsible for monitoring the implementation of key actions and reporting on achievements through the ERC.

The ERC can be used to monitor implementation of management responses and key action commitments.

Monitoring of key actions should be undertaken by country office M&E focal points on a quarterly basis, and the ERC should be updated accordingly, with supporting documentation where applicable.

Regional bureaux oversee and monitor implementation and follow-up with UNDP programme and project teams. The IEO reports on the number of management responses and key actions completed, initiated, overdue, or considered no longer applicable in its Annual Report on Evaluation. The BPPS monitors the completeness and timeliness of management responses in the ERC.

Step 2: Publish the report and management response

Who	Role
Evaluation Commissioner	Confirms that public disclosure aligns with transparency policy
Evaluation Manager / M&E Focal Point (ERC Focal Point)	Uploads publication to the ERC and confirms completeness and finalization of uploaded documentation, including the evaluation report, recommendations, management response, and key actions
Regional Evaluation Focal Point	Validates uploaded documents for completeness and correctness
IEO	Reports on completion # against planned Follows up with M&E focal points if ERC entries are incomplete

Publishing evaluation reports and management responses ensures transparency, accountability, and accessibility of UNDP's evaluative evidence. All finalized evaluations and related management responses must be uploaded to the ERC, which serves as the public repository of UNDP evaluations worldwide.

Ethics in evaluation: publication

In its commitment to transparency and accountability to its stakeholders, UNDP country offices are required to publish all evaluation reports within six weeks of completion. Transparency about what works, what does not, and why allows stakeholders to understand decisions, evidence, and results, and reinforces accountability for the use of resources. This is especially critical when public resources and people's lives are impacted by programme/project activities. Publishing evaluation reports reduces the risk of selective reporting, protecting the evaluator's independence and supporting broader learning and continuous improvement.

In some special cases, the publication of reports may jeopardize the safety and security of staff and beneficiaries, and/or cause significant reputational risks that may impact relations with partners. Additionally, some evaluations undertaken on behalf of donors may require internal approval before being published externally by UNDP.

The evaluation manager, in alignment with decisions taken to not publish at the senior management level, may wish to request to the IEO that an evaluation report not be published. For all evaluations, UNDP must ensure adherence to the "do no harm" principle, even when an evaluation has concluded.

Non-publication and special cases

When an evaluation report has the potential to jeopardize the "do no harm" principle, the evaluation manager may seek to not publish the report. In these cases, regional offices will work with the evaluation manager to understand and articulate the justification to the IEO, who will determine the final recommendation as an independent observer.

When an evaluation is included in planned evaluations but is not listed in its published form when completed, the evaluation manager must include a note in the ERC that identifies the reason for non-publication. The IEO will periodically check the ERC and follow up on the status of those identified as having non-publication status.

The ERC

The ERC is a corporate, web-based platform that hosts:

- All evaluation plans, TORs, final reports, and management responses and key actions evidence for UNDP evaluations.
- A public-facing interface for transparency and learning including guidelines, methodology, and the Artificial Intelligence for Development Analytics platform.
- An internal workspace for management tracking (evaluation plans, progress, management responses, and key actions).

Access to the ERC is granted through UNDP credentials. Each country and regional office designates one or more ERC focal point with appropriate permissions to upload and edit entries. Typically, the M&E Focal Point manages this access and supports the evaluation manager during upload.

What to upload

Document	Deadline	Notes
Final TOR	Within 2 weeks of TOR approval.	Uploaded as a PDF.
Final Evaluation Report	Within 2 weeks of final clearance.	Include annexes either embedded or as attachments.
Evaluation recommendations	Within 2 weeks of final clearance.	Each recommendation must be entered individually (not as a block).
Management responses and key actions evidence	Within 6 weeks of report finalization.	Enter corresponding management responses and key actions individually, linked to each recommendation.

Upload requirements

All uploads to the ERC must be final, approved, and free of track changes or comments. Reports must be in PDF format only, and comprise a maximum of 200 pages including annexes. Only the documents listed above are required.

Do not upload:

- Drafts or working versions;
- Audit trails; or
- Inception reports (these are internal working documents).

Step 3: Using evaluations

It would be misleading to view evaluation as a product that is only used once the report is finalized. Evaluation can be equally influential throughout the process by fostering reflection, learning, and course correction, and after completion when findings, conclusions, and recommendations are shared, discussed, and translated into action. There are two ways to effectively promote the use of decentralized evaluations:

1. Country offices can invest in stakeholder engagement throughout the evaluation, creating structured opportunities for dialogue where relevant actors can reflect on emerging findings, discuss implications, and jointly consider the best way forward. These exchanges - complementing formal management responses - help build shared understanding, brings practical constraints to the surface, and identify what it will take to implement recommendations in practice.

2. The value of evaluation can be enhanced by going beyond a single intervention. Looking across programmes, contexts, and time allows evaluations to identify patterns, generate transferable lessons, and inform policies and strategies at a higher level. This is most effective when evaluations clearly articulate the assumptions, drivers, and contextual factors behind success or failure, enabling others to understand under what conditions lessons can be applied or should be adapted.

To facilitate greater learning and opportunities to use evaluation evidence in programming work, the AIDA platform brings together UNDP's evaluation findings, conclusions, and recommendations into immediate, easy-to-consume information by all.

Country offices can search and identify best practices and approaches for learning, amongst many other subjects in a matter of minutes in English, Spanish, or French. Country offices can summarize, analyse, and identify learning from their own country's work, and from those across UNDP globally. Decentralized evaluations are vital to the platform's information bank, and AIDA allows country offices to make sense of them in a strategic way.

Examples of some of the queries that could be answered by the information available from AIDA include:

1. What interventions have been most effective in improving biodiversity in Asia and the Pacific?
2. What governance reforms have made a difference in reducing poverty?
3. What type of programming modalities have been the most efficient in Zimbabwe?

Step 7.

Quality assessment



This section describes the IEO system for quality assessment of decentralized evaluations, including the purpose, roles and responsibilities, processes, and tools.

It explains the SWAP EPI on Gender Equality and the Empowerment of Women, and how independent and decentralized evaluations are assessed to provide UNDP data for this indicator. It explains UNDP's requirements for reporting on the implementation of the United Nations Disability Inclusion Strategy.

- *Tool 7.1 Quality assessment questions*

Overview of the IEO quality assessment system

The IEO manages a quality assessment system for decentralized evaluations, providing feedback on performance to UNDP regional and headquarter bureaux and country offices, reported annually to the Executive Board.

Once an evaluation is uploaded to the ERC, the IEO assigns a quality review of the report. The quality assessment uses the same UNDP quality standards applied for quality assurance throughout the evaluation process. The IEO quality assessment provides the report with a rating from zero to six, with six indicating a report of outstanding quality. Details on evaluation quality ratings for each country office are available in the ERC, which can be used as examples of good practice and learning.

The objectives of this system are to make sure evaluations are useful, credible, gender-responsive, and comparable across the organization, and to generate data for UNDP's annual reporting to the Executive Board and for the UN-SWAP EPI on Gender Equality and the Empowerment of Women gender indicator. Specifically, the system:

- Confirms that UNDP country and regional offices are commissioning and uploading the evaluations in their plans;
- Independently rates the quality of each report (design, evidence, analysis, recommendations, and cross-cutting issues);
- Gives feedback to the commissioning unit through the ERC;
- Aggregates results by country and region, and globally;
- Provides quality assessment for GEF projects and supports technical advice for UNV, UNCDF, and UNOSSC; and
- Scores each evaluation against the SWAP EPI, so UNDP can report on how well evaluations integrate gender equality, the LNOB principle, and disability.

Purpose and scope

The quality assessment process evaluates the quality, credibility, and utility of each evaluation report. It examines:

- Design quality including clarity, logic, evaluability, and adherence to TORs;
- Credibility of evidence and the robustness of analysis; and
- Usefulness and feasibility of conclusions and recommendations.

For GEF evaluations, the quality assessment also validates the ratings of project achievements and expected outcomes.

The quality assessment system contributes to:

- Strengthening the quality of evaluative evidence across UNDP;
- Supporting accountability to the Executive Board and management;
- Ensuring consistency in reporting standards and methodological rigor;
- Providing timely feedback that can improve future evaluations; and
- Generating corporate lessons for the Annual Report on Evaluation.

The IEO conducts quality assessment for all decentralized evaluations undertaken by UNDP, UNOSSC, UNV, and UNCDF. Country and regional offices should use the feedback to strengthen evaluation design, management, and use.

Quality assessment for GEF evaluations

The scope of analysis for GEF terminal evaluation reports is broader than other UNDP evaluation reports. GEF analysis includes an assessment of project documentation (e.g. project objectives and programme/project planning and implementation) and an analysis of the validity of an evaluation's findings and conclusions. The ERC contains the same information for GEF project terminal evaluations. However, to further validate the terminal evaluation ratings for project implementation, the BPPS provides additional information to the quality assessment via the IEO. These documents are not available on the ERC at present. Additional documentation includes:

- *Project concept notes (project identification forms (PIFs);*
- *CEO endorsement requests;*
- *Project documents, including results framework;*
- *Annual project implementation reports (PIRs);*
- *Core indicators and tracking tools (as available).*

More information on the quality assessment of GEF, GCF, and AF evaluations can be found in their linked guidelines.

Quality assessment process

Step 1: Posting to the ERC

Units upload the final TOR, final evaluation report, and annexes within two weeks of completion. Only final documents may be posted, as the ERC is public. The management response and key actions are uploaded within six weeks of completion.

Step 2: Quality review

The IEO assigns the report for automatic quality review, followed by the IEO's own internal quality control.

Step 3: Feedback

The IEO reviews and finalizes the quality assessment results and makes them available to the business unit through the ERC.

Who	Role
IEO	Leads quality assessment and reporting, provides timely feedback, and maintains quality assessment tools and systems.
Regional Bureau	Monitors evaluation quality trends and supports country offices where quality assessment results highlight persistent weaknesses. Works with offices exhibiting low performance to address structural issues in evaluation design and implementation.
BPPS	Supports regional bureaux in resolving recurring evaluation quality challenges.

Reporting

The quality assessment process is regularly undertaken throughout the year. A quality assessment report for an individual evaluation is available as soon as the IEO performs quality checks on the assessment. Results at the global, regional, and country office levels are available through the ERC.

The IEO reports 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.

Quality assessment criteria and weighting

The quality assessment tool comprises 27 questions grouped into four weighted sections, as follows:

- TOR and evaluation design (5 questions, 15%)
- Report structure, methodology, and data sources (6 questions, 30%)
- Cross-cutting issues (5 questions, 15%)
- Findings, conclusions, and recommendations (11 questions, 40%)

Management response quality is reviewed but not scored.

For GEF terminal evaluations, an additional section validates the project's performance ratings.

Quality assessment questions are available in the quality assessment questions (Tool 7.1).

Rating scale

Each question is scored on a six-point scale. A report may score 4-6 on individual questions, but only a rating of 5 or above indicates that all core UNDP evaluation requirements have been met.

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

Supporting documentation

IEO reviewers access all relevant materials through the ERC, including:

- Evaluation TOR and final report (principal quality assessment documents);
- Project/evaluation metadata;
- Lessons and findings;
- Recommendations; and
- Management response and key actions.

UN-SWAP assessment

The United Nations System-wide Policy on Gender Equality and the Empowerment of Women (GEEW) was endorsed by the Chief Executives Board for Coordination in October 2006 as a means of furthering

the goals 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 and includes an EPI. In 2020, the UN-SWAP was updated. The UN-SWAP EPI reporting follows the UN-SWAP EPI Technical Notes published by the UNEG.

UNDP is required to report against the EPI annually, assessing both independent evaluations and decentralized evaluations. Detailed information on the EPI is available here.

EPI criteria and scorecard

A scorecard is always used to assess evaluation reports against the first three criteria below, while a fourth criterion applies at the agency level:

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.

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

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 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 the requirement

4-6 points = Approaching the requirement

7-9 points = Meets the requirement

The assessment process

The IEO is the focal point for the EPI.

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. A reviewer assesses the SWAP EPI for all IEO-led 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 under the "exceeds the requirement" category for the first time.

Box 1: Sample of evaluations that have met EPI requirements.

- Costa Rica (2024), Terminal Evaluation: Seventh Operational Phase of the GEF Small Grants Programme in Costa Rica.
- Indonesia (2024), Terminal Evaluation: Combating Illegal and Unsustainable Trade in Endangered Species in Indonesia.
- 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.
- Chad (2023), Final Evaluation: Projet de prévention de la féminisation des modes opératoires des groupes extrémistes au Tchad.
- Ukraine (2023), Final Evaluation: Digital, Inclusive, Accessible: Support to Digitalization of Public Services in Ukraine .

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.

Reporting on EPI

The IEO prepares a final synthesis report for UNDP, which is uploaded to the UN-SWAP GEEW 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. The 2020 report can be accessed at: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N20/092/53/pdf/N2009253.pdf?OpenElement

Disability and Inclusion Strategy

The IEO carries out yearly reporting on UNDP's work to support inclusion for people living with disabilities. It is responsible for providing details on the extent to which disability inclusion is considered throughout evaluation, including through the TORs, inception reports, and evaluation reports.

The evaluation indicator (EPI) assesses the extent to which UN entities integrate disability inclusion into all phases of the evaluation process and every type of evaluation it conducts. This indicator is part of the United Nations Disability Inclusion Strategy (UNDIS) entity accountability framework's 15 performance indicators and is specifically aimed at ensuring that evaluation systems meaningfully incorporate disability rights, perspectives, and data.

UNDIS Indicators for Evaluation

The following indicators are based on a three-tiered measurement scale:

1. Approaching the requirement;
2. Meets the requirement;
3. Exceeds the requirement.

According to the UNEG Guidance on Disability and Inclusion, the UNDIS evaluation indicator requires entities like UNDP to demonstrate how disability inclusion is systematically integrated throughout their evaluation processes including ensuring that:

- Evaluation guidelines provide clear direction on addressing disability inclusion;
- TOR, methods, and evaluation questions explicitly reflect disability-related considerations;
- Data collection involves people living with disabilities and organizations serving persons with disabilities; and
- Findings, conclusions, and recommendations meaningfully incorporate evidence on disability inclusion.

To meet and exceed the requirement, UNDP must not only apply these standards consistently across all evaluations, but also conduct periodic meta-analysis of disability-related evaluation evidence to strengthen learning and accountability.

Independent Evaluation Office
United Nations Development Programme
One UN Plaza, DC1-20th Floor
New York, NY 10017, USA
Tel. +1 (646) 781 4200

/www.undp.org/evaluation
/UNDP_Evaluation
/evaluationoffice
/company/undp-ieo