

UNDP EVALUATION GUIDELINES

Revised edition: June 2021



IEO | INDEPENDENT
Evaluation Office
United Nations Development Programme



UNDP EVALUATION GUIDELINES © IEO UNDP June 2021

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

Suggested citation: UNDP Evaluation Guidelines, Independent Evaluation Office of UNDP, New York, June 2021.

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

Web: <http://www.undp.org/evaluation>

Connect with us:



/UNDP_Evaluation



/ieoundp



/evaluationoffice



FOREWORD

It gives me great pleasure to introduce the UNDP Evaluation Guidelines in collaboration with the UNDP Independent Evaluation Office.

These guidelines reflect the commitment of UNDP to continuous learning and improvement, and they come at an important moment in time. The ambition of the 2030 Agenda and the Sustainable Development Goals demand a different kind of development: where problems and their solutions are connected and integrated, and where incremental change for good is not enough.

We are mandated to help our partners find and implement integrated, transformational solutions across the Sustainable Development Goals. As we do this in and across our six signature areas of poverty, governance, environment, resilience, energy and gender, we must learn and adapt with openness, transparency and accountability. Strong evaluations, thoughtfully used, are a fundamental part of that process and will be increasingly central to our work.

I would like to commend all those involved in preparing these guidelines, including colleagues across UNDP country, regional and global teams as well as the Independent Evaluation Office. I encourage you all to read them carefully and put them to work to enhance how we think, deliver, invest and manage as the next generation UNDP.

Achim Steiner

Administrator
UNDP



PREFACE

It gives me great pleasure to present the 2021 update of the United Nations Development Programme (UNDP) Evaluation Guidelines, to assist all UNDP colleagues involved in any aspect of evaluation across the organization.

The first UNDP handbook on monitoring and evaluation was developed in 1997, and subsequently revised in 2009 and again in 2019. The UNDP Evaluation Guidelines provide clear guidance and direction for those planning and commissioning evaluations, including step-by-step processes, templates and roles and responsibilities.

Evaluation is essential for UNDP, to ensure that the organization remains accountable, transparent and learns from its programme implementation. IEO believes that strong, quality, and credible evaluations are needed to ensure balanced and informed management decision making at all levels including Country, Regional and Headquarters.

High quality decentralized evaluations also strengthen the work of the Independent Evaluation Office (IEO) to provide evaluative evidence. This informs and enriches our Independent Country Programme Evaluations (ICPE) to strengthen new Country Programme Documents (CPDs), as well as our corporate and thematic evaluations which inform UNDP strategic planning. Recently, IEO has increasingly turned to decentralized evaluation to capture lessons across a broad range of thematic areas, informing the COVID-19 response as well as thematic programming.

IEO remains committed to support UNDP in strengthening and broadening evaluation processes and culture across the organization, and will continue to provide guidance and training to strengthen evaluation planning, implementation and production at all levels.

The completion of these Guidelines was made possible through the collective efforts of the IEO team and UNDP colleagues from headquarters, regional bureaux and country offices who participated in the process in 2019. Their comments, questions and suggestions have been considered and incorporated throughout this update, and will continue to be incorporated in future versions.

I urge UNDP management and colleagues at all levels to incorporate these Evaluation Guidelines into their everyday work, to ensure that our evaluations are of high quality and inform our strategic decision making, for the delivery of more inclusive and sustainable development results.

Oscar A. Garcia

Director,
Independent Evaluation Office, UNDP



ACRONYMS

BPPS	Bureau for Policy and Programme Support	OECD	Organization for Economic Cooperation and Development
CCA	Country Common Assessment	PAC	Programme Appraisal Committee
CPD	Country Programme Document	PBF	Peacebuilding Fund
CPPS	Country Programme Performance Survey	POPP	Programme and operations policies and procedures
D-CPE	Decentralized Country Programme Evaluation	PPM	Project and programme management
DAC	Development Assistance Committee	SDG	Sustainable Development Goal
EPI	Evaluation performance indicator	SWAP	System-wide Action Plan
ERC	Evaluation Resource Center	TOR	Terms of reference
FAQ	Frequently asked questions	UNCT	United Nations country team
GCF	Green Climate Fund	UNDAF	United Nations Development Assistance Framework
GEF	Global Environment Facility	UNDAP	United Nations Development Assistance Partnership
GPN	Global Policy Network	UNDP	United Nations Development Programme
ICPE	Independent Country Programme Evaluation	UNEG	United Nations Evaluation Group
IEO	Independent Evaluation Office	UNSDCF	United Nations Sustainable Development Cooperation Framework
M&E	Monitoring and evaluation		
NEC	National evaluation capacities		
OAI	Office of Audit and Investigation		



INTRODUCTION TO THE UPDATED EVALUATION GUIDELINES

Why update the Evaluation Guidelines?

Welcome to the 2021 updated Evaluation Guidelines. The Guidelines have been updated to reflect feedback from trainings and interviews, and recent changes in UNDP, bringing them into line with the new UNDP Evaluation Policy and the United Nations Sustainable Development Cooperation Framework (UNSDCF).

The following documents are of particular importance for the UNDP evaluation architecture:

- UNDP, 2019, **Revised UNDP Evaluation Policy**.¹ The Evaluation Policy sets out the purpose and basic principles of evaluation and defines the institutional architecture of evaluation for UNDP and its associated funds and programmes.
- UNDP, 2020, **Social and Environmental Standards**.² The revised Standards underpin the UNDP commitment to mainstream social and environmental sustainability.
- UNDP, 2018, **Gender Equality Strategy 2018-2021**.³ The Strategy provides a road map to integrate gender equality into all aspects of UNDP work.
- UNDP, 2018, **Disability Inclusive Development in UNDP. Guidance Note**.⁴ The Guidance Note incorporates a reflection on UNDP comparative advantage and the diversity and depth of interventions undertaken by UNDP.
- United Nations Evaluation Group (UNEG), 2020, **Ethical Guidelines for Evaluation**.⁵ This is a revision of the original document published in 2008.
- UNEG, 2018, **Guidance on Evaluating Institutional Gender Mainstreaming**.⁶ This practical guide was designed to advocate a common approach to assessing progress of institutional gender mainstreaming in the United Nations system.
- UNEG, 2016, **Norms and Standards for Evaluation**.⁷ The UNEG Norms and Standards are a foundational document intended for all United Nations evaluation bodies.
- UNEG, 2014, **UNEG Integrating Human Rights and Gender Equity in Evaluations**.⁸ This is an in-depth handbook designed to serve as a field guide.
- In addition, various **UNEG guidelines**⁹ have been introduced or adjusted in recent years.

¹ Access at: <http://web.undp.org/evaluation/policy.shtml>

² Access at: <https://www.undp.org/content/undp/en/home/librarypage/operations1/undp-social-and-environmental-standards/>

³ Access at: <https://www.undp.org/content/undp/en/home/librarypage/womens-empowerment/undp-gender-equality-strategy-2018-2021.html>

⁴ Access at: https://www.undp.org/content/undp/en/home/librarypage/democratic-governance/human_rights/disability-inclusive-development-in-undp.html

⁵ Access at: <http://www.unevaluation.org/document/detail/2866>

⁶ Access at: <http://www.unevaluation.org/document/detail/2133>

⁷ Access at: <http://unevaluation.org/document/detail/1914>

⁸ Access at: <http://www.unevaluation.org/document/detail/1616>

⁹ Access at: <http://www.unevaluation.org/document/guidance-documents>

- United Nations, 2018, **System-Wide Policy on Gender Equality and the Empowerment of Women (SWAP) Evaluation Performance Indicator. Technical Guidance**.¹⁰ SWAP was developed as a means of furthering the goal of gender equality and women’s empowerment within policies and programmes.
- Organization for Economic Cooperation and Development (OECD) Development Assistance Committee (DAC) Network on Development Evaluation, **Better Criteria for Better Evaluation, 2019, Revised Evaluation Criteria Definitions and Principles for Use**.¹¹ This update includes adapted definitions of the OECD evaluation criteria and reflects new policy priorities, including the 2030 Agenda for Sustainable Development and the Paris Climate Agreement.
- UNDP, 2018, Updated **UNDP programme and operations policies and procedures (POPP)**¹² for **project and programme management (PPM)**.¹³ These guidelines reflect changes resulting from streamlining of the POPPs in 2018.
- United Nations, 2019, **United Nations Sustainable Development Cooperation Framework**.¹⁴ The Cooperation Framework is the new planning and implementation instrument for all United Nations development activities within countries, replacing the United Nations Development Assistance Framework (UNDAF).
- **The Sustainable Development Goals (SDGs), 2030 Agenda and UNDP Strategic Plan, 2018-2021**¹⁵ are also reflected in these Guidelines.

Who are the Evaluation Guidelines for?

The Evaluation Guidelines target a variety of audiences:

- **Programme units**, including headquarters departments and bureaux, regional bureaux and country offices.
- **UNDP staff** in country offices, regional bureaux, regional centres and headquarters. This includes project and programme staff and managers involved in:
 - Planning evaluations;
 - Commissioning evaluations;
 - Managing evaluations;
 - Recruiting evaluators;
 - Using evaluation results.
- **UNDP senior management** who oversee and assure the quality of planning, monitoring and evaluation processes and products, and use monitoring and evaluation (M&E) for decision-making, including resident representatives, deputy resident representatives and outcome, sector or programme managers.

¹⁰ Access at: <http://www.unevaluation.org/document/detail/1452>

¹¹ Access at: <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>

¹² Access at: <https://popp.undp.org/SitePages/POPPRoot.aspx>

¹³ Access at: <https://popp.undp.org/SitePages/POPPBSUnit.aspx?TermID=1c019435-9793-447e-8959-0b32d23bf3d5&Menu=BusinessUnit>

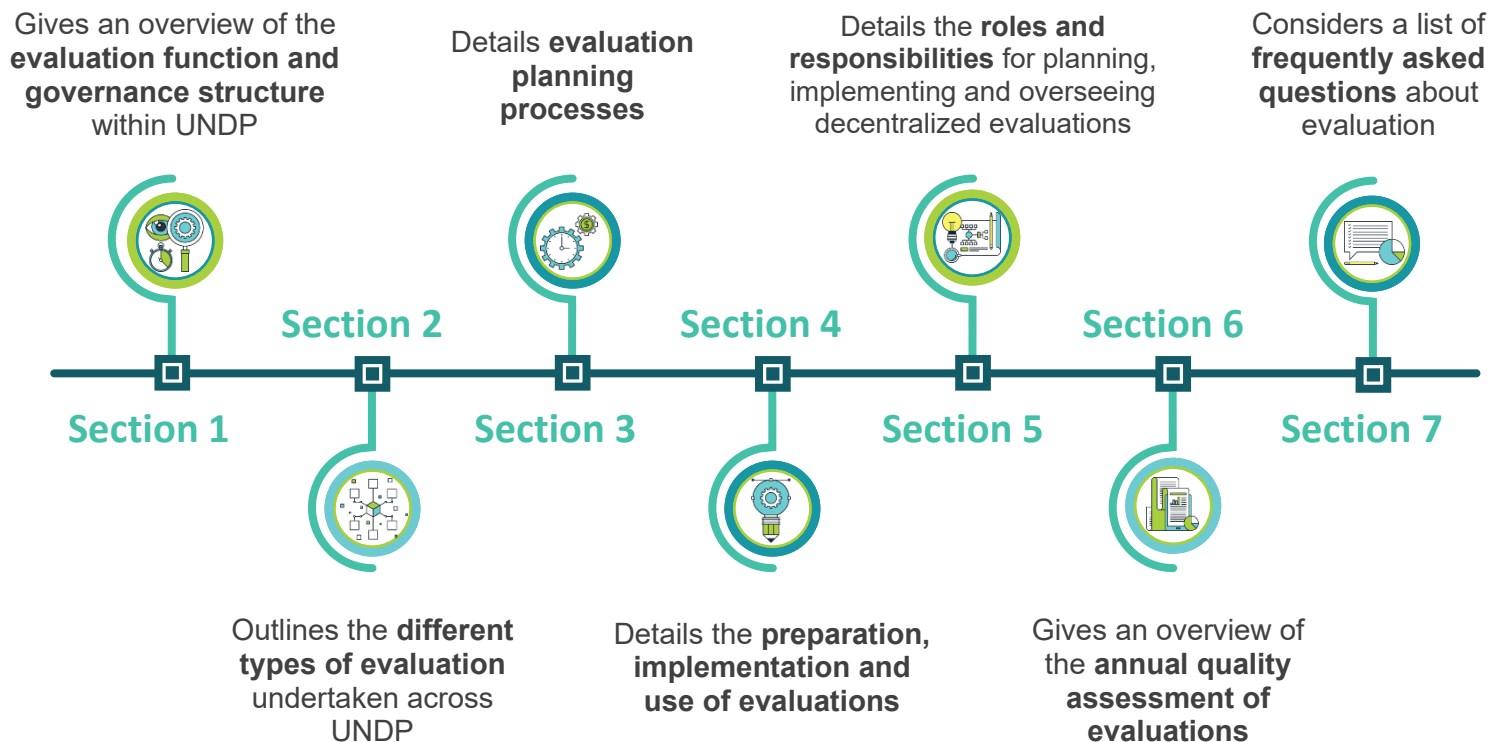
¹⁴ Access at: <https://unsdg.un.org/resources/united-nations-sustainable-development-cooperation-framework-guidance>

¹⁵ Access at: <http://strategicplan.undp.org/>

- **The UNDP Office for Audit and Investigations** can use the Guidelines in its audit function, as they provide detailed procedures on evaluation planning, content and implementation processes.
- **Stakeholders and partners**, such as governments, civil society organizations, the private sector, United Nations and development partners and beneficiaries involved in UNDP planning, monitoring and evaluation processes.
- The **UNDP Executive Board**, which oversees and supports the activities of UNDP, ensuring that the organization remains responsive **to** the evolving needs of programme countries.
- **Evaluators and researchers who** need to understand the guiding principles, standards and processes for evaluation within the UNDP context.

Structure of the updated Evaluation Guidelines

The Evaluation Guidelines are organized in a number of sections, which can be used in sequence or as stand-alone pieces, and are intended to be a living and continuously evolving document.



Sections will be updated regularly considering changing needs for evaluation guidance, as well as changes in UNDP policy and approaches. Additional sections or companion pieces will be developed in future.

Throughout the Evaluation Guidelines are links to other guidance and policies of relevance to evaluation in UNDP.



CONTENTS

THESE GUIDELINES ARE DIVIDED INTO SEVEN SECTIONS:

SECTION 1. THE UNDP EVALUATION FUNCTION



SECTION 2. DECENTRALIZED EVALUATION IN UNDP



SECTION 3. EVALUATION PLAN DEVELOPMENT



SECTION 4. EVALUATION IMPLEMENTATION AND USE



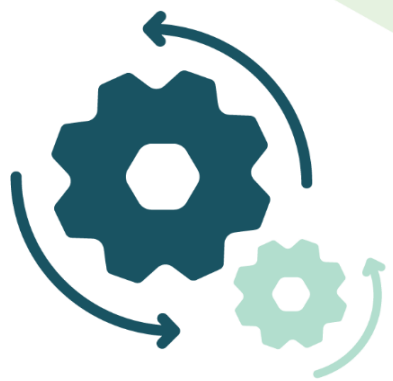
SECTION 5. ROLES AND RESPONSIBILITIES IN DECENTRALIZED EVALUATION



SECTION 6. EVALUATION QUALITY ASSESSMENT



SECTION 7. FREQUENTLY ASKED QUESTIONS (FAQs)

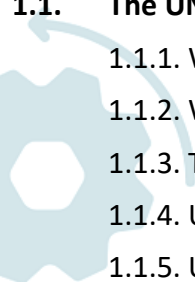


SECTION 1

THE UNDP EVALUATION FUNCTION



CONTENTS



1.1. The UNDP evaluation function	1
1.1.1. What is evaluation?	1
1.1.2. Why does UNDP evaluate?	1
1.1.3. Types of evaluations in UNDP	4
1.1.4. United Nations evaluation principles, norms and standards	5
1.1.5. UNDP evaluation governance structure	8

Figure 1. The UNDP evaluation function	1
---	----------

Figure 2. Examples of other oversight, accountability and assessment functions	3
---	----------

Table 1. Distinctions between monitoring and evaluation	4
--	----------

Table 2. UNEG norms for evaluations	6
--	----------

Table 3. UNEG evaluation standards	7
---	----------

1. THE UNDP EVALUATION FUNCTION

Section 1 introduces the updated UNDP Evaluation Guidelines and describes the UNDP evaluation function.

1.1. The UNDP evaluation function

Section 1 of the Evaluation Guidelines introduces the role of the evaluation function within UNDP.

1.1.1. What is evaluation?

DEFINITION: An evaluation is an assessment, conducted as systematically and impartially as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area or institutional performance. It analyses the level of achievement of both expected and unexpected results, by examining the results chain, processes, contextual factors and causality, using appropriate criteria such as relevance, coherence,¹⁶ effectiveness, efficiency, impact and sustainability. An evaluation should provide credible, useful, evidence-based information that enables the timely incorporation of its findings, recommendations and lessons into the decision-making processes of organizations and stakeholders.¹⁷

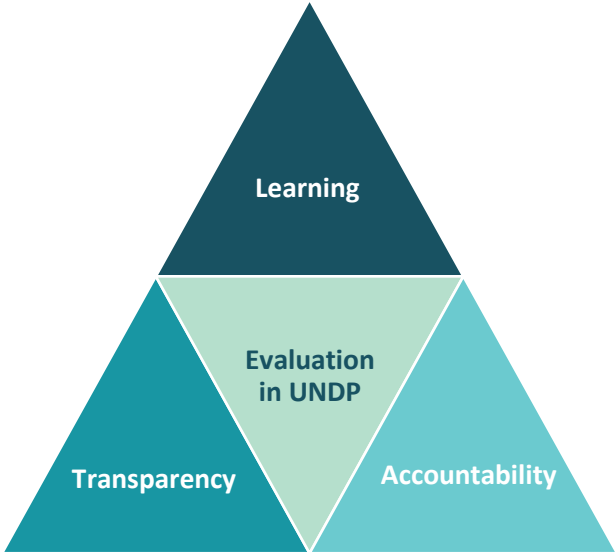


Figure 1. The UNDP evaluation function

1.1.2. Why does UNDP evaluate?

Evaluation is critical for UNDP to progress towards advancing sustainable human development. Evaluations help to ensure that organizational goals and initiatives are aligned with the UNDP Strategic Plan and support the 2030 Agenda and the SDGs, as well as other global, national and corporate priorities. When used effectively, evaluations support programmatic improvements, knowledge generation and accountability.

Evaluation is:

- A means to **strengthen learning** within our organization and among stakeholders, to support better decision-making.

¹⁶ The “Coherence” criterion was introduced by OECD/DAC in 2019. It refers to the compatibility of the interventions in a country, sector or institution. Also see: <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>

¹⁷ UNEG Norms and Standards for Evaluation, <http://www.unevaluation.org/document/detail/1914>

- Essential **for accountability and transparency**, strengthening the ability of stakeholders to hold UNDP accountable for its development contributions.
- Often **intended** to generate empirical knowledge about what has worked, what has not, and why. Through the generation of evidence and objective information, evaluations enable programme managers and other stakeholders to **make informed management decisions** and plan strategically.

Factors supporting effective evaluation

Engaging independent external evaluators is a means to avoid undue influence and bias, ensuring objective and credible evaluation results.

The likelihood of an evaluation being of high quality and useful is increased when the starting point is a good project or programme results framework (theory of change), articulating how activities and outputs are expected to lead to desired outcomes and results. Performance indicators should be SMART: Specific, Measurable, Achievable, Relevant and Time-bound.



All evaluations need to be built on explicit results frameworks and theories of change.

Since the promotion of gender equality, human rights and disability concerns are guiding principles for all United Nations entities, these interrelated issues need to be incorporated into all evaluations. This is clearly addressed in the UNDP Evaluation Policy. Further details will be provided in all relevant sections throughout these Guidelines.



All evaluations need to consider gender mainstreaming, human rights and disability concerns.

How does evaluation fit into broader oversight, accountability and assessment functions?

The United Nations and UNDP have a number of oversight, accountability and assessment tools and functions in addition to evaluation, which have different purposes. Figure 2 highlights some examples.



Figure 2. Examples of other oversight, accountability and assessment functions



There is a clear difference between monitoring and evaluation.

- **Monitoring** provides managers and key stakeholders with regular feedback on the consistency or discrepancy between planned and actual activities and programme performance and results.
- **Evaluation** is an independent judgement based on set criteria and benchmarks.

The importance of monitoring *for* evaluation resides in the availability of relevant and reliable data which can and should be used for evaluation. Table 1 sets out the differences between monitoring and evaluation in a nutshell:¹⁸

¹⁸ Taken from the UNEG Norms and Standards for Evaluation 2016. <http://www.unevaluation.org/document/detail/1914>

Table 1. Distinctions between monitoring and evaluation

 MONITORING	 EVALUATION
Continuous	Periodic
Part of regular management	Independent from management
During implementation	Before, during or after implementation
Process-oriented and focused on progress achieved according to implementation plans	Impact-oriented, objectives assessed at higher levels
Routine data collected as per results matrix	Monitoring data used. Data from multiple sources collected, analysed, interpreted and triangulated
Casual questions are not necessarily raised	Causal questions and theory of change are analysed and questioned
Not independent (internal)	Independent (external)

The programme or implementing unit **M&E plan** is an essential tool identifying the baselines and indicators for which data is to be collected. It also needs to define *when* each type of data will be collected.

1.1.3. Types of evaluations in UNDP

UNDP has a dual evaluation system, with:

1. **Independent evaluations**, such as independent country programme evaluations (ICPEs)¹⁹ and UNDP-wide thematic evaluations, undertaken by the Independent Evaluation Office (IEO) and fully independent of UNDP management and implementing agencies. Independent evaluations inform the decision-making process with credible recommendations, support learning and ensure accountability across the organization.
2. **Decentralized evaluations** undertaken by UNDP programme units such as the Bureau for Policy and Programme Support (BPPS), regional bureaux and country offices to capture lessons learned for future programming and planning, and to ensure accountability. Various types of decentralized evaluations exist in UNDP, but the most common are **project and outcome evaluations**. Programme units do not conduct evaluations themselves, but commission external evaluation consultants to do so.

Although the institutional arrangements (including mandates, accountability lines and operational modalities) of independent and decentralized evaluations are different, they complement and reinforce

¹⁹ Prior to 2018, independent country programme evaluations were known as assessments of development results.

each other. Decentralized evaluations, particularly outcome evaluations, provide relevant information for independent evaluations of country programmes and evaluations of thematic and regional programmes conducted by IEO. In conducting independent evaluations, IEO may carry out country case studies, including reviews of relevant decentralized evaluations, or apply a meta-evaluation approach and draw extensively from country- or region-specific decentralized evaluations. Similarly, decentralized evaluations may draw on the analysis provided in relevant independent evaluations and case studies.

Section two provides definitions of different types of evaluations.

Building national evaluation capacity

Apart from conducting independent and decentralized evaluations, IEO and UNDP support national evaluation capacity, which has also been identified as a programmatic priority in line with General Assembly resolution 69/237.

When appropriately tailored to national circumstances and priorities, the evaluation function can be an effective country-led vehicle for greater citizen accountability. This can accelerate progress towards national SDG priorities, drawing on contributions from indigenous peoples, civil society, the private sector and other stakeholders, including national parliamentarians.²⁰

In order to support governments to assess their national evaluation capacities, IEO developed an **online self-assessment tool**,²¹ which provides a flexible and practical framework to:

- a) facilitate development or strengthening of a national evaluation framework for the SDGs;
- b) respond to existing gaps in national evaluation thinking and practice;
- c) inform country-led evaluation processes and systems; and
- d) respond to countries' preferences for national evaluation diagnostics.

The IEO also organizes biannual National Evaluation Capacity conferences. Conference proceedings are available on the IEO website.²²

1.1.4. United Nations evaluation principles, norms and standards

Evaluations across UNDP and the wider United Nations system, both independent and decentralized, are guided by a set of clear principles, norms and standards developed by UNEG.²³ These evaluation principles adhere to impartiality, credibility and utility, and are interrelated.

The **UNEG Norms and Standards for Evaluation (2016)**²⁴ provide a detailed overarching framework for United Nations organizations in the implementation of evaluations and the evaluation function. The UNDP Evaluation Policy and these Evaluation Guidelines are built on the foundation of this agreed framework.

²⁰ <http://web.undp.org/evaluation/policy.shtml, para 7>

²¹ <http://web.undp.org/evaluation/guidance.shtml#dig>

²² <http://web.undp.org/evaluation/nec/nec.shtml>

²³ UNEG is an interagency professional network that brings together evaluation units from across the United Nations system, including United Nations departments, specialized agencies, funds and programmes, and affiliated organizations. It currently has 50 members and observers. Also, see: <http://www.unevaluation.org/about>

²⁴ UNEG Norms and Standards for Evaluation, 2016, <http://www.unevaluation.org/document/detail/1914>

Table 2. UNEG norms for evaluations

UNEG Norms

Summary explanations



It is the responsibility of evaluation managers and evaluators to uphold and promote the principles and values of the United Nations. This includes respect, promotion and contribution to the goals and targets set out in the 2030 Agenda for Sustainable Development.



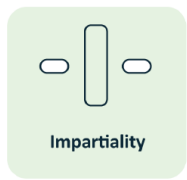
There must be a clear intention to use the evaluation's analysis, conclusions and recommendations to inform decisions and actions. This implies relevant and timely contributions to organizational learning, informed decision-making processes and accountability.



Evaluations must be credible. This requires independence, impartiality and a rigorous methodology. Key elements of credibility include transparent evaluation processes, involving relevant stakeholders and robust quality assurance systems as well as an ethical approach.



Independence allows evaluators to conduct their work without any influence from another party and without any negative effects on their careers. Additionally, the organization's evaluation function needs to be positioned independently from management functions, setting their own agenda and with adequate resources.



Impartiality includes objectivity, professional integrity and absence of bias at all stages of the evaluation process. Evaluators must not have been, or expect to be, directly responsible for the policy-setting, design or management of the evaluation subject.



Evaluation must be conducted with the highest standards of integrity and respect for the beliefs, manners and customs of the social and cultural environment; for human rights and gender equality and for the 'do no harm' principle.



Transparency is an essential element of evaluation that establishes trust and builds confidence, enhances stakeholder ownership, and increases public accountability. Evaluation products should be publicly accessible.



The universally recognized values and principles of human rights and gender equality need to be integrated into all stages of an evaluation., underpinning the commitment to the principle of 'no one left behind'.



Building capacity for the evaluation of development activities at the country level, national evaluation capacities should be supported upon the request of Member States.



Evaluations must be conducted with professionalism and integrity. These are supported by an enabling environment, institutional structures and adequate resources.

UNEG has also outlined **evaluation standards**, which provide a framework for the improvement of all United Nations evaluations functions. Table 3 sets out the five evaluations standards to be applied.

Table 3. UNEG evaluation standards

Standard 1: Institutional framework	1.1	An effective structure of the evaluation function
	1.2	An evaluation policy
	1.3	An evaluation planning and reporting system
	1.4	A management response follow-up mechanism
	1.5	An explicit evaluation disclosure policy
Standard 2: Management of the evaluation function	2.1	A head of evaluation who ensures that evaluation work adheres to norms and standards, secures that the evaluation function is fully operational and duly independent
	2.2	Evaluation guidelines
	2.3	Global leadership, setting standards and oversight, and adapting to new developments
Standard 3: Evaluation competencies	3.1	Individuals engaged in designing, conducting and managing evaluation activities possess the core competencies ²⁵ required for their role in the evaluation process
	3.2	All people engaged in the evaluation process conform to the agreed ethical standards and principles to ensure credibility
Standard 4: Conduct of evaluations	4.1	Evaluations are designed to provide timely, valid and reliable information, which is relevant to the subject being assessed
	4.2	An evaluability assessment is prepared
	4.3	Terms of reference are provided including the evaluation purpose, scope, design and plan
	4.4	The evaluation scope and objectives are identified
	4.5	Evaluation methodologies are sufficiently rigorous
	4.6	Diverse stakeholders are engaged, and reference groups defined
	4.7	The human rights-based approach and gender mainstreaming strategy were incorporated into the design of the evaluation

²⁵ Also, see UNEG Evaluation Competency Framework, <http://www.unevaluation.org/document/detail/1915>

	4.8	The evaluation team is selected through an open and transparent process
	4.9	The final evaluation report is logically structured and contains evidence-based findings, conclusions and recommendations
	4.10	Recommendations should be firmly based on evidence and analysis, clear, results-oriented and realistic in terms of implementation
	4.11	The evaluation function has an effective communication strategy for disseminating evaluation findings and enhancing evaluation use
Standard 5: Quality standards	5.1	There is a framework or system for quality assurance
	5.2	The quality of evaluations is controlled during design stages
	5.3	The quality of evaluations is controlled during final stages

UNDP evaluation policy

Evaluation in UNDP should follow the principles outlined in the 2019 Evaluation Policy,²⁶ which stem from General Assembly resolutions and UNDP Executive Board decisions.

The Evaluation Policy clearly outlines the **roles and responsibilities** for evaluation and its oversight within UNDP.

The Policy clearly calls for a distinction between *evaluation* and *monitoring*, both in function and budget, and establishes a **budget benchmark for evaluation**, separate from monitoring resources (financial and human), for the first time in UNDP.



Under the Evaluation Policy, UNDP aims at “allocating 1 per cent of combined programmatic (core and non-core) resources to the evaluation function on an annual basis, with 0.3 per cent reserved for the work of the Independent Evaluation Office”.²⁷

Furthermore, the Policy states that resources are allocated to evaluation through a series of **evaluation plans**, covering programmes at the country, regional and global levels, as well as through the IEO medium-term evaluation plan.

1.1.5. UNDP evaluation governance structure

This section provides an overview of the roles and responsibilities relating to evaluation, within implementing units. According to the UNDP Evaluation Policy:

- 1. The UNDP Executive Board** “is the custodian of the evaluation policy; annually considers its implementation, and periodically commissions independent reviews of the policy.”²⁸ The Board approves the biennial financial appropriation to IEO, as well as its annual programme of work. The

²⁶ Access at: <http://web.undp.org/evaluation/policy.shtml>

²⁷ <http://web.undp.org/evaluation/policy.shtml>, para 27.

²⁸ <http://web.undp.org/evaluation/policy.shtml>, para 36

IEO submits independent thematic and programmatic evaluations to the Executive Board, which approves or notes the management responses as appropriate.

- 2. The IEO** *“is a functionally independent unit with UNDP that supports the oversight and accountability functions of the Executive Board and the management of UNDP, the United Nations Capital Development Fund and United Nations Volunteers programme. The structural independence of the Office underpins and guarantees its freedom to conduct evaluations and report evaluation results to the Executive Board.”*²⁹

As custodian of the evaluation function, the IEO conducts independent evaluations, sets standards and guidelines, manages the systems for quality assessment and evaluation planning and use through the Evaluation Resource Centre, and develops products to support organizational learning, knowledge management and evaluation capacity development. The IEO also participates in UNEG, which works to strengthen the objectivity, effectiveness and visibility of the evaluation function across the United Nations system.

- 3. The UNDP Administrator** *“(a) safeguards the integrity of the evaluation function, ensuring its independence from operational management and activities; (b) ensures that adequate financial resources are allocated to the evaluation function across the organization, in accordance with the Executive Board-approved financial appropriation for Independent Evaluation Office, and reports to the Board annually on the volume of resources that the organization has invested in evaluation; (c) ensures that the Office has unfettered access to data and information required for the evaluation of UNDP performance; and (d) appoints the Director of the Office in consultation with the Executive Board, taking into account the advice of the Audit and Evaluation Advisory Committee.”*³⁰
- 4. UNDP programme and policy units (headquarters, regional and country offices)** *“commission decentralized evaluations according to evaluation plans that coincide with relevant programmes (regional and country) and global projects. The 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.”*³¹
- 5. The BPPS**, in addition to undertaking its own evaluations, *“coordinates communication between UNDP management and the Independent Evaluation Office and advises regional bureaux on the decentralized evaluation function for UNDP. The Bureau works with the monitoring and evaluation staff of UNDP units to ensure that evaluation plans are properly implemented. Together with the Office, the Bureau provides guidance to UNDP units on the use of evaluation findings and lessons to improve organizational decision-making and accountability and synthesizes evaluation lessons for institutional learning. It also monitors implementation of the management responses to independent evaluations and decentralized evaluations in UNDP.”*³²
- 6. Regional bureaux**, in addition to implementing their own evaluations, support country offices in the development of evaluation plans and implementation of evaluations and oversee implementation of evaluation plans through their appointed evaluation focal points.³³

²⁹ <http://web.undp.org/evaluation/policy.shtml>, para 41

³⁰ <http://web.undp.org/evaluation/policy.shtml>, para 37

³¹ <http://web.undp.org/evaluation/policy.shtml>, para 38

³² <http://web.undp.org/evaluation/policy.shtml>, para 39

³³ Regional bureaux must ensure that there is an M&E focal point, responsible for supporting and overseeing evaluation, based at the regional level. Evaluation focal points should have results-based management, M&E, planning or evaluation capacity.

7. **Bureau and country office senior management** (bureau directors, resident representatives and country directors) are responsible and accountable for the development of units' evaluation plans and ensuring their timely implementation.
8. **The Audit and Evaluation Advisory Committee**, expanded to include evaluation oversight functions, advises the UNDP Administrator on:
 - The Evaluation Policy;
 - Appointment and dismissal of the IEO Director;
 - IEO multi-year and annual workplans, budgets and periodic reports;
 - Thematic and programmatic evaluation reports and management responses;
 - The UNDP decentralized evaluation function and national evaluation capacity programming.³⁴

The Committee also periodically receives and comments on the IEO work programme, and annually appraises the performance of the IEO Director. It further helps to safeguard the Evaluation Policy.

More detail on roles and responsibilities can be found in section five of these Guidelines.

³⁴ <http://web.undp.org/evaluation/policy.shtml>, para 53 and 55



SECTION 2

DECENTRALIZED EVALUATION IN UNDP



CONTENTS

2.1	Introduction	1
2.2	UNSDCF/ UNDAF evaluations	2
2.3	Decentralized country programme evaluations	5
	Midterm D-CPEs	5
	Final D-CPEs	5
2.4	Independent Country Programme Evaluations	6
2.5	Outcome evaluations	7
2.6	Regional programme evaluations	8
2.7	Project evaluations	9
2.8	Global Environment Facility, Adaptation Fund and Green Climate Fund	11
2.9	Multi-country project evaluations	12
2.10	Portfolio evaluations	13
2.11	Thematic evaluations	13
2.12	Impact evaluations	14
2.13	Joint evaluations	15
2.14	Evaluations commissioned by donors or multilateral organizations	17
	Annex 1. Joint evaluation consideration checklist	20
	Table 1. Thresholds for mandatory project evaluations	10
	Table 2. Benefits and possible challenges of joint evaluations	16
	Figure 1. Types of evaluations undertaken by UNDP	1
	Box 1: Gender responsive evaluations	2
	Box 2: OECD/DAC evaluation criteria	2
	Box 3: UNSDCF evaluation guidance	3
	Box 4: Sample UNDAF evaluations	4
	Box 5: Sample decentralized country programme evaluations examples	6
	Box 6: Definition of an outcome	7
	Box 7: Sample outcome evaluations	8
	Box 8: Sample project evaluations	11
	Box 9: GEF evaluation guidelines	11
	Box 10: Sample GEF terminal evaluations	12
	Box 11: Sample thematic evaluations	14
	Box 12: European Commission Directorate evaluation guidance documents	19

2. DECENTRALIZED EVALUATION IN UNDP

Section 2 provides an overview of the different types of decentralized evaluations carried out by UNDP programme units.

2.1 Introduction

UNDP implements a variety of evaluations at different levels, using varied approaches, in line with evaluation plans. Programme units, especially country offices, should ensure that their evaluation plans include a variety of evaluation approaches to capture a broad spectrum of evaluation results during the country programme cycle. This supports accountability and the capture of experience and knowledge to strengthen work within the country, the region, and the wider organization. All UNDP evaluations must adhere to the UNEG Norms and Standards for Evaluation 2016.¹

Figure 1. Types of evaluations undertaken by UNDP



All evaluations conducted or commissioned by UNDP must integrate **human rights, gender equality and disability issues** to meet the requirements of the United Nations System-Wide Action Plan (SWAP) on Gender Equality and the Empowerment of Women Evaluation Performance Indicator,² and the United Nations Disability Inclusion Strategy.³ Integrating gender equality and women's empowerment in the scope, and throughout the terms of reference (TOR), is a critical first step in the evaluation process.

¹ <http://unevaluation.org/document/detail/1914>

² <http://www.uneval.org/document/detail/1452>

³ https://www.un.org/en/content/disabilitystrategy/assets/documentation/UN_Disability_Inclusion_Strategy_english.pdf

Box 1: Gender responsive evaluations

All evaluations should undertake a gender-responsive approach, even for projects that were not gender-responsive in their design. The UNEG guidance document, “Integrating Human Rights and Gender Equality in Evaluations” provides examples of how to incorporate these elements into the purpose, objectives, context and scope of the evaluation, and to incorporate a gender dimension into the standard evaluation criteria.

- UN-SWAP Evaluation Performance Indicator Technical Note, April 2018, <http://www.uneval.org/document/detail/1452>
- UN-SWAP - Individual Evaluation Scoring Tool, <http://www.uneval.org/document/detail/1452>

It is strongly recommended that all evaluations apply the Organization for Economic Cooperation and Development (OECD)/ Development Assistance Committee (DAC) evaluation criteria of relevance, coherence, effectiveness, efficiency and sustainability set out in Box 2 below. In 2019, OECD/DAC revised and further specified their evaluation criteria, published in “Better Criteria for Better Evaluations: Revised Evaluation Criteria Definitions and Principles for Use”,⁴ and added a new criterion, coherence.

Box 2: OECD/DAC evaluation criteria

RELEVANCE: **Is the intervention doing the right things?** The extent to which the intervention objectives and design respond to global and national needs, policies and priorities and those of beneficiaries and partner institutions, and continue to do so as circumstances change.

COHERENCE: **How well does the intervention fit?** The compatibility of the intervention with other interventions in a country, sector or institution.

EFFECTIVENESS: **Is the intervention achieving its objectives?** The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.

EFFICIENCY: **How well are resources being used?** The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.

SUSTAINABILITY: **Will the benefits last?** The extent to which the net benefits of the intervention continue or are likely to continue.

This section includes examples from evaluations that have achieved quality assessment ratings of satisfactory (5) or highly satisfactory (6).

2.2 UNSDCF/ UNDAF evaluations



The evaluation of the Cooperation Framework is an independent system-wide process at country level that contributes to system-level oversight, transparency, accountability and collective learning.

⁴ Access at: <http://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>

The UNSDCF – United Nations Sustainable Development Cooperation Framework - replaced the United Nations Development Assistance Framework (UNDAF) as the United Nations national-level guiding framework, following the Resident Coordinator delinking process in 2018 and 2019.⁵

Purpose and scope

UNSDCF evaluations use the United Nations Country Common Assessment (CCA) report as a benchmark to assess progress towards expected (and unintended) results, and whether the Cooperation Framework made a worthwhile, coherent, durable and cost-efficient contribution to collective United Nations system outcomes and national development processes towards the 2030 Agenda.⁶

UNSDCF evaluations are **undertaken independently** and are vital for ensuring greater transparency on results achieved, promoting joint work and efficiencies, and generating knowledge to inform and improve development programming.

UNSDCF evaluations assess the extent of conformity with **Cooperation Framework Guiding Principles** in terms of both process and results. By identifying synergies, gaps, overlaps and missed opportunities, these evaluations provide the basis for critical inquiry to support the continuous improvement of performance and results. They also play a role in supporting **social and environmental safeguarding efforts**.

Methodology and guidelines

UNSDCF evaluations must adhere to international best practices for evaluation and the UNEG Norms and Standards, and should reflect the evaluation indicators of the Funding Compact,⁷ the United Nations SWAP on Gender Equality and Women Empowerment⁸ and the United Nation Disability Inclusion Strategy.⁹

Box 3: UNSDCF evaluation guidance

Chapter 6 of the UNSDCF Internal Guidance, and chapter 8 of the UNSDCF Cooperation Framework Companion Package and Consolidated Annexes (forthcoming), provide further information for UNSDCF evaluation planning. The Companion Package will include tools and templates for Cooperation Framework evaluations including:

- Quality criteria;
- TORs for evaluators and the Evaluation Manager;
- Inception report template;
- Evaluation report template; and
- Management response and action plan templates.

Relevant Documents:

- [UNSDCF Guidance](#)¹⁰
- [Guidelines for the Evaluation of the United Nations Sustainable Cooperation Framework \(September 2021\)](#)¹¹

⁵ Following the delinking of the United Nations resident coordinator position from UNDP in 2019, the UNDAF was renamed the UNSDCF and new guidance was issued. UNSDCF development and guidance are implemented by the United Nations resident coordinator's office under the United Nations Secretariat.

⁶ https://unsdg.un.org/sites/default/files/2019-10/UN-Cooperation-Framework-Internal-Guidance-Final-June-2019_1.pdf, Paras 103-104

⁷ <https://www.un.org/ecosoc/sites/www.un.org.ecosoc/files/files/en/qcpr/DCO-FundingCompactUpdate-Indicators.pdf>

⁸ <https://www.unsystem.org/content/un-system-wide-action-plan-gender-equality-and-empowerment-women-swap>

⁹ <https://www.un.org/en/content/disabilitystrategy/>

¹⁰ <https://unsdg.un.org/resources/united-nations-sustainable-development-cooperation-framework-guidance>

¹¹ <http://www.unevaluation.org/document/detail/2972>

- [In Brief: United Nations Sustainable Development Cooperation](#)¹²

Management and implementation

UNSDCF evaluations are **mandatory** and should be commissioned in the **penultimate year** (year prior to completion) of the Cooperation Framework period. The evaluation findings and recommendations should be completed in time to feed into the development of the new Cooperation Framework.

UNSDCF evaluations are managed by the United Nations Resident Coordinator's office, in cooperation with the United Nations country team (UNCT). Their role is to ensure that the final UNSDCF evaluation is independent and useful by:

- a) Facilitating and promoting national ownership, through involvement of national partners throughout the process and the timely incorporation of findings, recommendations and lessons into decision-making around subsequent United Nations CCAs and Cooperation Frameworks, as well as within the country development programming of respective United Nations development entities.
- b) Providing an independent evaluation team with all required information and access to national stakeholders, and coordinate entity-specific evaluations so they are relevant and timely in supporting the Cooperation Framework evaluation.
- c) Issuing an evaluation management response and action plan as a key accountability tool. They are responsible for publicly disclosing the final evaluation at the country level as part of an overall communications and dissemination strategy.

Management responses and key actions

The UNSDCF management response is a crucial step to improve the timely and effective use of the evaluation. Through the management response process, the UNCT and other evaluation stakeholders can review the recommendations and agree the follow-up steps and actions to be taken, or reject recommendations (with justification). Management responses should include whether the recommendation is accepted, partially accepted, or rejected, and why. For those recommendations accepted, actions that will be taken, by whom and when should be indicated. If recommendations are rejected, an explanation and potential alternative actions to address the issues raised should be included.¹³ Final UNSDCF evaluations, its recommendations, management responses and key actions no longer need to be uploaded to the Evaluation Resource Center (ERC). UNSDCF evaluations are not quality assessed by UNDP.

Box 4: Sample UNDAF evaluations

- Niger, 2018, [Evaluation finale de l'UNDAF 2014-2018](#)
- Mauritania, 2017, [UNDAF Final Evaluation, 2012 to 2017](#)
- Cambodia, 2017, [Evaluation of the UNDAF cycles 2011-2015, 2016-2018](#)
- Mozambique, 2015, [Evaluation of the UNDAF 2012-2016](#)
- Uruguay, 2015, [Evaluación de medio término del UNDAF 2011-2015](#)
- Afghanistan, 2017, [Midterm review report, 2015-2019](#)

¹² <https://unsdg.un.org/sites/default/files/2020-01/In-Brief-UN-Sustainable-Development-Cooperation.pdf>

¹³ <http://www.unevaluation.org/document/detail/2972>, p.19.

2.3 Decentralized country programme evaluations



Decentralized country programme evaluations (D-CPEs) assess UNDP attainment of intended results and contributions to development results at country level. The evaluation examines key issues such as UNDP effectiveness in delivering and influencing the achievement of development results and UNDP strategic positioning. These evaluations contribute to UNDP accountability and learning.

D-CPEs can be scheduled during (midterm) or towards the end (final) of a programme cycle.

Midterm D-CPEs

The midterm evaluation assesses the **level of effectiveness** in delivering the intended results in the country programme document (CPD), as well as the **positioning** of UNDP. This provides an accountability tool as well as a means to review progress and adjust direction if needed (course correction). The process is also an opportunity to extend dialogue with the government and partners on UNDP progress and programme direction.

Management and implementation

It is **highly recommended** that country offices consider commissioning midterm evaluations of country programmes, as an opportunity to review the attainment of intended results across all (or most) outcome areas.

Regional bureaux and policy and practice units may also decide to carry out midterm evaluations of their global or regional programmes to allow for course correction.

Final D-CPEs

A **final evaluation** should be conducted in the penultimate year (year prior to completion) of a UNDP country programme, to feed into the process of developing the new country programme. It focuses at the **outcomes** defined in the CPD.

Management and implementation

All new CPDs being presented to the Executive Board of UNDP should be accompanied by either a D-CPE, an ICPE or a Country Programme Performance Survey (CPPS).¹⁴

Methodology and guidelines

Midterm and final D-CPEs should follow the OECD/DAC evaluation criteria (see Box 2). Section 4 includes a list of possible questions to guide these evaluations, and a sample is provided below.

Sample evaluation questions

Relevance

- To what extent have the intervention logic / theory of change and the underlying assumptions of the country programme integrated gender equality and other cross-cutting issues?
- To what extent are they still valid or do they need to be adapted to changes in the needs or priorities of the country?

Effectiveness

- To what extent has progress been made towards the programme goals, including gender equality, women's empowerment and other cross-cutting issues?
-

¹⁴ <https://popp.undp.org/SitePages/POPPSubject.aspx?SBJID=461&Menu=BusinessUnit>

Efficiency	<ul style="list-style-type: none"> • What key results and changes (stated in the CPD) have been attained for men, women and vulnerable groups? • To what extent has the country programme delivered, or is likely to deliver, its interventions and results in an economic¹⁵ and timely manner? • To what extent were resources (funds, expertise, time) sufficient?
Sustainability	<ul style="list-style-type: none"> • To what extent did UNDP establish mechanisms to ensure the sustainability of the programme benefits for women, men and other vulnerable groups? • To what extent have partners committed to provide continuing support (financial, female and male staff etc.) to sustain the programme results?

2.4 Independent Country Programme Evaluations

Purpose and scope

The UNDP Independent Evaluation Office (IEO) undertakes Independent Country Programme Evaluations (ICPEs)¹⁶ of selected countries as they are coming to the end of their country programme cycles. ICPE findings, conclusions and recommendations serve as inputs to the UNSDCF evaluation process, as well as the process of developing the new UNSDCF and UNDP country programme.

Management and implementation

ICPEs come under the IEO plan and budget and therefore **do not need to be included in programme unit evaluation plans or budgets**. The IEO will inform a regional bureau of those country programmes to be subject of an ICPE, and make every effort to coordinate ICPE implementation with the units responsible for the country programme and UNSDCF development processes.

ICPEs cover one country programme cycle, and are carried out in **the penultimate year** (year prior to completion) of the programme cycle. They accompany new CPDs presented to the UNDP Executive Board for approval.

Methodology and guidelines

Key ICPE questions are:

- What did the UNDP country programme intend to achieve during the period under review?
- To what extent has the programme achieved (or is likely to achieve) its intended objectives?
- What factors contributed to or hindered UNDP performance and, eventually, to the sustainability of results?

Box 5: Sample decentralized country programme evaluations examples

- Nepal, 2020, Mid Term Review of Country Programme Document, 2018 to 2022
- Cambodia, 2019, Evaluation of UNDP Country Programme Action Plan, 2016 to 2018
- Ghana, 2017, UNDP country programme, 2012-2017
- Niger, 2017, Evaluation mi-parcours CPAP 2014- 2018
- Togo, 2017, Evaluation du CPAP, 2014 to 2018
- All ICPEs are available on the ERC website.

¹⁵ Economic in this sense refers to the conversion of inputs – funds, expertise, time - into outputs, outcomes and impacts, in the most cost-effective manner.

¹⁶ Access at: <http://web.undp.org/evaluation/evaluations/assessment-of-development-results.shtml>

2.5 Outcome evaluations



Outcome evaluations focus on UNDP contributions to the achievement of outcomes identified in the CPD.

Box 6: Definition of an outcome

*“Outcomes are **actual or intended changes** in development conditions that interventions are seeking to support.”*

*“Outcomes describe the intended changes in development conditions that result from the interventions of governments and other stakeholders, including international development agencies such as UNDP. They are medium-term development results created through the delivery of outputs and the contributions of various partners and non-partners. Outcomes provide a clear vision of what has changed or will change globally or in a particular region, country, or community within a period of time. They normally relate to **changes in institutional performance or behaviour among individuals or groups**. Outcomes cannot normally be achieved by only one agency and are not under the direct control of a project manager.”*

IMPORTANT: Outcomes are not the sum of outputs delivered through UNDP programmes and projects, rather they occur when outputs are used by primary stakeholders to bring about change.

Purpose and scope

Outcome evaluations **provide evidence of UNDP contributions to outcomes**. This evidence supports the accountability of programmes, and can be used by UNDP in its accountability requirements to its investors.

Outcome evaluations **guide the improvement of performance** within ongoing programmes by identifying areas of strength, weaknesses and gaps, especially in respect to the appropriateness of the UNDP partnership strategy and obstacles to achievement of outcomes. This can support mid-course adjustments to the theory of change; and lessons learned for the next programming cycle.

Outcome evaluations can provide evidence for and inform **higher-level evaluations**, such as of the UNSDCF, country, regional and global programmes, as well as support subsequent planning.



It is highly recommended that country offices evaluate at least one outcome during the country programme cycle period.

Management and implementation

Outcome evaluations can be managed and commissioned at midterm or towards the end of the programming cycle.

- **Midterm outcome evaluations** can highlight progress towards an outcome, and provide an opportunity to identify outcome implementation challenges and opportunities for course correction. **Midterm outcome evaluations can also support and inform D-CPEs and ICPEs.**

- **Final outcome evaluations** are ideally timed so that the findings and recommendations can support the development of the new UNSDCF and UNDP country programme. This should be in the **penultimate year** of the programme, and before the UNSDCF evaluation.

Methodology and guidelines

Outcome evaluations begin with the outcome(s) to be assessed and work backwards.¹⁷ Taking the outcome(s) as the starting point, the evaluation can assess a number of variables, including:

- (a) whether an outcome has been achieved, or the progress made towards it;
- (b) how, why and under what circumstances the outcome has changed;
- (c) the UNDP contribution to progress towards, or achievement of, the outcome; and
- (d) the UNDP partnership strategy in pursuing the outcome.



Outcome evaluations do not start by analyzing individual projects, as this approach is unlikely to yield useful or complete information. They only enquire what is happening at the outcome level.

Outcome evaluations are expected to apply the OECD/DAC evaluation criteria of relevance, coherence, effectiveness, efficiency and sustainability. Some possible outcome evaluation questions might include:

- 1) Were the stated outcomes or outputs achieved?
- 2) What progress has been made towards the outcomes?
- 3) Which factors have contributed to achieving (or not) the intended outcomes?
- 4) To what extent have UNDP outputs and assistance contributed to outcomes?
- 5) Has the UNDP partnership strategy been appropriate and effective? What factors contributed to effectiveness?

Box 7: Sample outcome evaluations

- Mauritania, 2020, Analyse des interventions du PNUD en appui à la conception et à la mise en œuvre, S&E de politiques publiques pro-pauvres
- Turkmenistan, 2019, Partnership Framework for Development evaluation
- El Salvador, 2019, Evaluación de la contribución del PNUD a la generación de capacidades de planificación y ejecución de la inversión pública a nivel nacional y local y su efecto en la cobertura y acceso universal a los servicios básicos de salud"
- Lebanon, 2019, Energy and Environment Outcome Evaluation
- Tanzania, 2019, Midterm evaluation of democratic governance outcome

2.6 Regional programme evaluations

¹⁷ UNDP Outcome Level Evaluation, 2011.

http://web.undp.org/evaluation/documents/guidance/UNDP_Guidance_on_Outcome-Level%20Evaluation_2011.pdf



Regional evaluations assess the attainment of intended results, as well as UNDP contributions to development results, as articulated in the regional programme document.

Purpose and scope

Regional programme evaluations are intended to reinforce the substantive accountability of UNDP to the Executive Board, and are timed to contribute to the preparation and approval of the subsequent regional programme.

Methodology and guidelines

Regional programme evaluations are similar to D-CPEs, may follow a similar methodology, and can be at midterm or towards the end of the programme cycle.

2.7 Project evaluations



Project evaluations assess the performance of a project in achieving its intended results, and its contribution to outcomes and associated theories of change.

Purpose and scope

Project evaluations yield useful information on project implementation arrangements and the achievements of outputs, and also draw linkages between a project's outputs and its contribution to broader programme outcomes.

The primary purpose of a project evaluation is to improve, continue or scale up an initiative; to assess its sustainability and replicability in other settings; to demonstrate accountability for results; or to consider alternatives. **Project evaluations play an important role in accountability to donors and governments involved in financing projects.** For their own accountability reasons, donor agencies and other cost-sharing partners may request UNDP to include evaluation requirements in the UNDP-donor partnership agreements.

Management and implementation

Project evaluation budgets should be agreed with partners and stakeholders and included in project documents and plans. **For all mandatory evaluations, plans and budgets need to be included in the respective project document.**

When to conduct a project evaluation

To ensure learning and accountability, and that results are being achieved, projects representing a significant financial investment and/or extending over a longer period should be evaluated (see table 1). Project evaluations are mandatory when required by partnership protocols, as in the case of the Global Environment Facility (GEF), Adaptation Fund and Green Climate Fund (GCF).



Project evaluations are mandatory when projects are expected to reach or pass certain thresholds as indicated in the table below:

Table 1. Thresholds for mandatory project evaluations

MANDATORY PROJECT EVALUATION THRESHOLDS	EVALUATIONS
Projects with a planned budget or actual expenditure of more than US\$ 5 million. If the project is under four years, then only one evaluation is required. ¹⁸	Midterm and final evaluation
Projects with a planned budget or actual expenditure of between \$3 million and \$5 million.	Midterm or final evaluation
Projects with duration of more than five years.	At least one evaluation: midterm or final
Projects entering a second or subsequent phase ¹⁹	One evaluation before moving into the new phase
Development initiatives being considered for scaling up	An evaluation before expansion

Methodology and guidelines

Project evaluations can use some of the following guiding questions based on OECD/DAC evaluation criteria and cross-cutting issues (see also Section 4, Annex 1):

Relevance	<ul style="list-style-type: none"> To what extent was the project in line with national development priorities, country programme outputs and outcomes, UNDP Strategic Plan and the SDGs?
Effectiveness	<ul style="list-style-type: none"> What have been the key results and changes attained for men, women and vulnerable groups? In which areas has the project had greatest achievements? Why and what have been the supporting factors? How can the project build on or expand these achievements?
Efficiency	<ul style="list-style-type: none"> To what extent was the management structure outlined in the project document efficient to generate the expected results? To what extent were the resources used to address inequalities in general, and gender issues in particular?
Sustainability	<ul style="list-style-type: none"> To what extent will targeted men, women and vulnerable people benefit from the project interventions in the long term? To what extent will financial and economic resources be available to sustain the benefits achieved by the project?

¹⁸ <https://popp.undp.org/SitePages/POPPSubject.aspx?SBJID=448&Menu=BusinessUnit&Beta=0>

¹⁹ A project is entering a second phase when it proposes to scale up results, through a substantive project revision or a new project.

Box 8: Sample project evaluations

- Pakistan, 2020, Midterm evaluation of upscaling of Glacial Lake Outburst Floods Project
- Tunisia, 2020, Évaluation finale - Assistance électorale
- Nepal, 2020, Final Evaluation of Resilient Reconstruction and Recovery of Vulnerable Communities Severely Affected by 2015 Earthquake (EU II)
- Regional Bureau for Africa, 2020, Midterm Evaluation of project “Accelerating the Ratification and Domestication of African Union Treaties”
- Haiti, 2020, Évaluation finale du projet Promotion de la Cohésion sociale à Jérémie
- El Salvador, 2020, Evaluación final del Proyecto "Apoyo al Ciclo Electoral de El Salvador" 2018-2019

2.8 Global Environment Facility, Adaptation Fund and Green Climate Fund evaluations

GEF evaluations

Management and implementation

Terminal evaluations are mandatory for **all medium- and full-sized projects** financed by GEF.

- All **full-sized projects** (with a GEF grant value of more than \$2 million) and all programmes must complete a midterm review *and* a terminal evaluation.
- All **medium-sized projects** (with a GEF grant value of between \$500,000 and \$2 million) must complete a terminal evaluation. Midterm reviews are optional for medium-sized projects.

Joint agency projects require just one evaluation, managed by the lead agency.



Both midterm and terminal evaluations should be included in UNDP evaluation plans (for the country office or regional bureau) and uploaded to the ERC.

Methodology and guidance

Separate GEF guidance for terminal²⁰ and midterm²¹ evaluations outline the procedures and approaches that must be followed. These include guidance on evaluation processes, roles and responsibilities, templates and outlines for TORs and evaluation reports, and sample evaluation criteria matrices.

Box 9: GEF evaluation guidelines

- [Guidance for conducting terminal evaluations of UNDP-supported, GEF-financed projects, 2020](#)
- [Guidance for conducting midterm reviews of UNDP-supported, GEF-financed projects, 2014](#)

GEF evaluations must be **independently implemented and quality-assured** as the GEF Independent Evaluation Office compares the quality of terminal evaluations between GEF agencies.

²⁰ http://web.undp.org/evaluation/guideline/documents/GEF/TE_GuidanceforUNDP-supportedGEF-financedProjects.pdf

²¹ http://web.undp.org/evaluation/documents/guidance/GEF/midterm/Guidance_Midterm%20Review%20EN_2014.pdf

Box 10: Sample GEF terminal evaluations

- India, 2020, Mainstreaming Coastal & Marine Biodiversity Conservation in Andhra Pradesh
- Bangladesh, 2020, Terminal Evaluation of Expanding the Protected Area System to Incorporate Important Aquatic Ecosystems project
- China, 2019, Final Evaluation for Wetlands Portfolio - National
- Jordan, 2019, Mainstreaming Rio Convention Provisions into National Sectoral Policies

Adaptation Fund evaluations

Management and implementation

All Adaptation Fund regular projects (those with a grant budget of over \$1 million) are subject to a final evaluation by an external evaluator selected by the implementing entity. The Adaptation Fund Evaluation Framework states that: *“Projects and programmes that have more than four years of implementation will conduct an independent midterm evaluation after completing the second year of implementation.”*²² All small size projects (of up to \$1 million), as well as readiness grant projects, are subject to a final evaluation if deemed appropriate by the Adaptation Fund Board.

Final evaluation reports will be submitted to the Adaptation Fund Board as stipulated in the project agreement.²³ **Adaptation Fund midterm and final evaluations should be included in UNDP (country office and regional bureau) evaluation plans and uploaded to the ERC.**

Methodology and guidelines

Evaluations will be conducted following a process decided by the Board and using templates approved by the Board.

Green Climate Fund evaluations

Methodology and guidelines

The GCF Evaluation Policy is currently under development.²⁴

Management and implementation

All GCF evaluations should be included in UNDP evaluation plans (country office or regional bureau evaluation plans) and uploaded to the ERC.

2.9 Multi-country project evaluations

²² Access at: <https://www.adaptation-fund.org/document/evaluation-framework-4/>

²³ Access at: <https://www.adaptation-fund.org/wp-content/uploads/2017/08/OPG-amended-in-October-2017-1.pdf>

²⁴<https://ieu.greenclimate.fund/evaluation-policy-of-the-gcf>



Multi-country project evaluations are of projects being implemented in multiple countries.²⁵

Management and implementation

Though considered a single project for project management purposes, the offices involved are accountable for their contributions to joint results on equal terms. In this scenario, the implementation modality is **similar to a joint evaluation (see below)**, with participating UNDP offices agreeing on the management structure, and collaborating in drafting the TOR, selecting evaluators, reporting, dissemination, developing management responses, and following up and implementing recommendations. The UNDP country office managing and commissioning the evaluation is designated the “coordination office”.

2.10 Portfolio evaluations



Portfolio evaluations refer to those evaluating the work of a group or portfolio of projects designed to contribute to a country programme output or outcome(s).

Purpose and scope

A portfolio evaluation is similar to an outcome evaluation, but may focus on part of an outcome, or a group of projects that are linked to different outcomes in the country programme results framework. For instance, a country office may evaluate a set of interventions contributing to strengthening access to justice and the judiciary, which is a component of a larger outcome referring to strengthened governance.

Management and implementation

From a practical perspective, conducting a portfolio evaluation may allow several projects to contribute to the financing of the evaluation, rather than funding separate project evaluations. This should be planned in advance and budgets agreed across projects if needed, and the approach included in the evaluation plan of the implementing unit.

Portfolio evaluations are managed and commissioned by the UNDP country office, and **should be included in UNDP evaluation plans and uploaded to the ERC.**

2.11 Thematic evaluations



Thematic evaluations assess UNDP performance in areas that are critical to ensuring sustained contribution to development results. They may focus on one or several cross-cutting themes that have significance beyond a particular project or initiative, across several outcomes or results areas in a country, such as gender mainstreaming, capacity development, human rights or democratic governance.

²⁵

https://popp.undp.org/_layouts/15/WopiFrame.aspx?sourcedoc=/UNDP_POPP_DOCUMENT_LIBRARY/Public/PPM_Design_Multi-country%20and%20South-South%20Cooperation%20Projects.docx&action=default

The IEO undertakes thematic evaluations as part of its mandate to look closely at the achievements and challenges of the organization. Examples of thematic evaluations in UNDP include areas such as democratic governance, or cross-cutting themes such as gender mainstreaming.

Methodology and guidance

The objectives, scope and questions for thematic evaluations will vary depending on the subject matter. However, evaluation questions should adhere to the OECD/DAC evaluation criteria of relevance, coherence, effectiveness, efficiency and sustainability (see Box 2).

The following evaluation questions were developed for a **thematic evaluation on gender mainstreaming and women’s empowerment**. More guiding questions are available in Section 4.

Relevance	<ul style="list-style-type: none"> ▪ To what extent has UNDP contributed to gender equality and women’s empowerment development results at the policy and implementation level? ▪ To what extent does UNDP operationalize gender concerns in its programme/ project designs (results frameworks, theories of change) and identify indicators for projects to ensure “quality-at-entry” of gender-mainstreamed interventions?
Effectiveness	<ul style="list-style-type: none"> ▪ How effective has UNDP been in implementing gender mainstreaming and contributing to institutional change? ▪ To what extent have UNDP programmes, projects and other interventions contributed to promoting gender equality that benefits women, men and other vulnerable groups?
Efficiency	<ul style="list-style-type: none"> ▪ To what extent are UNDP resources (financial, time, male/female staff, technical and gender expertise) adequate to address gender inequalities and root causes?
Sustainability	<ul style="list-style-type: none"> ▪ To what extent will the benefits of UNDP programmes, projects and other interventions in respect to gender mainstreaming and women’s empowerment continue, or are likely to continue in future?

Box 11: Sample thematic evaluations

- [Mauritania, 2020, Analyse des interventions du PNUD en matière de promotion de l’égalité genre.](#)
- [Tanzania, 2018, Gender impact midterm evaluation across the three programme outcomes](#)
- [Panama, 2018, Evaluación temática de diálogos facilitados por PNUD en Panamá en marco de Programa de País 2016-2020](#)
- [BPPS, 2017, External Assessment of the UNDP Gender Seal](#)

2.12 Impact evaluations



An impact evaluation explores the effects (positive or negative, intended or not) on individual households and institutions, and the environment created, by a given development activity such as a programme or project. Such an evaluation refers to the long-term impact as well as medium-term effects at the outcome level.

Purpose and scope

Impact evaluations do not simply measure whether objectives have been achieved or assess direct effects on intended beneficiaries. They include the full range of impacts at all levels of the results chain, including ripple effects on families, households and communities, on institutional, technical or social systems and

on the environment. In terms of a simple logic model, there can be multiple intermediate (short- and medium-term) outcomes over time that eventually lead to impact, some or all of which may be included in an evaluation of impact at a specific moment in time. By identifying whether development assistance is working or not, impact evaluations serve an accountability function.

Methodology and guidance

UNDP undertakes very few impact evaluations as many of our projects contribute to a broader outcome or development goal or play a role for which it is difficult to attribute impact.²⁶ When projects are being designed and an impact evaluation is expected, programme units should consider the type of impact that is expected and the indicators which can illustrate this. A baseline measure will give the pre-project levels and realistic targets can then be built into the project goals, monitored regularly and finally checked and validated (or not) through an impact evaluation.

UNEG has developed detailed impact evaluation guidance.²⁷

2.13 Joint evaluations



Joint evaluations refer to evaluative effort by more than one entity of a topic of mutual interest, or of a programme or set of activities which are co-financed and implemented. The degree of 'jointness' varies from cooperation in the evaluation process, to the pooling of resources for combined reporting.²⁸

Any evaluation can be conducted as a joint evaluation.

UNDP is increasingly being asked to undertake joint evaluations. Depending on the extent to which individual partners cooperate in the evaluation process, it requires discussions with the evaluation partners as to:

- **whose procedures** should be used, both for the evaluation and for the procurement;
- **which funding modalities** and contributions from different parties should be applied; and
- **how** the process is managed and reported to all parties.

At country level, an obvious example of a joint evaluation is the UNSDCF evaluation, where several United Nations organizations and national government participate. In addition, a UNDP country office may wish to conduct a joint outcome evaluation together with a partner, government and / or donor.

Purpose and scope

Joint evaluations can have greater **benefits and challenges**, highlighted in the UNEG Resource Pack on Joint Evaluations,²⁹ 2014 and summarised in Table 2. In general, joint evaluations take longer to plan, implement and finalize in comparison to other evaluations, and require more coordination efforts and commitment among evaluation stakeholders. All steps in the evaluation process, such as agreeing on TORs, selecting the evaluation team, commenting on the draft inception and final reports need to be conducted jointly.

²⁶ Between 2016 and 2017, UNDP implemented 600 decentralized evaluations, which included only one impact evaluation.

²⁷ Access at: <http://www.uneval.org/document/detail/1433>

²⁸ Access at: <http://www.uneval.org/document/detail/1620>

²⁹ Access at <http://www.unevaluation.org/document/detail/1620>

Table 2. Benefits and possible challenges of joint evaluations

Benefits	Possible Challenges
Increased objectivity and legitimacy	Different terminology
Broader scope and picture of situation	Developing TORs
Mutual capacity development and peer learning	Complexity of coordination arrangements
Helps foster a collective sense of accountability	Power differentials among partners
Promotes cost-sharing and a sense of partnership among United Nations evaluation offices and units	Longer time frames
Useful for developing consistent policy messages	Findings may not address individual agency accountability requirements
Greater credibility and broader ownership of findings and recommendations	Diffusion of responsibility for follow-up can weaken evaluation impact
Builds coordination and collaboration	Multisector data may be more complicated to interpret
Often yields higher quality evaluations	
May help reduce the overall number of evaluations undertaken – thereby reducing transaction costs and administrative demands on aid recipient countries.	

Management and implementation

Experiences from previous joint evaluations have provided the following lessons and recommendations for effective joint evaluation:³⁰ A check list is available in the annexes of this section that will help to plan and organize any joint evaluations.

- Keep a simple and light management structure
- It is critical to have a core group of four or five (or less) agencies involved at an early stage to move it forward
- Delineate clear roles and responsibilities
- Decide which evaluation guidelines will be used
- Be clear of any donor or funding agency evaluation requirements and processes.
- Ensure adequate resources since joint evaluations require full-time coordination, administration, and research support
- Develop and apply a good communications and participation strategy to keep agencies involved
- Agree to speak with one voice to the evaluation team
- Ensure adequate funding, including a contingency budget (if, indeed, the dissemination and follow-up are not funded upfront)

³⁰ Adopted from UNEG, 2014, Resource Pack on Joint Evaluations. <http://uneval.org/document/detail/1620>

- Start writing the TOR as early as possible
- Be as specific as possible about the purpose and objective of the evaluation in the TOR
- Ensure that all partner organizations/agencies have started early in collecting relevant information for the evaluation team
- Ensure sufficient time for drafting TORs, selecting evaluators, a kick-off meeting, and reviewing the draft inception and final reports.



Joint evaluations follow the same approach as other evaluations, but each step in the evaluation process needs to be agreed by all parties. These steps are described in the annex of this section.

2.14 Evaluations commissioned by donors or multilateral organizations

Purpose and scope

Donor-funded projects and programmes may require evaluations, and these must be planned in advance and included in project documents. If evaluations are mandatory, this will be stated in the agreement between UNDP and the donor, and the timing and funding will be agreed in the project contract. These evaluations may be commissioned directly by the donor agencies or by UNDP.

Management and implementation

Evaluations commissioned by donors should be planned and completed collaboratively, including developing the scope and methodology, and procedural matters should be agreed in advance with the donor. UNDP should ensure that donor partners share the draft report for comment prior to final issuance and completion.

These evaluations should be included in programme unit evaluation plans and uploaded to the ERC website. All UNDP evaluations are treated as public documents and this should be communicated to the donor. Where an evaluation is not planned and has not been included in the evaluation plan, the plan must be revised to include the new evaluation, and this must be uploaded to the ERC.

Recommendations, management responses and key actions which are specifically applicable to UNDP and the programme unit should be uploaded to the ERC within six weeks of completion of the evaluation report. Other non-UNDP related recommendations can be omitted from the ERC, though they remain in the evaluation report. The programme unit needs to monitor the implementation and report on the progress of the planned key actions until they have all been completed.

Methodology and guidance

Commissioning donors and organizations may have their own evaluation guidance documents and standards, which will need to be followed alongside UNDP and UNEG guidance.

Peace Building Fund (PBF) project evaluations

Purpose and scope

According to the Guidance Note on PBF Project Operational and Financial Closure,³¹ an independent project evaluation is one of the requirements for closing a project. The objectives of a PBF project evaluation are set out in the standard TOR and include:

- Assess the **relevance and appropriateness** of the project in terms of:
 - addressing key drivers of conflict and the most relevant peacebuilding issues;
 - alignment with National Peacebuilding Policy and national priorities of the country;
 - whether the project capitalized on the United Nations added value in the country; and
 - the degree to which the project addressed cross-cutting issues such as cross-border conflict dynamics, youth- and gender-responsiveness in the country.
- Assess the extent to which the PBF project has made a concrete **contribution to reducing a conflict factor** in the country and advancing achievement of the SDGs, in particular SDG 16.
- Evaluate the project's **efficiency**, including its implementation strategy, institutional arrangements, management and operational systems and value for money with particular regard to distribution of labour across multiple fund recipients through joint projects.
- Assess whether the support provided by the PBF has promoted the **Women, Peace and Security agenda**, allowed a specific focus on women's participation in peacebuilding processes, and was accountable to gender equality.
- Assess whether the project has been implemented through a **conflict-sensitive approach**.
- Document **good practices, innovations and lessons** emerging from the project.
- Provide **actionable recommendations** for future programming.

Management and implementation

The evaluation is managed by the recipient agencies, under the supervision of the PBF Secretariat where this exists. Recipient agencies should be reminded about the evaluation requirements and need to provide support and oversight to the evaluation process, including on TORs and evaluation deliverables. PBF guidance stresses the need to ensure sufficient budgets relative to the investment to support high-quality evaluations, and PBF policies stipulate a minimum of 5-7 percent of the total project budget dedicated to monitoring and evaluation.

Methodology and guidance

PBF has developed a template for TORs of project evaluations, which are available to guide evaluation managers in the particular features of evaluations of peacebuilding interventions. PBF project evaluations can be structured around the following types of guiding questions:

³¹

https://www.un.org/peacebuilding/sites/www.un.org.peacebuilding/files/documents/pbf_guidance_note_on_project_closure_-_2020.pdf

Relevance	<ul style="list-style-type: none"> Was the project relevant in addressing the conflict drivers and factors for peace identified in the conflict analysis? Was the project appropriate and strategic to the main peacebuilding goals and challenges in the country at the time of the project design? Did relevance continue throughout implementation?
Effectiveness	<ul style="list-style-type: none"> To what extent did the PBF project achieve its intended objectives and contribute to the Fund's strategic vision? To what extent did the PBF project substantively mainstream gender and support gender-responsive peacebuilding? How appropriate and clear was the PBF project targeting strategy in terms of geographic and beneficiary targeting? Was the project monitoring system adequately capturing data on peacebuilding results at an appropriate outcome level?
Efficiency	<ul style="list-style-type: none"> How efficient was the overall staffing, planning and coordination within the project (including between fund recipients, implementing partners and stakeholders)? Have project funds and activities been delivered in a timely manner?
Sustainability	<ul style="list-style-type: none"> To what extent did the PBF project contribute to the broader strategic outcomes identified in nationally-owned strategic plans, legislative agendas and policies? Did the project design include an appropriate sustainability and exit strategy (including promoting national/local ownership, use of national capacity etc.) to support positive changes in peacebuilding after the end of the project? To what extent was the project financially catalytic?
Conflict-Sensitivity	<ul style="list-style-type: none"> Did the PBF project have an explicit approach to conflict-sensitivity? Were United Nations and non-UN recipient organizations' internal capacities adequate for ensuring an ongoing conflict-sensitive approach? Was the project responsible for any unintended negative impacts? Was an ongoing process of context monitoring and a monitoring system that allows for monitoring of unintended impacts established?

The PBF has also developed a checklist for evaluative exercises in the context of COVID-19.³²

Box 12: European Commission Directorate evaluation guidance documents

- Directorate-General for International Cooperation and Development (DG DEVCO)³³
- Directorate-General for European Civil Protection and Humanitarian Aid Operations (DG ECHO)³⁴
- Directorate-General Regional and Urban Development (DG REGIO)³⁵

³² Access at: <https://www.un.org/peacebuilding/fund/documents/guidelines>

³³ https://ec.europa.eu/international-partnerships/our-impact/monitoring-and-evaluation_en

³⁴ https://ec.europa.eu/echo/funding-evaluations/evaluations_en

³⁵ https://ec.europa.eu/regional_policy/en/policy/evaluations/guidance/

Annex 1. Joint evaluation consideration checklist



1. Deciding on the need for a joint evaluation

It is important to assess whether the programme or project warrants a joint evaluation.

- Is the focus of the programme on an outcome that reaches across sectors and agencies?
- Is the programme co-financed by multiple partners?
- Would a joint evaluation reduce evaluation transaction costs?
- Can the project be evaluated (evaluability)?



2. Determining partners

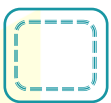
Choose evaluation partners at an early stage to ensure their involvement and ownership.



3. Management structure

A recommended structure for a joint evaluation could have a steering group that oversees the evaluation process and a smaller management group to ensure smooth implementation.

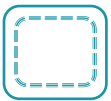
- The steering group comprises a representative from each partner organization and government entity.
- The steering group approves the TOR, and the evaluation team ensures oversight of the evaluation, introduces balance in the final evaluation judgements and takes responsibility for the use of results.



4. Division of work

Senior management of the UNDP programme unit should agree with the evaluation partners on the decision-making arrangements and the division of labour at the outset of the evaluation process.

- This includes who in the management group takes the lead role in each of the subsequent steps in the evaluation.
- A conflict resolution process should be determined to deal with any problems that may arise.



5. Drafting the TOR

It is practical for one party to take the lead in drafting the evaluation TOR, which define the scope of work. The draft should be discussed and agreed upon by the partner organizations and the interests of all parties concerned should be included and agreed in the TOR.



6. Determining whose procedures will be used

Different organizations take different approaches to evaluation, and it is important to allow flexibility to adapt and allow additional time to accommodate delays due to such differences. Implementing agencies could:

- Agree that the evaluation will be managed using the systems and procedures of one agency or
- Split the evaluation into components and agree whose systems will be used to manage which components.

Whichever approach is taken will determine appropriate templates, budgeting norms and approaches and report finalization procedures. These approaches should be agreed prior to the evaluation starting.



7. Funding modalities

If UNDP is taking the lead, the preferred funding approach should be to pool partners' financial support into a fund (akin to a trust fund) that is administered by one agency and that covers all costs related to the exercise. Alternatively, individual partner(s) could finance certain components of the evaluation while UNDP covers others. This approach increases transaction and coordination costs.



8. Selecting evaluators

One of the joint evaluation partners could take responsibility for recruiting the evaluation team, in consultation with the other partners. Another option is for each of the partners to contribute their own experts. However, an evaluation team leader should be hired and agreed by partners to aid the smooth implementation, organization and final report development of the evaluation. Guidance on evaluator recruitment can be found in section 5.



9. Report dissemination strategies

For a joint evaluation, partners should agree that they have the opportunity to correct factual errors in the report; where it is impossible to resolve differences on the findings and conclusions, dissenting views should be included in the report; and the conclusions and recommendations should be the responsibility of the evaluators. However, sometimes measures such as allowing for separate evaluation products may be beneficial for the partners who have certain accountability or reporting requirements.



10. Management response, follow-up and implementing recommendations

All managers must follow up on the findings and recommendations of each evaluation report in UNDP. Partners need to agree on what to do individually and collectively and decide upon a follow-up mechanism that monitors the status of the changes being implemented. In line with the Evaluation Policy requirement, UNDP may select recommendations that are pertinent to UNDP and prepare a management response focusing on these recommendations.



SECTION 3

EVALUATION PLAN DEVELOPMENT



CONTENTS

3.1	Introduction	1
3.2	Step One: Developing an evaluation plan	2
3.3	Step Two: Evaluation plan content	3
3.4	Step Three: Costing and identifying sources for the evaluation plan	4
3.5	Step Four: Evaluation plan template	8
3.6	Step Five: Evaluation plan review and quality assurance process	9
3.7	Step Six: Evaluation plan completion and approval	10
3.8	Step Seven: Making changes to the evaluation plan	11
3.9	Step Eight: Monitoring compliance	12
Table 1. Budget considerations and calculation for evaluations		5
Table 2. Evaluation plan template		8
Table 3. Evaluation plan checklist		9
Table 4. Sample evaluation plan scheduling checklist		10
Table 5. Examples of well-balanced evaluation plans		13
Figure 1. The steps in evaluation plan development		1

3. DEVELOPING AN EVALUATION PLAN

Section 3 provides guidance on how to develop and use an evaluation plan that will provide appropriate evaluative coverage of a programme (e.g., a country programme, in the case of country offices). This section details what should be included in an evaluation plan, how to cost the plan, the review and approval process, and how the evaluation plan can be updated.



An evaluation plan is a mandatory and strategic document outlining the evaluations planned for a country programme, and is used to monitor progress. As the evaluation plan and its evaluations support management decision-making, they should reflect programmatic priorities.

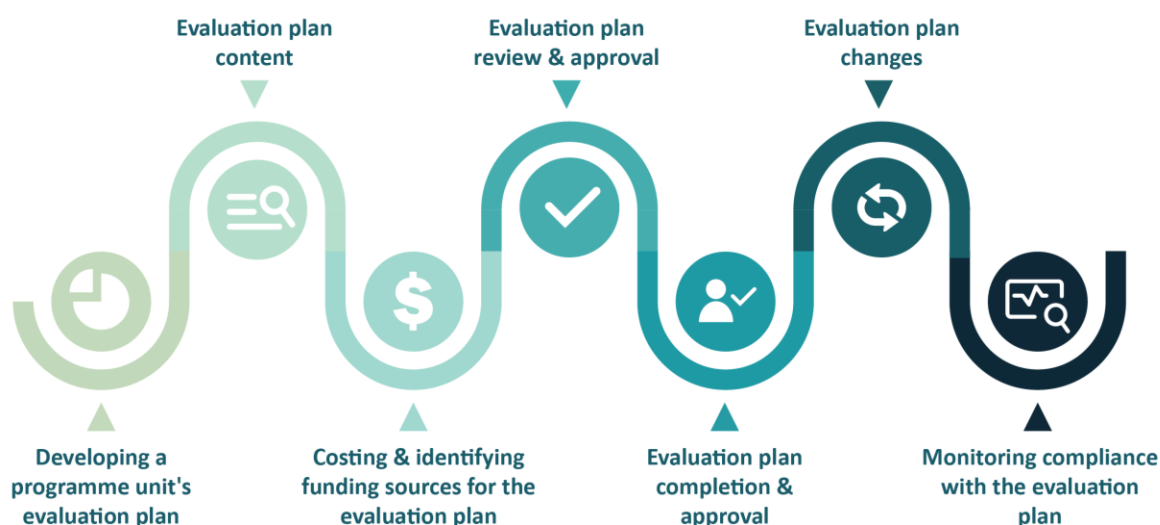


Figure 1. The steps in evaluation plan development

3.1. Introduction

This section gives details on how a programme unit (such as a country or regional office) can develop its mandatory evaluation plan, including who should be involved, what the plan should contain, budget considerations, and how the plan is managed throughout the programme cycle.¹

As a programme unit plans its activities over a strategic period (for example, the country programme period), it is important also to plan how it will **check its progress towards agreed development goals and outcomes** at all levels (project, programme, outcome, etc.). Evaluation planning is necessary in order to:

¹ This section refers to the country office as the key programme unit, but the guidance is equally applicable to other programme units such as regional offices.

i) support course correction if needed; ii) check progress (in the case of midterm evaluations and reviews); or iii) capture results (in the case of final or terminal evaluations).

An evaluation plan is a strategic document that is constantly used to check progress towards agreed evaluation commitments, produce evaluation findings to support change, aid knowledge-gathering and inform the work of UNDP. The evaluation plan accompanies the draft country programme document (CPD) as an annex when it is submitted to the Executive Board for approval.² Programme units should ensure that the evaluation plan is an effective learning and accountability tool, not just a compliance document containing only mandatory evaluations.

The evaluation plan should be reviewed annually and refined and adjusted as needed. The annual country office **business planning meeting** at the beginning of the year offers a good opportunity to review the evaluation plan. A formal **midterm review of the evaluation plan** is also highly recommended.

3.2. Step One: Developing an evaluation plan

Programme units must **present a timed and fully costed evaluation plan to the Executive Board** with each country, regional and global programme document considered for approval. The plan should be strategic, practical, cost-effective and include evaluations of different types (project, programme, outcome, etc.) that will generate the most critical and useful information for UNDP and its partners in future programming. The plan should ensure accountability and learning from implementation.

When submitted to the Programme Appraisal Committee (PAC) for review, all evaluation plans **must be accompanied by an evaluation rationale**: a brief text (maximum 300 words) explaining the justification for the evaluations included in the plan. This is for internal use only and should not be submitted to the Executive Board. The evaluation rationale should explain:

- How the evaluations contribute to learning and accountability, and the achievement of strategic results.
- How the evaluations provide sufficient and balanced coverage of the programme unit's areas of engagement.

As with the country programme development process, government, partners and stakeholders need to be included in the development of the evaluation plan. **Therefore, the evaluation plan should be developed through the same process as the country programme.**



The programme unit senior management leads the development of the evaluation plan and is accountable for its implementation. Typically, the programme unit monitoring and evaluation (M&E) focal point coordinates with programme teams and other stakeholders in the development of the evaluation plan, in order to identify which evaluations should be carried out and why. Regional evaluation focal points should also be included in the review of draft country office evaluation plans.

²https://popp.undp.org/_layouts/15/WopiFrame.aspx?sourcedoc=/UNDP_POPP_DOCUMENT_LIBRARY/Public/PPM_Formulate%20Programmes%20and%20Projects_Evaluation%20Plan%20Template.docx&action=default

3.3. Step Two: Evaluation plan content

In deciding what to evaluate, programme units should first determine the purpose of proposed evaluations, as well as other factors (such as country office priorities, emerging areas of engagement or potential scale-up opportunities) that may influence the relevance and use of evaluations.

The evaluation plan should reflect the goals and outcomes of the country programme and take a balanced approach, ensuring evaluation of all programmatic areas to provide the broadest accountability and learning.

The contents of the evaluation plan should be checked against the following criteria:

1. Planned evaluations are strategic:

- (a) Evaluations that provide substantive information for decision-making and learning;
- (b) Evaluations that address the programme unit priorities, emerging areas of engagement, potential scale-up opportunities and cross-cutting issues.³

2. Evaluation coverage is as inclusive and balanced as possible:

- (a) **A range of evaluations (outcome, project, thematic and others) are included in the evaluation plan** to provide comprehensive evaluation coverage of the programme.⁴ Any revisions should ensure that a comprehensive evaluation focus is retained.

3. All mandatory evaluations are included⁵:

- (a) Global Environment Facility (GEF) terminal evaluations for all GEF-financed medium-size projects and full-size projects;⁶
- (b) GEF midterm reviews for full-size projects;
- (c) Adaptation Fund and Green Climate Fund projects as required;
- (d) Donor/cost-sharing agreement evaluations.

4. Inclusion of project evaluations meeting the following criteria:⁷

- (a) Projects with a planned budget or actual expenditure of over US\$ 5 million must plan and undertake both a midterm and final evaluation;⁸
- (b) Projects with a planned budget or actual expenditure between \$3 million and \$5 million must plan and undertake either a midterm or final evaluation;⁹
- (c) Projects with a duration of more than five years¹⁰ must plan and undertake either a midterm or final evaluation;

³ For example, gender, crisis prevention and recovery, youth empowerment, HIV/AIDS, human rights or governance.

⁴ For instance, if a programme unit has a strong focus on/ large portfolio of disaster risk management, then its evaluation plan should reflect this.

⁵ UNSDCF evaluations do not need to be included in the UNDP evaluation plans (as of June 2022). These evaluations will not be quality assessed by UNDP IEO.

⁶ GEF medium-size projects are those with up to \$2 million in grant funds, GEF full-size projects are those of over \$2 million in grant funds.

⁷ Country offices may request the regional evaluation focal point waive evaluations based on reasonable justification. At the same time, if a project is due to be evaluated as part of an outcome, portfolio or thematic evaluation, a separate project evaluation may not be necessary.

⁸ If the project has a duration of under four years, only one evaluation is required.

⁹ This covers projects and not development services. While it is recommended that programme units evaluate large development service projects, delivery efficiency can be covered through audits.

¹⁰ Projects exceeding five years should be evaluated within six months if they have not yet been evaluated.

- (d) Projects entering a second phase should plan and undertake an evaluation;¹¹
- (e) Development initiatives being considered for scaling up should be evaluated before expansion.

5. Timing, costs, resources and sequencing are realistic:

- (a) The evaluation plan should consider the timing of evaluations across a full **evaluation calendar**. When developing an evaluation calendar, it is important to ensure that timing allows for completion and contribution to key planning activities and other evaluations being undertaken by the implementing unit, such as outcome evaluations, UNSDCF evaluations, and independent country programme evaluations conducted by IEO;
- (b) The calendar should ensure that evaluations are not ‘bunched together’ for completion at the same time, such as the end of the country programme period or the end of a calendar year (when other reporting is required), which will overstretch human resources and impact oversight;
- (c) Evaluation plans and calendars should consider that evaluations should be completed and uploaded to the Evaluation Resource Center (ERC) by December;
- (d) Evaluation costs should be realistic and funds for evaluations made available. For further detail, see subsection 3.4 (costing).

6. Influencing and constraining factors have been fully considered:

Socioeconomic, political and environmental risks should be considered when outlining the evaluation plan and calendar. Examples include elections (national and local), cultural and religious festivals, rainy seasons (which can impact travel) and planting and harvesting times, when community members can be extremely busy. All of these could impact on the availability of interviewees and the scope of data collection.

3.4. Step Three: Costing and identifying sources for the evaluation plan

Costing of the evaluation plan is important and should be realistic, in relation to the requirements and scope of the evaluation, as well as the realities of the country office budget. The Independent Evaluation Office (IEO) annual report on evaluation gives average annual costs for different types of evaluations across the globe, as well as at regional level (in the annexes), which should be used as a guide, although there will be differences between country offices.

Programme units should estimate and indicate financial requirements and financing sources for each evaluation in the evaluation plan. When estimating the cost for an evaluation, it is important to consider the scope, depth and duration of the evaluation, as well as the composition of the planned evaluation team.



The greater the complexity and scope of an evaluation, the longer time and more detailed work will be needed for preparation by the responsible programme unit, and for data collection by the evaluation team, which will increase evaluators’ overall fees and therefore total evaluation costs.

¹¹ A project is entering a second phase when it is proposed to scale up results through a substantive project revision or a new project.

A further consideration is the cost of the travel by the evaluation team. Programme units should be realistic in terms of the scope and complexity of the evaluation vis-à-vis available resources. Finally, programme unit should consider communication and dissemination costs for wider dissemination of the evaluation report.



Underfunding evaluations will seriously constrain their scope, results, quality and credibility. When identifying the sources of funds for evaluations, the following should be considered:

- **Decentralized country programme evaluations/ outcome evaluations** should have resources set aside in the country programme budget. Alternatively, related projects should contain a budget line to allow for sufficient resources for an outcome evaluation.
- **Project evaluations** should have a budget line for evaluation activities, exclusive of monitoring activities.
- **Portfolio and thematic evaluations** could take funds from across related projects to evaluate results.
- **GEF terminal and midterm evaluation** guidelines give suggested budget outlines.

Individual evaluation budget considerations include:

- **Professional fees** for all evaluators or thematic experts undertaking the evaluation (international and national). There are often additional management costs when hiring a professional firm.
- **Travel** to and from the evaluation country, where applicable
- **Additional and non-professional costs** such as daily subsistence allowance for time in country for data collection and terminal expenses.
- **Translation costs** for interviews, field visits, validation and dissemination workshops.
- **Travel costs** within the country during the evaluation (evaluator, translator, UNDP accompanying staff and other participants).
- Any focus group meeting or data-collection **meeting costs** (venue hire, snacks, participant transport costs etc.).
- **Communication costs** including editing, publication and dissemination costs.
- Stakeholder, validation, or evaluation reference group **workshop costs**.
- Additional **contingency costs** for any unknown expenses during the evaluation.

Table 1. Budget considerations and calculation for evaluations

A. EVALUATION TEAM COSTS		# DAYS	DAILY RATE	TOTAL COST
Professional fees	Team Leader/ Evaluator 1			
	Evaluator 2			
	TOTAL			
Flights (international)	Evaluator 1			
	Evaluator 2			

	TOTAL			
Per diem costs (time in the field)	Evaluator 1			
	Evaluator 2			
	TOTAL			
TOTAL A				

B. EVALUATION IMPLEMENTATION AND DATA-COLLECTION COSTS		#	COST	TOTAL
Internal flights				
Car hire				
Translation				
Focus group and workshop-related costs				
Other costs				
TOTAL B				

C. EVALUATION DISTRIBUTION COSTS		#	COST	TOTAL
Report production (editing, design, printing)				
Report dissemination (outreach, shipping etc.)				
Stakeholder meeting				
TOTAL C				

TOTAL EVALUATION COSTS A+B+C	
-------------------------------------	--

Where an individual or a group of individuals is hired to undertake an evaluation, most of the costs cited above will be managed by UNDP. Where UNDP engages a firm to undertake an evaluation, some of the costs (such as flights and per diems) may be managed by the firm on behalf of UNDP.

It is important that an evaluation be fully costed and budgeted for, to allow for adequate scope and duration of the evaluation, and also to ensure that additional incidental costs are included.

In all cases, whether an individual or a firm is engaged, the budget and financing expectations and responsibilities must be clarified and agreed prior to the evaluation starting.



Evaluation budgets are separate to monitoring budgets and should be detailed under a separate budget line. Delineation of monitoring and evaluation budgets is required under the 2019 Evaluation Policy.¹²

Joint evaluations require evaluation partners to agree whose procedures should be used, both for the evaluation and for procurement, the funding modalities and contributions from different parties, and how the process is managed and reported to all parties. Section 2 Annex 1 outlines these considerations and others related to joint evaluations.

¹² Access at: <http://web.undp.org/evaluation/policy.shtml>

3.5. Step Four: Evaluation plan template

The completed evaluation plan template (see Table 2)⁸² accompanies the draft CPD as Annex 2 Fully Costed Evaluation Plan.⁸³

Table 2. Evaluation plan template

UNSDCF (or equivalent) outcome	UNDP Strategic Plan outcome	Evaluation title	Partners (joint evaluation)	Evaluation commissioned by (if not UNDP)	Type of evaluation	Planned evaluation completion date	Estimated cost	Provisional source of funding
Copied verbatim from the UNSDCF/ equivalent/ CPD	Cite relevant Strategic Plan outcome	E.g., Midterm outcome evaluation: Energy and Environment Portfolio	List all partners, e.g., United Nations organizations; government partners such as national ministries; donors etc.	E.g. Ministry of Environment; GEF	E.g., country programme, outcome, thematic, programme / project, GEF, etc. Note: Evaluative exercises may vary in size and scope, but they should all help produce information towards outcome-level. All evaluations should meet the United Nations Evaluation Group evaluation and gender standards	E.g., June 2015 Note: Timing and nature of evaluation will be determined by learning and performance needs based on testing the theory of change that underpins each outcome	Consider the following expenses: Evaluators and external advisers, and expenses related to their duties; expert advisory panel members (if any); travel; stakeholder consultations; data collection, and analysis tools and methods; supplies (office, computer, software, etc.); communication costs; publication and dissemination	E.g., project budget; donor; M&E budget; etc.

⁸² This template should be accompanied by a brief text explaining the rationale behind the plan.

⁸³https://popp.undp.org/_layouts/15/WopiFrame.aspx?sourcedoc=/UNDP_POPP_DOCUMENT_LIBRARY/Public/PPM_Formulat e%20Programmes%20and%20Projects_Evaluation%20Plan%20Template.docx&action=default

3.6. Step Five: Evaluation plan review and quality assurance process

All evaluation plans go through a pre-PAC and headquarters PAC review process. The reviewers use a checklist of requirements for the evaluation plan (see table 3),⁸⁴ to verify that the criteria and requirements for the content of the plan, as detailed above, have been fully considered and included.

Table 3. Evaluation plan checklist

#	Quality assurance criteria	YES	NO
1	Is the evaluation plan complete, i.e., noting the following? <ul style="list-style-type: none"> ▪ The commissioning unit ▪ Evaluation partners (only for joint evaluations) ▪ Evaluation type (programme, project, outcome, thematic, GEF, etc.) ▪ Planned evaluation completion dates ▪ Are evaluations aligned to UNSDCF and Strategic Plan outcomes? ▪ Estimated budget and source of the funding 		
2	Are all mandatory evaluations included? ⁸⁵ <ul style="list-style-type: none"> ▪ GEF terminal evaluations for all GEF-financed medium-size projects and full-size projects ▪ GEF midterm reviews for full-size projects ▪ Adaptation Fund and Green Climate Fund projects as required ▪ Donor/cost-sharing agreement evaluations ▪ Projects with a budget or expenditure of over \$5 million - midterm <i>and</i> final evaluation ▪ Projects with a budget or expenditure between \$3 million and \$5 million - midterm or final evaluation ▪ Projects with a duration of more than five years - midterm or final evaluation ▪ Projects entering a second or subsequent phase ▪ Development initiatives being considered for scaling up 		
3	Is there a brief text explaining the rationale for including the evaluations in the plan (maximum 300 words)?		
4	Is there inclusive and balanced coverage of the country programme content?		
5	Are the timing and sequencing of evaluations in the plan realistic?		
6	Does costing properly reflect the scope, depth and duration of each evaluation? Is it realistic?		

A further template is used to check the scope and balance of the evaluation plan (see table 4). By categorizing evaluations by year, type or outcome, the reviewer can quickly identify evaluation gaps, where lessons are not being captured or where a year may see significant bunching of evaluations and therefore pose implementation challenges.

⁸⁴ Access at:

<https://intranet.undp.org/unit/office/exo/sp2014/SP201417/PAC%20Library/2018/New%20guidance%20notes/Sec%203%20Template%20%20PAC%20Evaluation%20plan%20checklist.docx>

⁸⁵ Exceptions and further details can be found in section 2.6. UNSDCF evaluations do not need to be included in the UNDP evaluation plans. These evaluations are not under UNDP control and will not be quality assessed by UNDP IEO.

Table 4. Sample evaluation plan scheduling checklist

Number of evaluations planned							
		Year 1	Year 2	Year 3	Year 4	Year 5	Total
		2019	2020	2021	2022	(If applicable)	
Outcomes	Outcome 1 evaluation			1			1
	Outcome 1 project evaluations	1		1			2
	Outcome 2 evaluation			1			1
	Outcome 2 project evaluations						0
	Outcome 3 evaluation			1			1
	Outcome 3 project evaluations						0
GEF	GEF terminal evaluation	4	3	1	2		10
	GEF midterm evaluations						0
	Other evaluations						0
	TOTAL	5	3	5	2		15

3.7. Step Six: Evaluation plan completion and approval



The country office senior management team must review and endorse the evaluation plan before its submission to the headquarters PAC.

Once the evaluation plan has been finalized and endorsed through the pre-PAC and headquarters PAC, it accompanies the CPD as an annex and submitted to the Executive Board for approval.



Once the CPD and annexed evaluation plan have been approved by the Executive Board, the programme unit should upload the details of each evaluation to the ERC. The

Plan detail (2016-2020)	
Commissioning unit:	Indonesia
Period:	2016-2020
Status:	Posted
Comments:	UNDP CO Indonesia Evaluation Plan 2016-2020
Plan document:	30 May Indonesia Evaluation Plan.docx

evaluation plan should also be uploaded as a supporting document under the “Plan details” heading of the programme unit evaluations plan on ERC.⁸⁶

3.8. Step Seven: Making changes to the evaluation plan

Once an evaluation plan has been approved, and is entered in the ERC for tracking, the regional bureau will use the plan as a basis for monitoring compliance.



The evaluation plan is not a static document and may require adjustment as circumstances change.

Adjustments to individual evaluations and the evaluation plan should be considered annually as part of the programme unit’s stocktaking exercise. Changes that can be made with approval include:

- Extending the completion date for evaluations.
- Changing the scope and purpose of evaluations due to changes in the context (e.g., crisis settings).
- Addition of new evaluations. New projects may require new and additional evaluations that need to be included in the evaluation plan.
- Deletion (in exceptional circumstances).⁸⁷

Any adjustments to the plans including date changes, deletions and additions need to be clearly supported with a detailed rationale validated and approved by the regional evaluation focal point. As changes are made to the evaluation plan, it is also important to ensure that the overall goals, scope, coverage and timing remain reflective of the programme unit’s work, capture its results and are realistic for implementation.



The evaluation plan should be reviewed annually and refined and adjusted as needed. The annual country office business planning meeting at the beginning of the year offers a good opportunity to review the evaluation plan.

As part of the annual review, programme units should also ensure that all completed evaluations have been uploaded to ERC together with their management response, and that all management responses and key actions are up to date.

Programme units should discuss possible changes with regional evaluation focal points prior to making and requesting adjustments to plans through the ERC. Changes, particularly deletions, to individual project evaluations should be discussed, agreed and noted in minutes with project

⁸⁶ For more information, see the ERC user guide, <https://erc.undp.org/guidance>

⁸⁷ Evaluations can be deleted in instances such as: (a) evaluations were mistakenly added to the plan or ERC, such as duplicates; (b) the planned completion date is out of the country programme period, in which case the evaluation is deleted and added to the next evaluation plan; (c) evaluations are combined into other evaluations such as outcome, thematic or regional evaluations; (d) the funds available are too limited to make an evaluation usable or credible; and (e) the security, political, environmental, health or social situation is such that the evaluation cannot occur safely or meet its goals.

management boards or their equivalents such as a steering committee. Change requests can be made through the ERC by the M&E focal points. Regional evaluation focal points will review these requests and approve or reject as needed.

A formal **midterm review of the evaluation plan** is highly recommended. Changes to the evaluation plan during the midterm review ensure that:

- (a) the evaluation plan remains balanced and covers all aspects of the CPD in some way;
- (b) all completion dates are realistic and attainable;
- (c) all new evaluations have been included; and
- (d) all management responses and key actions are up to date.

Changes to evaluation plans are recorded and kept in the ERC and programme units can see a full picture of the changes and adjustments through the life of an evaluation plan. If a country programme period is extended, this change must also be reflected in the ERC.

3.9. Step Eight: Monitoring compliance

A programme unit M&E focal point, together with regional evaluation focal points, monitors the implementation of the evaluation plan to ensure that evaluations are completed, have management responses and that key actions are implemented.

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

Annually, IEO reports to the Executive Board on the number of evaluations planned during a given year, the number completed, the number of changes to evaluation plans and the reasons for those changes. IEO also reports on management responses to recommendations and key actions completed.

Table 5. Examples of well-balanced evaluation plans

Country	Decentralised Country Programme Evaluation	Outcomes	Other	Projects	GEF projects (terminal & midterm evaluations)	TOTAL
<u>Cambodia, 2019 to 2023</u>	1	3	1	8	5	18
<u>Philippines, 2019-2023</u>		2	1	7	10	20
<u>Ukraine, 2018-2022</u>	1	4		12	5	22
<u>Kenya, 2018-2022</u>	1	2		6	6	15



SECTION 4

EVALUATION IMPLEMENTATION AND USE



CONTENTS

4.1	Evaluation implementation	1
4.2	Step One: Pre-evaluation - initiating the evaluation process	7
4.3	Step Two: Evaluation preparation	11
4.3.1	Evaluation terms of reference	12
4.3.2	Supporting documentation for evaluations	18
4.3.3	Gender- and exclusion-sensitivity and rights-based approach	20
4.3.4	Choosing evaluators	21
4.4	Step Three: Managing an evaluation	26
4.4.1	Briefing the evaluation team	26
4.4.2	Supplying supporting information	27
4.4.3	Evaluation inception report	27
4.4.4	Evaluation and data collection mission	28
4.4.5	Draft report and review process	29
4.4.6	Evaluation review processes	30
4.4.7	Dispute settlement and reporting wrongdoing	31
4.5	Step Four: Using the evaluation	34
4.5.1	Preparing the management response for decentralized evaluations	34
4.5.2	Publication of the final evaluation report	36
	Annex 1. Evaluation TOR template	39
	Annex 2. Summary of common data-collection methods/sources used in UNDP evaluations	52
	Annex 3. UNDP Evaluation dispute resolution process	55
	Annex 4. UNDP evaluation report template and quality standards	56
	Annex 5. Management response template	61




Table 1. Evaluation process checklist.....	4
Table 2. Evaluability checklist	7
Table 3. Sources of information for an evaluation team	19
Table 4. Audit trail form template	31
Table 5. Sample evaluation matrix.....	51
Box 1: Role of the M&E focal point, specialist or officer	3
Box 2: Planning, monitoring and evaluation in a crisis setting	8
Box 3: UNDP quality standards for programming	13
Box 4: Integrating gender equality and women’s empowerment perspectives in an evaluation TOR	16
Box 5: Integrating disability issues in evaluations	17
Box 6: Sample TORs	18
Box 7: Evaluator Databases.....	22
Box 8: Sources of conflict of interest in evaluation	23
Box 9: Inception report content.....	28
Box 10: Terms of reference and recommendations	35
Figure 1. Key steps in the evaluation process.....	1
Figure 2. Steps in preparing an evaluation	11
Figure 3. The phases of managing an evaluation.....	26

4. EVALUATION IMPLEMENTATION AND USE

Section 4 provides detailed guidance on the implementation of decentralized evaluations, beginning with the roles and responsibilities of the evaluation manager and other actors. The following subsections include: pre-evaluation steps, such as checking the readiness for evaluation; preparing for the evaluation; managing the evaluation and the evaluation team; and using the evaluation, including preparation of the management response.

The process for developing evaluations commissioned by programme units includes the following four key steps, outlined in detail in this section.

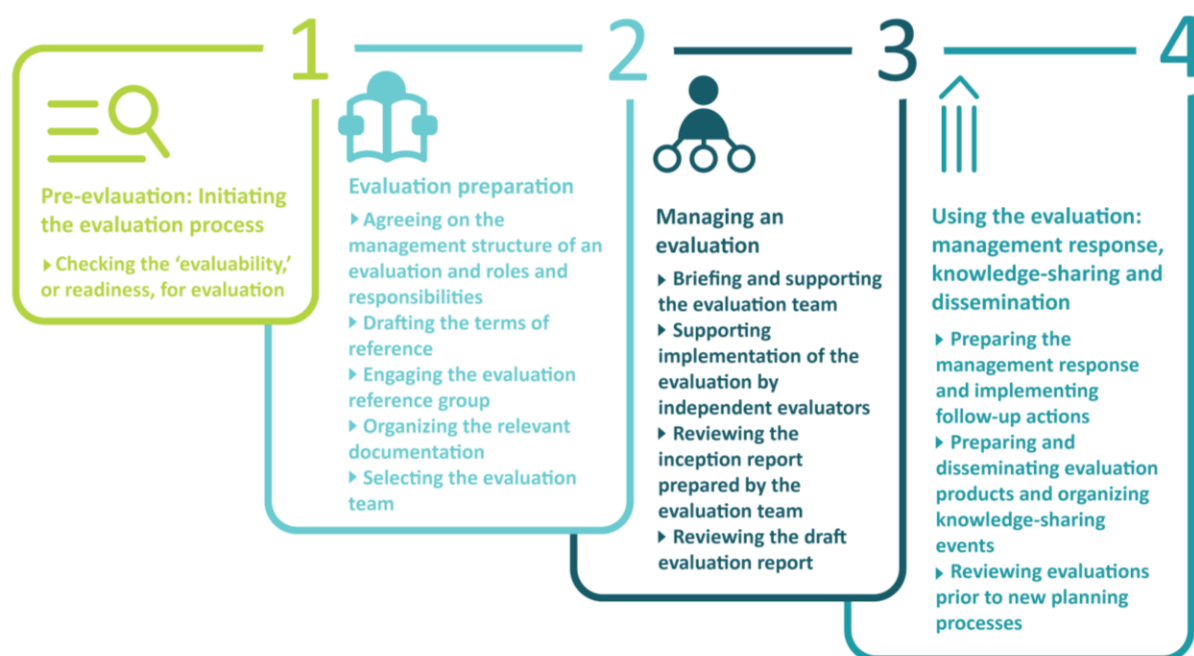


Figure 1. Key steps in the evaluation process

4.1 Evaluation implementation

Roles and Responsibilities



All evaluations should have a clearly defined organization and management structure, and well established and communicated roles and responsibilities, including an evaluation manager responsible for oversight of the whole evaluation process. Who this is will depend on the human resources available within the programme unit. To avoid conflicts of interest, the evaluation manager cannot be the manager of the programme/ project being evaluated.

This section defines and describes key members of the evaluation team.

Evaluation commissioner: in the context of these Guidelines, the evaluation commissioner is the agency or entity that calls for the evaluation to be conducted, in this case UNDP, and within UNDP, the senior manager that “owns” the evaluation plan under which the decentralized evaluation is being carried out. The evaluation commissioner, for example the resident representative for a country office, appoints the evaluation manager and approves the final terms of reference (TOR).

Programme/project manager: This is the manager responsible for the programme, outcome, portfolio or project under evaluation (the “evaluand”).¹ The programme/project manager should take a supporting role in the implementation of the evaluation but, in order to ensure independence and credibility, will not manage the evaluation. They will provide documents and data as requested, support the overall evaluation and evaluation manager, and have a clear plan for using the evaluation results.

Evaluation manager: Evaluation management should be separate from programme/project management. Where the UNDP implementing office has a monitoring and evaluation (M&E) specialist or focal point, they should take the evaluation management role. Where there is no such position, an evaluation manager should be assigned by senior management (e.g. the resident representative).

The evaluation manager can recommend final sign-off and approval of all aspects of the evaluation process including: (a) ensuring evaluability; (b) the evaluation TOR; (c) the evaluation team structure and recruitment; (d) the inception report; (e) coordinating comments on the draft evaluation report; and (f) the final evaluation report.

For a joint evaluation, there may be a co-commissioner and co-manager from the partner agency. The evaluation management structure, roles and responsibilities should be agreed prior to the evaluability stage of the evaluation process.

¹ Typically, this includes senior management for country programme evaluations, global programme managers for global programme evaluations, outcome leads for outcome evaluations and/or programme officers (programme team leaders, programme analysts) for project evaluations.

Box 13: Role of the M&E focal point, specialist or officer

Whether or not the M&E focal point/ specialist/ officer is the evaluation manager, they should still ensure the quality of all evaluations - outcome, project, vertical-funded projects (Global Environment Facility [GEF] and Green Climate Fund [GCF]), donor project evaluations, etc.

The M&E focal point/ specialist/ officer should approve each stage before moving to the next, including:

- Developing and reviewing the evaluation TOR, ensuring that they meet UNDP guidance requirements;
- Reviewing and approving the evaluation inception report, ensuring that it meets UNDP requirements;
- Reviewing and recommending acceptance of the draft and final evaluation reports; and
- Reviewing the management responses and key actions.

In addition, the M&E focal point or specialist maintains the programme unit evaluation plan on the Evaluation Resource Center (ERC), including:

- Uploading the evaluation plan and updating as required;
- Managing changes to the evaluation plan and getting approval from the regional evaluation focal point;
- Uploading evaluation documents (TOR, evaluation reports etc.) to the ERC within the timelines outlined;
- Uploading management responses and key actions and updating on a quarterly basis; and
- Reporting to management on compliance with the evaluation plan, completion of management responses and key actions and results of the quality assessment.

Evaluation reference group: The evaluation commissioner and evaluation manager should consider establishing an evaluation reference group made up of key partners and stakeholders who can support the evaluation and give comments and direction at key stages in the process. An evaluation reference group ensures transparency in the evaluation process and strengthens the credibility of the results.

Regional evaluation focal points oversee the implementation of country office evaluation plans, approve any adjustments to the plans with valid justification, and ensure that country offices meet the evaluation commitments made in the plans. The regional evaluation focal point also offers technical guidance on the implementation of evaluations to country offices, primarily to their management and M&E focal points or specialists, to ensure that commitments under evaluation plans are met and that evaluations are credible, independent and of the required quality. Evaluation focal points at central bureau level have the same role, overseeing central bureau evaluation plans and changes uploaded to the ERC.

In country offices where there is no dedicated M&E officer or specialist, the regional evaluation focal points should provide additional support to the assigned M&E focal points. Technical support can include: advice on the development of TORs, including the integration of gender equality perspectives; recruitment of evaluators; feedback on inception reports; implementation of evaluations; finalization of evaluations; and feedback on draft evaluation reports and management responses. Regional evaluation focal points are the main contacts when disputes arise in the evaluation process.

More details of roles and responsibilities in evaluation implementation can be found in section 5.

Table 1 details the roles and responsibilities and expected completion schedules for the entire evaluation process.

Table 1. Evaluation process checklist

STEP	ACTIVITY	TIME SCHEDULE	RESPONSIBILITY
ONE	Evaluability check	Six months before proposed commencement	<ul style="list-style-type: none"> ▪ Evaluation commissioner ▪ Evaluation manager ▪ M&E specialist/ officer or focal point ▪ Programme/ project officer
	Draft TOR	Three to six months before proposed commencement	<ul style="list-style-type: none"> ▪ Evaluation commissioner ▪ Evaluation manager ▪ M&E specialist/ officer or focal point ▪ Evaluation reference group ▪ Programme/ project officer
	Final TOR	❖ Uploaded to ERC two weeks after completion of the TOR	<ul style="list-style-type: none"> ▪ M&E specialist or focal point
TWO	Recruit evaluation team	One month prior to proposed commencement or earlier	<ul style="list-style-type: none"> ▪ Evaluation commissioner ▪ Evaluation manager ▪ M&E specialist or focal point ▪ Operations team ▪ Programme/ project officer
	Inception report review	According to the TOR (two to four weeks after contract signing)	<ul style="list-style-type: none"> ▪ Evaluation commissioner ▪ Evaluation manager ▪ M&E specialist/ officer or focal point ▪ Evaluation reference group ▪ Programme/ project officer
THREE	Data collection and field visits	According to the TOR and inception report	<ul style="list-style-type: none"> ▪ Evaluation team
	Draft report review	Immediately on reception according to the TOR and inception report	<ul style="list-style-type: none"> ▪ Evaluation commissioner ▪ Evaluation manager ▪ M&E specialist or focal point ▪ Evaluation reference group ▪ Programme/ project officer
	Audit report and comments	According to the TOR and inception report	<ul style="list-style-type: none"> ▪ Evaluation team
	Final report completion	According to the TOR and inception report	<ul style="list-style-type: none"> ▪ Evaluation team
	Final report uploaded to the ERC	Uploaded to ERC within two weeks of receipt	<ul style="list-style-type: none"> ▪ M&E specialist or focal point
	Management response and key actions	Project and outcome evaluations: within six weeks of the final report	<ul style="list-style-type: none"> ▪ Evaluation manager ▪ Evaluation reference group ▪ Programme/ project officer
	Final management response	Uploaded to ERC within six weeks of receipt of the final evaluation report	<ul style="list-style-type: none"> ▪ M&E specialist or focal point
	Quarterly follow-up on key actions	Update ERC at the end of every quarter	<ul style="list-style-type: none"> ▪ Evaluation manager ▪ M&E specialist or focal point based on inputs provided by programme units

	Management response and key actions closed	When all planned actions have been completed or after five years	▪ M&E specialist or focal point
--	--	--	---------------------------------



Timings and schedules for each stage can be set by the programme units. However, the dates for completion and uploading to the ERC are set.

Step One: Evaluability



4.2 Step One: Pre-evaluation - initiating the evaluation process

Checking the “evaluability” or readiness of a programme/ project for evaluation



An evaluability assessment examines the extent to which a project, programme or other intervention can be evaluated in a reliable and credible way. It calls for the early review of a proposed project, programme or intervention in order to ascertain whether its objectives are adequately defined and its results verifiable.

UNDP programme units and stakeholders should undertake an evaluability assessment of any proposed evaluation (six months) prior to its commencement, to ensure that the activity (whether a programme, outcome, project, portfolio or thematic area) is in a position to be evaluated. This should be undertaken jointly by the evaluation commissioner, evaluation manager and/ or M&E focal point. Key stakeholders in the project, especially national counterparts, should be fully involved in the development of an evaluation and contribute to the evaluation design and results, including the evaluability assessment.

Table 2 provides a checklist which can guide the evaluability assessment, and highlights areas that may need to be improved and strengthened for an evaluation to move ahead.

Table 4. Evaluability checklist

		Y	N
1.	Does the subject of the evaluation have a clearly defined theory of change ? Is there common understanding as to what initiatives will be subject to evaluation?		
2.	Is there a well-defined results framework for the initiative(s) that are subject to evaluation? Are goals, outcome statements, outputs, inputs and activities clearly defined? Are indicators SMART? ⁸⁹		
3.	Is there sufficient data for evaluation ? This may include baseline data, data collected from monitoring against a set of targets, well-documented progress reports, field visit reports, reviews and previous evaluations.		
4.	Is the planned evaluation still relevant , given the evolving context? Are the purpose and scope of the evaluation clearly defined and commonly shared among stakeholders? What evaluation questions are of interest to whom? Are these questions realistic, given the project design and likely data availability and resources available for the evaluation?		
5.	Will political, social and economic factors allow for effective implementation and use of the evaluation as envisaged?		
6.	Are there sufficient resources (human and financial) allocated to the evaluation?		

⁸⁹ Specific, Measurable, Assignable, Relevant and Time-bound.



If the answer to one or more of **questions 1 to 3** above is ‘no’, the evaluation can still go ahead. The programme unit management, evaluation commissioner, evaluation manager and/ or the M&E focal point or specialist and stakeholders will need to make the appropriate adjustments and updates to bring the programme/ project into a position to be evaluated (which may cause implementation delays). Working with implementing partners, results models and frameworks and overall documentation should be brought up to date. **A well-managed and monitored programme/ project should have these prerequisites in place by the time of the evaluation.**

The relevance of an evaluation (question 4) may be a consideration where a project or outcome area has been reduced in importance due to resource mobilization limitations or changes in the organizational or country context that have led to a reduced focus for UNDP.

If **political and socioeconomic situations (question 5)** do not allow the team to carry out an evaluation in a meaningful manner, UNDP management, together with national stakeholders, may decide to wait for a more conducive environment to be secured. The evaluation may need to be flexible in its data collection approach and methodology to accommodate issues that arise (for example changing field visit sites). In crisis settings (see Box 2), such decisions should be made based on good, current analyses of the context, to ensure that the evaluation will be relevant to fast-changing situations. Factors such as security situations (safety of evaluators, staff and interviewees) and the potential impact of the evaluation on existing tensions should be carefully assessed.

Box 2: Planning, monitoring and evaluation in a crisis setting

If an initiative is being implemented in a crisis setting (relating to conflicts and disasters), this will have ramifications for all aspects of programming including planning, monitoring and evaluation. In general, the planning and M&E methods and mechanisms presented in these guidelines are transferable to crisis settings, with several important caveats:

- **Crisis situations are dynamic**, and UNDP programming should quickly respond to radical changes that can take place in such circumstances. Therefore, the situation should continually be analysed and monitored to ensure that programming remains relevant. Changes should be documented so that monitoring and evaluation of the relevance and appropriateness of development initiatives take into consideration the fluid situations in which they were conceived and implemented. This will involve continuous situational and conflict analysis.
- **Crisis situations are characteristically of (potentially) high tension** between different parties. Thus, crisis- and conflict-sensitivity should be exercised in all aspects of programming, including planning, monitoring and evaluation, to ensure that both the substance and process reduce, or at least do not heighten, tensions between different parties. The security of programme staff, beneficiaries and M&E staff can be a constant concern, and risk analysis for all those involved should be constantly monitored and factored into M&E activities.
- **It is important to keep a “big picture” perspective**, considering how projects and programmes connect to the wider peace process is critical, particularly for conflict prevention and peacebuilding programming. Planning, monitoring and evaluation should always factor this in to avoid a situation where a project is “successful” in terms of meeting the desired results, but has no – or negative - impacts on wider peace.

The 'Compendium on Planning, Monitoring and Evaluation in Crisis Prevention and Recovery Settings' provides further guidance.⁹⁰ Other resources are also available to support evaluation in crisis and humanitarian contexts.⁹¹

Covid-19 specific guidance is also available, providing tools and approaches for undertaking evaluations in fragile and crisis settings.⁹²

Finally, sufficient resources (**question 6**) should have been assigned at the time of the design and approval of the country programme document (CPD) and evaluation plan. Where adequate resources are not available for the full scope of an evaluation, it is more prudent to delay implementation until adequate resources are available than to push ahead with an evaluation that is under-resourced and likely to suffer from reduced scope, utility and credibility.

Delaying an evaluation: If a project, programme or outcome is found not to be ready for evaluation, and a delay is required, adjustments can be made to the evaluation plan with a new evaluation completion date. The adjustment and justification should be submitted to the ERC for review and approval by the regional evaluation focal point.

Deletion of an evaluation: Programme units should make every effort to implement all evaluations in an evaluation plan. Only in exceptional circumstances should an evaluation be deleted from an evaluation plan (see section 3.8). If an evaluation is believed to be no longer relevant or is not expected to meet evaluability requirements, then UNDP senior and programme unit management 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 implementation.

⁹⁰ ALNAP, 2016, "Evaluation of Humanitarian Action Guide", <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>

⁹¹ The ALNAP network has further guidance and tools for evaluation in crisis settings on its website: <https://www.alnap.org/>

⁹² <http://web.undp.org/evaluation/guideline/covid19.shtml>



Step Two: Evaluation preparation



4.3 Step Two: Evaluation preparation



Figure 2. Steps in preparing an evaluation

Budgets and sources of funding for an evaluation should be agreed with partners during the drafting of the evaluation plan, and detailed in the plan.

- **Project evaluation budgets** should be detailed in project and programme documents. GEF projects have suggested budgets for midterm reviews and terminal evaluations.
- **Outcome evaluation budgets** can come from country office funds or be part-funded by individual projects and programmes.



Budgets should be realistic and enable credible and independent evaluations that produce usable results for the organization. A reduced or limited budget will limit the scope and depth of an evaluation and could limit its use and credibility. The annual report on evaluation from the Independent Evaluation Office (IEO) gives average budgets for different evaluation approaches globally and by region. These can be used as a reference.⁹³

Individual evaluation budget considerations include:

- **Professional fees** for all evaluators or thematic experts undertaking the evaluation. There are often additional costs when hiring a professional firm.
- **Travel costs**, including flights to and from the evaluation country, where applicable, and travel within the country (for the evaluator, translator, UNDP accompanying staff and other participants).
- Additional and **non-professional costs** such as daily subsistence allowance for time in country for data collection and terminal expenses.
- Any **meeting costs** related to workshops (stakeholder, validation or evaluation reference group workshops) and focus group or data collection meetings (such as venue hire, snacks, participant transport costs etc.).
- **Translation costs** for interviews, field visits, and validation and dissemination workshops.
- **Communications costs** including editing, publication and dissemination costs.
- Additional **contingency costs** for unknown expenses arising during the evaluation.

Section 3 of this guidance includes an evaluation budget template.

4.3.1 Evaluation terms of reference



The TOR is a written document which defines the scope, requirements and expectations of the evaluation and serves as a guide and point of reference throughout the evaluation.

Quality TOR should be explicit and focused, providing a clear mandate for the evaluation team on *what* is being evaluated and *why*, *who* should be involved in the evaluation process, and the expected *outputs*. **TORs should be unique to the specific circumstances and purpose of each individual evaluation.** Since the TOR play a critical role in establishing the quality criteria and use of the evaluation report, adequate time should be allocated to their development.

The outcome, project, thematic area, or any other initiatives selected for evaluation, along with the timing, purpose, duration, available budget and scope of the evaluation, will dictate much of the substance of the TOR. However, because an evaluation cannot address all issues, developing the TOR involves strategic choices about the specific focus, parameters and outputs for the evaluation, given available resources.

The initial draft TOR should be developed by the evaluation manager with input from the evaluation commissioner and shared with the evaluation reference group for review and comment. Regional evaluation focal points and others with the necessary expertise may comment on the draft TOR to ensure that they meet corporate quality standards.

Writing TORs and engaging relevant stakeholders can be a time-consuming exercise. Therefore, it is recommended that this process is started **three to six months before** the proposed *commencement* of the evaluation, depending on the scope and complexity of the evaluation and the numbers of stakeholders involved.

The TOR template is intended to help UNDP programme units create TORs based on quality standards for evaluations consistent with evaluation good practice. When drafting TORs, programme units should consider how the evaluation covers UNDP quality standards for programming, as relevant and required (see Box 3).⁹⁴

The TOR should retain enough flexibility on the evaluation methodology for the evaluation team to determine the best methods and tools for collecting and analysing data. For example, the TOR might suggest using questionnaires, field visits and interviews, but the evaluation team should be able to revise the approach in consultation with the evaluation manager and key stakeholders. These changes in approach should be agreed and reflected clearly in the inception report.

⁹⁴ Access at:

https://popp.undp.org/layouts/15/WopiFrame.aspx?sourcedoc=/UNDP_POPP_DOCUMENT_LIBRARY/Public/PPM_Programming%20Standards_Quality%20Standards%20for%20Programming.docx&action=default

Box 14: UNDP quality standards for programming

Strategic

Programming priorities and results contribute to the Sustainable Development Goals (SDGs), are consistent with the UNDP Strategic Plan and aligned with United Nations Sustainable Development Cooperation Frameworks (UNSDCFs). Programmes and projects are based on clear analysis, backed by evidence and theories of change. The latter justify why the defined approach is most appropriate and will most likely achieve, or contribute to, desired development results along with partner contributions. The role of UNDP vis-à-vis partners is deliberately considered. New opportunities and changes in the development context are regularly reassessed, with any relevant adjustments made as appropriate.

Relevant

Programming objectives and results are consistent with national needs and priorities, as well as with feedback obtained through engaging excluded and/ or marginalized groups as relevant. Programming strategies consider interconnections between development challenges and results. A gender analysis is integrated to fully consider the different needs, roles and access to/ control over resources of women and men, and appropriate measures are taken to address these when relevant. Programmes and projects regularly capture and review knowledge and lessons learned to inform design, adapt and change plans and actions as appropriate, and plan for scaling up.

Principled

All programming applies the core principles of human rights, gender equality, resilience, sustainability and leaving no one behind. Social and environmental sustainability are systematically integrated. Potential harm to people and the environment is avoided wherever possible, and otherwise minimized, mitigated and managed. The complete Social and Environmental Standards can be found [here](#).

Management and monitoring

Outcomes and outputs are defined at an appropriate level, are consistent with the theory of change, and have SMART, results-oriented indicators, with specified baselines and targets and identified data sources. Gender-responsive, sex-disaggregated indicators are used when appropriate. Relevant indicators from the Strategic Plan integrated results and resources framework have been adopted in the programme or project results framework. Comprehensive, costed M&E plans are in place and implemented to support evidence-based management, monitoring and evaluation. Risks, in terms of both threats and opportunities, are identified, with appropriate plans and actions taken to manage them. Governance of programmes and projects is defined, with clear roles and responsibilities, and provides active and regular oversight to inform decision-making.

Efficient

Programming budgets are justifiable and valid and programming design and implementation includes measures to ensure the efficient use of resources. The size and scope of programmes and projects are consistent with available resources and resource mobilization efforts. Plans include consideration of scaling up and links with other relevant initiatives to achieve greater impact. Procurement planning is done early and regularly reviewed. Monitoring and management include analysis, and actions, to improve efficiency in delivering the desired outputs with the required quality and timeliness, such as country office support to national implementation modalities. Costs are fully recovered (see the [cost-recovery policy](#)).

Effective

Programming design and implementation are informed by relevant knowledge, evaluation and lessons learned to develop strategy and inform course corrections. Targeted groups are systematically identified and engaged, prioritizing the marginalized and excluded. Results consistently respond to gender analysis and are accurately rated by the gender marker. Managers use monitoring data for making decisions that maximize the achievement of desired results. South-South and triangular cooperation are used when relevant and captured in the results framework. Required implementing partner assessments have been conducted and the implementation modality is consistent with the results.

Sustainability and national ownership

Programming is accomplished in consultation with relevant stakeholders and national partners, who are engaged throughout the programming cycle in decision-making, implementation and monitoring. Programming includes assessing and strengthening the capacity and sustainability of national institutions. A strategy for use of national systems is defined and implemented, if relevant. Monitoring includes the use of relevant national data sources, where possible. Sustainability of results is accomplished through tracking capacity indicators and implementing transition and scale-up plans.

The TOR should, at a minimum, cover the elements described below which are outlined in more detail in annex 1.

1. Background and context

- Detail the social, economic, political, geographic and demographic factors (at the time of programme / project design *and* of the evaluation).
- Clearly describe the intervention to be evaluated (scale and complexity).
- Provide details of project beneficiaries (gender, disability, vulnerable groups, and human rights issues).
- Present outcomes, outputs, main achievements, results and the theory of change.

2. Evaluation purpose, scope and objectives

- Detail *why* the evaluation is being conducted, *who* will use or act on the evaluation findings and recommendations and *how*.
- Define the parameters and focus of the evaluation.
- **Include gender equality and women's empowerment, disability and other cross-cutting issues** (see box 4).

3. Evaluation criteria and key questions

- The Organization for Economic Cooperation and Development (OECD) Development Assistance Committee (DAC) criteria of relevance, coherence,⁹⁵ effectiveness, efficiency and sustainability (see Section 2) can be used to frame the evaluation questions.⁹⁶
- Include specific questions to be answered through the evaluation. All questions should be clear, well defined and manageable.
- **Include at least one evaluation question related to gender equality and women's empowerment, and ideally at least one *per evaluation criterion*.**
- Include evaluation questions related to disability and other cross-cutting issues.⁹⁷

4. Methodology

- Consider the overall approach and methodology.

⁹⁵ Coherence' is a new OECD DAC evaluation criteria introduced in 2019. For UNDP, 'coherence' is not a required criterion, but could be considered, if found pertinent.

⁹⁶ Access at: <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>. We recommend that relevance, coherence, effectiveness, efficiency and sustainability are covered to ensure the full scope of an evaluation. If the commissioning unit chooses not to use one or more of the criteria, this should be explicitly stated in the TOR, including the reasons for the omission.

⁹⁷ Technical Notes. Entity Accountability Framework of the United Nations Disability Inclusion Strategy. Access at: <https://www.un.org/en/disabilitystrategy/resources>

- Include data sources, suggested collection methods and analysis approaches (with flexibility for refinement by the evaluation team).
- Refer to funding analysis requirements and sources of funding.
- **Methodologies for addressing gender-specific issues as well as inclusion of the SDGs should be requested from the evaluation teams.**

5. Evaluation products (key deliverables)

- Details key evaluation products, which may include:
 - (a) an evaluation inception report, including a workplan and evaluation schedule;
 - (b) a draft evaluation report for comment;
 - (c) an audit trail detailing how comments, questions and clarifications have been addressed;
 - (d) a final report (addressing comments, questions and clarifications); and
 - (e) any presentations or other knowledge products.
- **All evaluation products need to address gender, disability and human rights issues.**

6. Evaluation team composition and required competencies

- State the structure of the evaluation team and the number of evaluators required.
- Detail specific skills, knowledge and expertise, competencies and characteristics required of the evaluator and each member of the evaluation team.
- **Gender expertise and competencies are a must.**
- At least one evaluation team member or reference group member needs to have knowledge and/ or experience of disability inclusion.

7. Evaluation ethics

- Evaluation consultants will be held to the highest ethical standards and are required to sign a ***pledge of ethical conduct*** upon acceptance of the assignment.
- UNDP evaluations are conducted in accordance with the principles outlined in the United Nations Evaluation Group (UNEG) Ethical Guidelines for Evaluation.⁹⁸

8. Management and implementation arrangements

- Describe the management and implementation structure and define the roles, key responsibilities and lines of authority of all parties involved in the evaluation process.
- Outline the feedback mechanisms for the different evaluation products.

9. Time frame for the evaluation process

- Details the number of days planned for the evaluation and the number of days allocated across the evaluation team.
- Should also cover an evaluation timetable including dates for: (a) the start of the evaluation; (b) the evaluation deliverables; (c) fieldwork and data collection; and (d) completion of the evaluation.

10. Submission process and basis for selection

- Details the structure and procedures for application, supporting documents and submission documents required and the criteria for review of applications.

⁹⁸ Access at: <http://www.unevaluation.org/document/detail/2866>

11. TOR annexes - links to supporting background documents and UNDP evaluation guidelines

- Intervention results framework and theory of change.
- Key stakeholders and partners.
- Documents to be reviewed and consulted.
- Evaluation matrix template.
- Outline of the evaluation report format.
- Pledge of ethical conduct forms.
- UNDP Evaluation Guidelines, highlighting
 - Inception report template (section 4)
 - Evaluation report template and expected content (Section 4)
 - Quality Assessment process (Section 6)



The TOR will be assessed in the evaluation quality assurance. All points summarized above conform to the rating system (Annex 6).

Box 15: Integrating gender equality and women's empowerment perspectives in an evaluation TOR

In principle, all evaluations conducted or commissioned by UNDP must integrate human rights and gender equality, and aim to meet the requirements of the United Nations System-wide Action Plan (UN-SWAP) on Gender Equality and the Empowerment of Women Evaluation Performance Indicators⁹⁹ (see section 6 for more details).

IMPORTANT: Integrating gender equality and women's empowerment in the scope of the evaluation, as expressed in the TOR, is a critical first step. A gender-responsive evaluation should be carried out, even if the subject of evaluation was not gender-responsive in its design.

The UNEG guidance document, 'Integrating Human Rights and Gender Equality in Evaluations',¹⁰⁰ provides examples of how to incorporate these elements into the definition of the evaluation purpose, objectives, context and scope, and how to add a gender dimension to the standard evaluation criteria. Examples of tailored evaluation questions are also provided. (Annex 2)

What makes an evaluation gender-responsive?

Gender-responsiveness includes and relates to both *what* the evaluation examines and *how* it is undertaken. This means:

- assessing the degree to which gender and power relationships, including structural and other causes of inequities, discrimination and unequal power relations, change as a result of an intervention; and
- using a process that is inclusive, participatory and respectful of all stakeholders (rights holders and duty bearers).

A gender-responsive evaluation promotes accountability to gender equality, human rights and women's empowerment commitments by providing information on the way in which development programmes are affecting women and men differently and contributing towards the achievement of these commitments. It is

⁹⁹ Access at: <https://www.unwomen.org/en/how-we-work/un-system-coordination/promoting-un-accountability/key-tools-and-resources>

¹⁰⁰ Access at: <http://www.uneval.org/document/detail/980>

applicable to all types of development programming, not just gender-specific work (See UN Women Independent Evaluation Office, 2015, How to Manage Gender Responsive Evaluation. Evaluation Handbook).¹⁰¹

What does a gender-responsive evaluation methodology include?

- The evaluation specifies how gender issues are addressed in the methodology, including how data collection and analysis methods integrate gender considerations and ensure that data collected is disaggregated by sex.
- The evaluation methodology employs a mixed-methods approach, appropriate to evaluating gender equality and women's empowerment considerations.
- The evaluation employs a diverse range of data sources and processes (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility.
- The evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate.
- Ethical standards are considered throughout the evaluation and all stakeholder groups are treated with integrity and respect for confidentiality.¹⁰²

IMPORTANT: A gender equality and women's empowerment perspective needs to be included throughout the TOR.

Where applicable, evaluations need to consider disability issues deriving from the United Nations Disability Inclusion Strategy and Technical Notes.¹⁰³ Box 5 details the elements to be addressed:

Box 5: Integrating disability issues in evaluations

- The evaluation TOR pays adequate attention to disability inclusion.
- Evaluation teams have knowledge and/ or experience of disability inclusion where relevant.
- Evaluation questions cover different aspects of disability inclusion.
- Evaluation stakeholder mapping and data collection methods involve persons with disabilities and their representative organizations.
- Evaluation findings and analysis provide data and evidence on disability inclusion.
- The conclusions and recommendations of evaluations reflect findings on disability inclusion.

GEF terminal evaluations and midterm reviews have their own TOR structure and requirements, which cover the above with additional GEF-specific requirements. These requirements are outlined in the GEF guidelines for terminal evaluations and midterm reviews.¹⁰⁴

All TORs undergo a post-evaluation quality assessment conducted by IEO through the ERC quality assessment process. Details of the five quality assessment questions for TORs are given in section 6.

¹⁰¹ Access at: <https://www.unwomen.org/en/digital-library/publications/2015/4/un-women-evaluation-handbook-how-to-manage-gender-responsive-evaluation>

¹⁰² Further guidance on key elements of a gender-responsive evaluation methodology is provided on p. 37-41 and p. 91-110 in the UNEG Guidance Integrating Human Rights and Gender Equality in Evaluations. Access at: <http://www.uneval.org/document/detail/980>

¹⁰³ Access at: <https://www.un.org/en/content/disabilitystrategy/>

¹⁰⁴ Access at: <http://web.undp.org/evaluation/guidance.shtml>

All TORs are uploaded to the ERC and can be searched by evaluation type as well as their quality assessment ratings. Examples of TORs that can be used as a guide are detailed in Box 6 below.¹⁰⁵

Box 16: Sample TORs

Country programme evaluation TORs

- Nepal, 2020, [Mid Term Review of Country Programme Document](#)

Outcome evaluation TORs

- El Salvador, 2019, [Evaluación de la contribución del PNUD a la generación de capacidades de planificación y ejecución de la inversión pública a nivel nacional y local y su efecto en la cobertura y acceso universal a los servicios básicos de salud"](#)
- Tanzania, 2019, [Mid-term evaluation of democratic governance outcome](#)
- Kazakhstan, 2018, [Kazakhstan Country Programme Outcome Evaluation 2016-2018 Diversification of the economy provides decent work opportunities for the underemployed, youth, and socially vulnerable women and men](#)

Project evaluation TORs

- Bangladesh, 2020, [Final Evaluation of Partnership for a Tolerant, Inclusive Bangladesh \(PTIB\) project](#)
- Nepal, 2020, [Final Evaluation of Resilient Reconstruction and Recovery of Vulnerable Communities Severely Affected by 2015 Earthquake \(EU II\)](#)
- Haiti, 2020, [Évaluation finale du projet Promotion de la Cohésion sociale à Jérémie](#)
- DPK Korea, 2020, [End of Project Evaluation: Strengthening the Resilience of Communities through Community-Based-Disaster Risk Management](#)

4.3.2 Supporting documentation for evaluations

Once the scope of an evaluation has been defined, the evaluation manager, with help from the project or programme manager, will compile basic documentation that will be provided to the evaluation team. Preliminary desk work may be carried out to gather information on the activities and outputs of partners, previous UNDP-related assistance and the current context of the project, programme or outcome. Table 3 provides more detail on the types and sources of information to be provided.

¹⁰⁵ The sample TORs provided in Box 6 were chosen due to their quality assessment ratings. Only TORs with a rating of satisfactory (5) or highly satisfactory (6), were chosen. Efforts will be made to update these annually.

Table 5. Sources of information for an evaluation team

SOURCES OF INFORMATION	DESCRIPTION OF INFORMATION
Country, regional and global programme results frameworks and theories of change	<ul style="list-style-type: none"> ▪ Addressing the key outcomes that UNDP plans to achieve in a three- to five-year period ▪ Relevant theories of change at country and regional levels ▪ CPDs provide background information and UNDP perspective on development in a given country
Monitoring (regular reporting, reviews, steering committee meeting minutes) and evaluation reports	<ul style="list-style-type: none"> ▪ Evaluation reports on related subjects commissioned by IEO, UNDP programme units, government or other development partners and stakeholders ▪ Annual and quarterly progress reports, field visit reports, and other outcome and key programme or project documentation ▪ The ERC can be used to search for relevant evaluations carried out by other UNDP units on similar topics
Reports on progress of partner initiatives	<ul style="list-style-type: none"> ▪ Reports which show progress made by partners towards achieving the same outcome, and information about how they have strategized their partnership with UNDP
Data from official sources	<ul style="list-style-type: none"> ▪ Information on progress towards outcome achievements may be obtained from sources from the government, private sector, academia and national, regional and international research institutes, including those in the United Nations system ▪ In many cases, nationally adopted DevInfo systems and the websites of national statistical authorities are good sources for national statistics ▪ Data disaggregated by sex and other relevant social categories should be provided whenever available
Research papers	<ul style="list-style-type: none"> ▪ Topics related to the outcome being evaluated may have been addressed in research papers from the government, non-governmental organizations (NGOs), international financial institutions and academia
National, regional and global reports	<ul style="list-style-type: none"> ▪ Data can be found in various reports such as the national Human Development Report (HDR), national SDG report and other reports published by national, regional and subregional organizations, international financial institutions, and United Nations organizations ▪ National strategies and sectoral strategies as well as progress reports ▪ Reports on the status of gender equality and the empowerment of women may be useful, such as the <u>Gender Inequality Index</u> of the HDR¹⁰⁶
Financial and management information	<ul style="list-style-type: none"> ▪ Relevant reports from Atlas, audit, Corporate Planning System, Executive Snapshot, Integrated Work Plan, Transparency Dashboard (International Aid Transparency Initiative) ▪ A number of corporate tools provide financial and other management information that is relevant to evaluation. They include delivery, resource mobilization and human resource management ▪ Programme and project quality assurance reports

¹⁰⁶Access at: <http://hdr.undp.org/en/composite/GII>

Additional sources at country level

Reports of related regional and subregional projects and programmes	<ul style="list-style-type: none"> These reports indicate the extent to which these projects and programmes have complemented contributions by UNDP and partners to progress towards the outcome
CPD or United Nations Development Assistance Partnership (UNDAP) in full Delivering as One countries and results-oriented annual report	<ul style="list-style-type: none"> The results-oriented annual report provides a summary of the contributions of projects, programmes, sub-programmes and soft assistance that contribute to each outcome, on an annual basis Also included is information on key outputs, strategic partners, the partnership strategy, how much progress has been reported in previous years, the quality of outcome indicators, the need for further work and baseline information
UNDAF/ UNSDCF, UNDAP and country programme annual reviews and Common Country Assessments (CCAs) as well as evaluations	<ul style="list-style-type: none"> These documents include baseline information on the country's development situation, partnerships and joint activities of UNDP and other United Nations organizations

4.3.3 Gender- and exclusion-sensitivity and rights-based approach

UNDP evaluations are guided by the principles of equity, justice, gender equality and respect for diversity.¹⁰⁷ As appropriate, UNDP evaluations assess the extent to which initiatives have addressed these principles through their various development efforts in a given country; and incorporated the UNDP commitment to gender mainstreaming and the rights-based approach in their design and implementation.

Mainstreaming a gender perspective is the process of assessing the implications for women and men of any planned action, including legislation, policies or programmes, in all areas and at all levels. It is a strategy for making gender equality-related concerns an integral dimension of the design, implementation, monitoring and evaluation of policies and programmes in all political, economic and societal spheres, so that women and men benefit equally, and inequality is not perpetuated.

UNDP evaluations should assess the extent to which UNDP initiatives have considered addressing gender equality issues in their design, implementation and outcomes, and whether women and men can equally access the programme's benefits to the degree intended. Similarly, evaluations should address the extent to which UNDP has advocated for the principles of equality and inclusive development, considering disability issues, and has contributed to empowering and addressing the needs of the most disadvantaged and vulnerable populations in a given society.

The rights-based approach to development requires ensuring that development strategies facilitate the claims of rights-holders and the corresponding obligations of duty-bearers. This approach emphasizes the need to address the immediate, underlying and structural causes for such rights not being realized. The concept of civic engagement, as a mechanism to claim rights, is an important aspect in the overall framework. When appropriate, evaluations should assess the extent to which an initiative has facilitated the capacity of rights-holders to claim their rights, and of duty-bearers to fulfil their obligations.

¹⁰⁷ UNDP Evaluation Policy 2019, <http://web.undp.org/evaluation/policy.shtml>

Evaluations should also address other cross-cutting issues, depending on the focus of the evaluation.

4.3.4 Choosing evaluators

The choice of evaluators is paramount to the quality and credibility of an evaluation. UNDP selects evaluators through a competitive and transparent process in accordance with the organization's rules and regulations for procurement.¹⁰⁸ Areas of expertise to be considered in the team composition include:

- Proven expertise and experience in conducting and managing evaluations.
- Proven experience in data analysis and report writing.
- Technical knowledge and experience in UNDP thematic areas, with specifics depending on the focus of the evaluation.
- Technical knowledge and experience in cross-cutting areas such as gender equality, disability issues, the rights-based approach and capacity development.
- Knowledge of the national/ regional situation and context.

Depending on the scope and resources for the evaluation, as well as its complexity, a programme unit may choose:

- (a) an individual evaluator who may be national or international;
- (b) a team of evaluators, made up of national and/or international experts; or
- (c) a firm which provides an individual evaluator or team.

When procuring an evaluation through a firm, the implementing agency should still review the qualifications of the proposed team members to ensure that they meet the needs of the evaluation.

The selection process should start at least **one month prior to proposed commencement** of the evaluation, though it is recommended to start earlier as good evaluators are in high demand and not always available at the time requested.

It is advisable to have a team comprised of at least two evaluators, preferably national and international. This will allow for the evaluators to compare notes, verify the accuracy of information collected and recorded, divide efforts to interview more people and bounce ideas off of each other. In addition, **evaluation teams should be balanced, to the extent possible, in terms of gender and geographical composition.**

¹⁰⁸ As detailed in the Programme and Operations Policies and Procedures (POPP), <https://popp.undp.org/SitePages/POPPBSUnit.aspx?TermID=254a9f96-b883-476a-8ef8-e81f93a2b38d&Menu=BusinessUnit>

Box 7: Evaluator Databases¹⁰⁹

There are several evaluator databases available across UNDP which can provide details of evaluators with sectoral, regional and country-level experience. The search for evaluators needs to consider knowledge and expertise in **evaluation** (including data analysis and report writing), relevant **thematic** areas and **cross-cutting** areas such as gender equality, disability issues, rights-based approach, capacity development and national/regional expertise.

➤ *Global Policy Network (GPN) /ExpRes Roster¹¹⁰*

The GPN/ExpRes Roster is a recruitment and deployment mechanism which maintains technically-vetted consultants across 21 profiles and 79 sub-profiles to support the work of UNDP country offices, programme units and partner agencies. A similar roster was established in 2019 for midterm evaluations and reviews for vertical fund-financed projects which currently includes over 40 vetted candidates.

ExpRes Roster recruitment process

- Submit a request via the [GPN/ExpRes Deployment Request Form](#). In the comments section you may mention the names of any specific candidates that you would like to consider, if any.
- The deployments team will share curriculum vitae (CVs) for you to rank informally for crisis-related assignments, where no formal desk review is needed. **For non-crisis assignments, a quick scoring index will be used.**
- Upon ranking, the deployments team will check availability and link you with the highest ranked available candidate to negotiate with and contract based on their fee on the roster.
- The 'bottom' fee range of the candidates provided will remain within 30% of the 'top' fee range. For example, if the fee of the highest priced consultant is \$500 per day, the lowest priced consultant option cannot be less than \$350 per day. Country offices may select an expert whose fee is no more than 30% higher than the next most suitable candidate, as long as they are the best technical option.
- No financial proposals are required, as fee information is provided on the roster. The deployments team will provide guidance on fee negotiations for longer assignments (over one month). The fee cannot exceed the stated rate on the roster.
- Signed individual contracts are to be shared with the deployments team. Any contract amendments also need to be shared, ensuring that the total cumulative amount stays below \$100,000. If the case needs to go to the local Contracts, Assets, and Procurement Committee or Regional Advisory Committee on procurement, these approval documents will also need to be shared.
- Feedback/ written assessment on the work of each expert has to be provided via the GPN/ExpRes Evaluation or Individual Contract Evaluation Form.

¹⁰⁹ The ERC contains a referral-based Database of Evaluators. This is not a fully vetted list of consultants. The consultants included in this roster are either referred by UNDP staff members working in the area of M&E or they were a team member working on one of the UNDP evaluations. UNDP IEO does not maintain any long-term agreement with any of the consultants in the database. Therefore, the hiring unit must do requisite due diligence in following the UNDP procurement process.

¹¹⁰ This Roster is managed by the Global Policy Network/ ExpRes.

➤ **Regional evaluator rosters**

Several regional service centers maintain databases and rosters of evaluators, which can be accessed and used by programme units. These are updated regularly. In addition, regional evaluation focal points can share advice and experience on finding suitable evaluators.

➤ **The Evaluation Resource Center**

The IEO offers a database of evaluation experts via the ERC (login required). Programme units can also add to this database. The database can be searched by name, thematic area, region and country, and can also link evaluators to their evaluation work for UNDP. The database is not vetted and programme units will need to request references.

Evaluators not in the main database can be searched using a general search of all reports. If they have undertaken evaluations for UNDP and their work is in the ERC, the reports can be found alongside the quality assessment review scores for their evaluations. A search of the ERC by regional, country or thematic area may also bring to light experts and evaluators who could be used.

A guidance note called “Finding good evaluators using the ERC and Quality Assessment scores” has also been made available.¹¹¹ A new roster is currently being developed.

In order to ensure the credibility and usability of an evaluation, programme units need to ensure the independence and impartiality of evaluators, and that they are free from conflicts of interest (see box 8).



Evaluators should not have worked or contributed to the project, programme, outcome or UNDAF/ UNSDCF under evaluation, at any time, in any way. Equally, there should be no possibility of future contracts for the evaluator in the area under evaluation. In either of these cases, the evaluator would not be able to provide objective and impartial analysis of the evaluation subject.

Box 17: Sources of conflict of interest in evaluation

Conflict of interest due to past engagement

UNDP commissioning units may not assign consultants to the evaluation of former UNDAFs, country programmes, outcomes, sectors or thematic areas in which they have had prior involvement, whether in design, implementation, decision-making or financing. Following this principle, UNDP staff members, including advisers based in regional centres and headquarters units, civil servants or employees of NGOs that may be or have been directly or indirectly related to the programme or project, should not take part in the evaluation team.

More broadly, UNDP programme units should consider whether conducting multiple assignments could create a conflict of interest. Many consultants and evaluators undertake numerous assignments for UNDP and its partners over their professional careers. This can include a mixture of evaluation and

¹¹¹ <https://erc.undp.org/resources/docs/guidance/ERC-Guide-finding-good-evaluator.pdf>

advisory roles with multiple agencies at different levels. Programme units should make a judgement as to whether a consultant with a high reliance on work with UNDP may preclude them from producing an impartial evaluation. The ERC provides a recent history of evaluations undertaken by an evaluator.

Conflict of interest due to potential future involvement

Programme units must ensure that the evaluators will not be rendering any service (related or unrelated to the subject of the evaluation) to the programme unit of the project or outcome being evaluated in the immediate future. Evaluators should not subsequently be engaged in the implementation of a programme or project that was the subject of their evaluation. Equally, evaluators should not be engaged to design subsequent phases of projects that they have evaluated.

Evaluator’s obligation to reveal any potential conflicts of interest

Evaluators must inform UNDP and stakeholders of any potential or actual conflict of interest. The evaluation report should address any potential or actual conflict of interest and indicate measures to mitigate any negative consequences. If a conflict of interest is uncovered or arises during the evaluation, the organization should determine whether the evaluator should be dismissed and/ or the evaluation terminated.

It is good practice to share the CVs of potential candidates with wider stakeholders and partners before engagement. This will help to ensure that there is no potential conflict of interest or objection to the selection. Check references by talking to colleagues and partners who have worked with the candidates previously to verify their competency as evaluators. The ERC and quality assessment ratings give a further check on the quality of an evaluator’s work.



Step Three: Managing an evaluation



4.4 Step Three: Managing an evaluation



Figure 3. The phases of managing an evaluation

4.4.1 Briefing the evaluation team

To safeguard independence, implementing agencies should ensure that there is no interference in the implementation of an evaluation, but this does not exclude any cooperation, support or direction to the evaluation team. A successful evaluation requires a good level of cooperation and support from the commissioning unit to the evaluation team.

Supporting roles of the programme unit, evaluation manager and project manager include:

- **Briefing the evaluation team** on the purpose and scope of the evaluation and explaining the expectations of UNDP and its stakeholders in terms of the required quality standards for the evaluation process and products. Reaching a joint understanding on the TOR and objectives of the evaluation.
- **Providing the evaluation team with relevant UNDP Evaluation Policy guidelines**, including the quality standards for evaluation reports, UNEG Norms and Standards for Evaluation and the quality assessment guidance. In particular, evaluators must understand the requirement to follow ethical principles set out in the UNEG ethical guidelines for evaluators, and sign the pledge of ethical conduct for evaluators in the United Nations system.¹¹²
- **Ensuring that all relevant information is available to the evaluators.** If they encounter any difficulty in obtaining information that is critical for the conduct of evaluation, provide necessary support.
- **Providing preliminary partner, stakeholder and beneficiary information** for the evaluation team. While the evaluation team is responsible for identifying who they wish to meet and UNDP cannot interfere with their decisions, further suggestions can be made and access to partners, stakeholders and beneficiaries facilitated.
- **Organizing a kick-off meeting to introduce the evaluation team to the evaluation reference group and other partners and stakeholders** and facilitate initial contact.
- **Supporting the arrangement of interviews, meetings and field visits.** Programme units should support contact and send interview requests as needed to ensure that meetings are held.
- **Providing comments on and assuring the quality of the workplan and inception report**, including the elaborated evaluation methodology prepared by the evaluation team.

¹¹² UNEG, Pledge of Ethical Conduct in Evaluation, Access at: <http://www.unevaluation.org/document/detail/2866>.

- **Ensuring the security of consultants, stakeholders and accompanying UNDP staff**, particularly in crisis situations. The evaluation team members should have passed relevant United Nations security exams and be aware of and compliant with related security protocols, including passing the United Nations Department of Safety and Security training courses on basic security in the field¹¹³ and advanced security in the field.¹¹⁴

4.4.2 Supplying supporting information

Following the introductory meetings and briefings outlined above, the evaluation team will undertake a **thorough** desk review of all relevant reports and data. These should be supplied by the programme unit in a timely manner, and all efforts should be made to access missing reports and data prior to the development of the inception report and the data collection mission.

4.4.3 Evaluation inception report



The inception report is a written document prepared by the evaluator after an initial review of relevant documentation. It sets out the conceptual framework to be applied in the evaluation. It includes the understanding of the evaluation objectives, theory of change, evaluation questions and possible sub-questions, defines the methodology, and provides information on data sources and collection, sampling and key indicators.

Evaluators will commence the evaluation process with a desk review and preliminary analysis of the information supplied by the implementing agency. Based on the TOR, initial meetings with the UNDP programme unit/ evaluation manager (reference group) and the desk review, evaluators should develop an inception report. The evaluation manager needs to share the UNDP template for the inception report at the beginning of the assignment with the evaluation team leader. This includes a description of what is being evaluated, and illustrates the evaluators' understanding of the logic or theory of how the initiative is supposed to work, including strategies, activities, outputs and expected outcomes and their interrelationships (see box 9).



The inception report provides an opportunity to clarify issues and understanding of the objective and scope of the evaluation, such as resource requirements and delivery schedules. Any issues or misunderstandings identified should be addressed at this stage, prior to any data collection or field missions.

The commissioning unit and key stakeholders should review and assure the quality of the inception report and its adherence to the TOR and goals of the evaluation, as well as discussions held with the evaluation team. The inception report needs to be submitted according to the time frame in the TOR, usually two to four weeks after contract signing.

¹¹³ Access at: <https://training.dss.un.org/course/category/1>

¹¹⁴ Access at: <https://training.dss.un.org/course/category/2>



The inception report is a key milestone in the evaluation process, and it is important that the evaluation manager and reference group pay adequate attention to it. A separate meeting or call should be scheduled to discuss the methodological approach of the evaluation. The inception report needs to be officially approved by the evaluation manager / reference group before field missions can commence.

Box 18: Inception report content

1. **Background and context**, illustrating the understanding of the project/ outcome to be evaluated.
2. **Evaluation objective, purpose and scope**. A clear statement of the objectives of the evaluation and the main aspects or elements of the initiative to be examined.
3. **Evaluation criteria and questions**. The criteria the evaluation will use to assess performance and rationale. The stakeholders to be met and interview questions should be included and agreed, as well as a proposed schedule for field visits.
4. **Evaluability analysis**. Illustrates the evaluability analysis based on formal (clear outputs, indicators, baselines, data) and substantive (identification of problem addressed, theory of change, results framework) approaches, and the implications for the proposed methodology.
5. **Cross-cutting issues**. Provide details of how cross-cutting issues will be evaluated, considered and analysed throughout the evaluation. The description should specify how methods for data collection and analysis will integrate **gender considerations**, ensure that data collected is disaggregated by sex and other relevant categories, and employ a diverse range of data sources and processes to ensure the inclusion of diverse stakeholders, including the most vulnerable where appropriate.
6. **Evaluation approach and methodology**, highlighting the conceptual models to be adopted, and describing the data collection methods,¹¹⁵ sources and analytical approaches to be employed, including the rationale for their selection (how they will inform the evaluation) and their limitations; data-collection tools, instruments, and protocols; and discussing their reliability and validity for the evaluation and the sampling plan.
7. **Evaluation matrix**, identifying the key evaluation questions and how they will be answered through the selected methods.
8. A revised **schedule of key milestones**, deliverables and responsibilities, including the evaluation phases (data collection, data analysis and reporting).
9. Detailed **resource requirements**, tied to evaluation activities and deliverables detailed in the workplan. Include specific assistance required from UNDP, such as providing arrangements for visiting particular field offices or sites
10. **Outline of the draft/ final report** as detailed in the guidelines, and ensuring quality and usability (outlined below). The agreed report outline should meet the quality standards outlined in these guidelines and the quality assessment requirements outlined in section 6.

4.4.4 Evaluation and data collection mission

Once the inception report has been agreed, with a detailed list of stakeholders and beneficiaries to be interviewed or met, UNDP should prepare a detailed workplan of meetings agreed with the evaluation team.

To ensure independence and confidentiality, **UNDP staff should not participate in any stakeholder or beneficiary meetings**. Interviews and meetings are confidential and evaluation reports should not assign any statements or findings to individuals.

¹¹⁵ Annex 2 outlines different data collection methods.

Following the field missions, and prior to the drafting of the evaluation report, the evaluation team should **debrief the UNDP project/ programme and management teams** with preliminary findings. Debriefings with key stakeholders and the evaluation reference group may also be organized. This gives an opportunity to discuss the preliminary findings and address any factual errors or misunderstandings, prior to writing the evaluation report.

4.4.5 Draft report and review process

The evaluation manager is ultimately responsible for the quality of the evaluation and should plan to review the relevance and accuracy of any reports, and their compliance with the TOR. The evaluation manager needs to share the UNDP evaluation report template with the evaluation team leader and inform them about the UNDP evaluation quality assurance system. Where the template has been shared earlier in the evaluation process, it is recommended to remind the evaluation team leader to adhere to the template.

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, lessons and recommendations, and be presented in a way that makes the information accessible and comprehensible. **A quality evaluation report should:**

- Have a concise executive summary (maximum four pages).
- Be well structured and complete.
- Describe sufficiently what exactly is being evaluated and why.
- Include an analysis of the project design, theory of change or results framework.
- Identify the evaluation questions of concern to users.
- Explain the steps and procedures used to answer those questions.
- Acknowledge limitations and constraints in undertaking the evaluation.
- Identify target groups covered by the evaluation, whether or not the needs of the target groups were addressed through the intervention, and why.
- Address gender equality and women's empowerment, disability and other cross-cutting issues.
- Present findings supported by credible evidence in response to the questions.
- Draw conclusions about the findings based on the evidence.
- Propose concrete and usable recommendations derived from the conclusions.
- Be written with the report users and how they will use the evaluation in mind.

Standard outline for an evaluation report

Annex 1 provides further information on the standard outline for the evaluation report. In brief, the minimum contents of an evaluation report include:

1. **Title and opening pages** with details of the project/ programme/ outcome being evaluated and the evaluation team.
2. **Project and evaluation details**, including the project title, Atlas number, budgets and project dates and other key information.
3. **Table of contents.**

4. **List of acronyms and abbreviations.**
5. **Executive summary**, a stand-alone section of maximum 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 the basis for readers to understand the design, general logic, results framework (theory of change) and other relevant information of the initiative being evaluated.
8. **Evaluation scope and objectives**, to provide a clear explanation of the evaluation scope, primary objectives and main questions.
9. **Evaluation approach and methods**, describing in detail the selected methodological approaches and methods.
10. **Data analysis**, describing the procedures used to analyse the data collected to answer the evaluation questions.
11. **Findings and conclusions**, setting out the evaluation findings, based on analysis of the data collected, and the conclusions drawn from these findings.
12. **Recommendations**. The report should provide a reasonable number of practical, feasible recommendations directed to the intended users of the report about what actions to take or decisions to make.
13. **Lessons learned**. As appropriate and when requested in the TOR, the report should include discussion of lessons learned from the evaluation of the intervention.
14. All **findings, conclusions, recommendations and lessons learned** need to consider gender equality and women’s empowerment, disability, and other cross-cutting issues.
15. **Annexes**. At a minimum these should include:
 - a. TOR for the evaluation.
 - b. Evaluation matrix and data collection instruments
 - c. List of individuals or groups interviewed or consulted, and sites visited.
 - d. List of supporting documents reviewed.



When reviewing the evaluation report and its structure and content, evaluation managers should consider the requirements of the quality assessment ratings¹¹⁶ which IEO uses to assess the quality of evaluation reports. More details of the quality assessment process and criteria can be found in section 6.

GEF terminal evaluations and midterm reviews have their own reporting requirements, which can be found in their evaluation guidelines.¹¹⁷

4.4.6 Evaluation review processes

¹¹⁶ Access at: <http://web.undp.org/evaluation/guideline/section-6.shtml>

¹¹⁷ Access at: <http://web.undp.org/evaluation/guidance.shtml#gef>

Providing comments on the draft evaluation report

The evaluation manager should coordinate the collection of all comments, questions and requests for clarification into one document, which should be shared with the evaluation team within the agreed time schedule. This should include inputs, contributions and comments from UNDP, the evaluation reference group, and external stakeholders.



The evaluation manager needs to ensure that the draft evaluation report reflects the TOR, relates to the inception report including issues agreed during the inception phase, and is checked for factual accuracy.

Comments, questions, suggestions and requests for clarification on the evaluation draft should be provided in an evaluation “audit trail” document and not directly in the draft report (see Table 4). Where errors of fact or misunderstanding of approaches are found, documentation should be provided to support comments and requests.

The evaluator or evaluation team should reply to the comments through the evaluation audit trail document. If there is disagreement in findings, these should be documented through the evaluation audit trail and effort made to come to an agreement. If UNDP continues to disagree with the findings, conclusions or recommendations in an evaluation report, this should be clearly stated in the management response, with supporting reasons for the disagreement.

The evaluation audit trail is not part of the evaluation report and is not a public document, but is part of the process for completion of the evaluation report. The evaluation audit trail should not be included in the final report or uploaded to the ERC. In some circumstances where the commissioning unit and/ or stakeholders disagree with a finding, a note can be made in the report as to the disagreement.

Table 6. Audit trail form template

Chapter and section number	Paragraph number/ line number	Comments	Evaluation team responses and/ or actions taken

Programme units should not make any adjustments to an evaluation report but should address any disagreement of findings, conclusions or recommendations through the management response.

4.4.7 Dispute settlement and reporting wrongdoing

Dispute settlement

Disputes between evaluators and those being evaluated are not uncommon. The audit trail provides an avenue to highlight issues with an evaluation, and also for the evaluator to provide further evidence to support their findings.

If there is a continued disagreement, then either UNDP 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 UNDP Programme and Operations Policies and Procedures (POPPs).¹¹⁸

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 to receive allegations.¹¹⁹

Anyone with information regarding fraud 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.

When reporting to the Investigations Hotline, people are encouraged to be as specific as possible, including the basic details of who, what, where, when, and how any of these incidents occurred. Specific information will allow OAI to properly investigate the alleged wrongdoing.

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 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

¹¹⁸ Access at:

https://poppp.undp.org/layouts/15/WopiFrame.aspx?sourcedoc=/UNDP_POPP_DOCUMENT_LIBRARY/Public/PSU_Individual%20Contract_Individual%20Contract%20Policy.docx&action=default

¹¹⁹ <https://www.undp.org/accountability/audit/investigations>

To ensure evaluators are fully aware of the dispute resolution and wrongdoing reporting processes in UNDP, details should be given to them at the time of the signing their contracts. A standard form to be included in all evaluator contracts is detailed in Annex 4.



Step Four: Using the evaluation - management response, knowledge sharing and dissemination



4.5 Step Four: Using the evaluation

4.5.1 Preparing the management response for decentralized evaluations



A management response is a formal mechanism to ensure that evaluation findings, conclusions, lessons learned and recommendations will be used. A management response should contribute to the effectiveness, learning and accountability of the intervention being evaluated, and the organizations and institutions involved.

Why do we prepare a management response?

All UNDP evaluations have to develop management responses for each recommendation. This ensures the effective use of evaluation findings and recommendations, through considered follow-up actions.

Management responses should include detailed actions and highlight which agency or unit is responsible for each key action by when.

Management responses and key follow-up actions are closely monitored by headquarters, IEO and regional bureaux to ensure that lessons are being learned from evaluations. The preparation of a management response should not be seen as a one-time activity. Learning emanating from the management response process should be documented and reflected upon when designing a new project or programme or defining an outcome. The process of developing a management response to terminal project evaluations (specifically for projects that have been completed) allows key stakeholders to reflect on the project results and generate lessons that are applicable beyond a particular project to support other activities, projects and outcomes of the programme units.

What is in a management response?

Firstly, the management response outlines whether the programme unit accepts each recommendation and how it will deal with it. **Programme units can fully accept, partially accept or reject a recommendation**, and must provide justification for this.

- **Fully accepted:** agrees entirely with the whole recommendation and will seek actions to achieve it.
- **Partially accepted:** agrees with elements of the recommendation. The management response should detail the elements of agreement and those of disagreement, and give reasons why parts of the recommendations are not considered valid.
- **Rejected:** management must state why they reject the recommendation and will not follow up on it (no key actions need to be included in the response).

Key actions



When recommendations are fully or partially accepted, they require a corresponding management response and key follow-up actions.

Recommendations can have several key actions to ensure the evaluation recommendation is met. It is important that key actions:

- Clearly state the nature of the action and how it will address the recommendations.
- Indicate the parties (unit or organization) responsible for implementing the key action and accountable for its implementation.
- Are time-bound with clear deadlines and schedules for completion. Ideally, key actions should be completed within 18 months of an evaluation.
- Are of a reasonable number to facilitate implementation, tracking and oversight.

Management responses and key follow-up actions to evaluation recommendations need to be discussed and agreed within the project boards (where relevant). If UNDP or its implementing partner disagrees with a recommendation, they can so state in the management response and no key follow-up actions need to be added.



Management responses and key actions need to be agreed and entered into the ERC within six weeks of completion of the evaluation report. The management response template can be found in the annexes.

Box 19: Terms of reference and recommendations

It is important that the evaluation TOR clearly request recommendations that are targeted and anticipate actual follow-up and implementation. At the same time, the TOR should call for a realistic set of recommendations that are implementable and manageable in number (7-10) and, when draft reports are reviewed, evaluators should consider grouping recommendations under broad area headings to ensure this.

If there are too many recommendations they can be difficult to implement and manage and will considerably stretch resources in: (a) developing management responses and key actions; (b) entering recommendations, management responses and key actions to the ERC; and (c) monitoring and reporting on implementation of the management responses and key actions.

Joint project management responses

For joint projects and UNSDCF evaluations, UNDP should cooperate and coordinate with project partners in the development of management responses and key actions. UNDP programme units are only responsible for those recommendations targeted at them, and should develop management responses and key actions only for these.

Monitoring implementation of key actions

M&E specialists or focal points are responsible for monitoring the implementation of key actions and reporting on achievements through the ERC. This should be undertaken **on a quarterly basis**, and the ERC should be updated accordingly, with supporting documentation where applicable. Regional bureaux also oversee and monitor implementation and follow up with programme units on implementation. The ERC can be used to monitor implementation of management responses and key action commitments.



The IEO reports on the number of management responses and key actions completed, initiated, overdue or considered no longer applicable in its annual report.



The Evaluation Policy requires management responses to all evaluations regardless of the status of the initiative that was evaluated. The management response template can be found in the annexes.

4.5.2 Publication of the final evaluation report

All evaluation plans and evaluation reports must be uploaded to the ERC.¹²⁰ The ERC is an open information site accessible to all to view evaluation plans and evaluations. Some information on the ERC is for internal UNDP use only and requires IEO to provide access. Each programme unit can nominate several ERC focal points who will have different permissions across the site for uploading and accessing data for oversight purposes.

Access is required to upload evaluation plans, TORs and evaluations. Normally the M&E focal point or officer has this level of access to the ERC using their UNDP email and password.

Only the following documents need to be uploaded for completed evaluations:

- Final TOR for the evaluation, **uploaded within two weeks** of completion of the TOR.
- Final evaluation report **uploaded within two weeks** of agreement and completion of the report.
- Supporting annexes, **uploaded with the evaluation report** if not as part of the main report.

No other documents need to be uploaded. **Evaluation audit trails and inception reports should not be uploaded to the ERC.** Programme units should remember that the ERC site is public, and therefore only final documents should be uploaded. Documents should not contain comments or track changes and should ideally be uploaded as PDF files, although Word documents are acceptable. Therefore:

- Evaluation audit trails should not be uploaded to the ERC.
- Inception reports should not be uploaded separately to the ERC

Once the evaluation is complete and uploaded, information previously entered will need to be updated including the completion date, evaluation expenditure and the individual names and roles of the evaluation team members. More information is available in the ERC guidance.¹²¹

Recommendations, management responses and key actions

The following information also needs to be uploaded to the ERC by the evaluation manager within the schedule outlined below:

- Evaluation recommendations addressed to UNDP entered into the ERC when the final report is uploaded (**within two weeks of completion**).
- Evaluation management responses entered into the ERC **within six weeks** of receipt of the final evaluation report.
- Evaluation key actions entered into the ERC along with the management responses.

¹²⁰ Access at: <https://erc.undp.org/>

¹²¹ Access at: <https://erc.undp.org/resources/docs/guidance/ERC-User-Guide.pdf>

Recommendations are entered into the ERC individually and should not be cut and pasted as a single block of recommendations. Individual recommendations, management responses and key actions are reported and tracked separately so must be entered to the ERC individually. Where the TOR requires an evaluation team to produce lessons learned, these should also be uploaded. A separate page is available for this information.



Annexes



Annex 1. Evaluation TOR template

The TOR should, at a minimum, cover the elements described below.

1. Background and context

The background section makes clear what is being evaluated and identifies critical social, economic, political, geographic and demographic factors that have a direct bearing on the evaluation. This description should be focused and concise (a maximum of two pages) highlighting **only** those issues most pertinent to the evaluation. The key background and context descriptors that should be included are:

- Description of the intervention (outcome, programme, project, group of projects, themes, soft assistance) that is being evaluated.
- The name of the intervention (e.g., project name), purpose and objectives, including when and how it was initiated, who it is intended to benefit and what outcomes or outputs it is intended to achieve, the duration of the intervention and its implementation status within that time frame.
- The scale and complexity of the intervention, including, for example, the number of components, if more than one, and the size and description of the population each component is intended to serve, both directly and indirectly.
- The geographic context and boundaries, such as the region, country or landscape, and challenges where relevant.
- Total resources required for the intervention from all sources, including human resources and budgets comprising UNDP, donor and other contributions and total expenditures.
- Key partners involved in the intervention, including the implementing agencies and partners, other key stakeholders, and their interest, concerns and relevance for the evaluation.
- Observed changes since the beginning of implementation and contributing factors.
- State details of project beneficiaries (gender, disability, vulnerable groups, human rights issues, etc.)
- How the subject fits into: the partner government's strategies and priorities; international, regional or country development goals; strategies and frameworks; the SDGs, UNDP corporate goals and priorities; and UNDP global, regional or country programmes, as appropriate.
- Key features of the international, regional, and national economies and economic policies that have relevance for the evaluation.
- Description of how this evaluation fits within the context of other ongoing and previous evaluations and the evaluation cycle.

More detailed background and context information (e.g., initial funding proposal, strategic plans, **logical framework or theory of change**, monitoring plans and indicators) should be included or referenced in annexes.

Basic project information can also be included in table format as follows:

PROJECT/OUTCOME INFORMATION		
Project/outcome title		
Atlas ID		
Corporate outcome and output		
Country		
Region		
Date project document signed		
Project dates	Start	Planned end
Project budget		
Project expenditure at the time of evaluation		
Funding source		
Implementing party ¹²²		

2. Evaluation purpose, scope and objectives

This section of the TOR explains clearly *why* the evaluation is being conducted, *who* will use or act on the evaluation results and *how* they will use or act on the results. The purpose should include some background and justification for why the evaluation is needed at this time and how the evaluation fits within the programme unit’s evaluation plan. A clear statement of purpose provides the foundation for a well-designed evaluation.

Scope and objectives of the evaluation should detail and include:

- aspects of the intervention to be covered by the evaluation. This can include the time frame, implementation phase, geographic area and target groups to be considered and, as applicable, which projects (outputs) are to be included.
- the primary issues of concern to users that the evaluation needs to address or objectives the evaluation must achieve.

Issues relate directly to the questions the evaluation must answer so that users will have the information they need for pending decisions or action. An issue may concern the relevance, coherence, efficiency, effectiveness or sustainability of the intervention. In addition, UNDP evaluations must address how the intervention sought to **mainstream gender in development efforts**, considered disability issues and applied the rights-based approach.

¹²² This is the entity that has overall responsibility for implementation of the project (award), effective use of resources and delivery of outputs in the signed project document and workplan.

3. Evaluation criteria and key guiding questions

Evaluation questions define the information that the evaluation will generate. This section proposes the questions that, when answered, will give intended users of the evaluation the information they seek in order to make decisions, take actions or increase knowledge. Questions should be grouped according to the four or five OECD-DAC evaluation criteria: (a) relevance; (b) coherence; (c) effectiveness; (d) efficiency; and (e) sustainability (and any other criteria used).



Individual evaluation questions should be developed by the evaluation manager to address the key concerns of the evaluation and should not just copied from the list below, which is illustrative.

The TOR should contain a reasonable and not exhaustive range of questions which can be realistically covered under a limited time evaluative exercise.

Sample questions for different types of evaluation:

Guiding evaluation questions need to be outlined in the TOR and further refined by the evaluation team and agreed with UNDP evaluation stakeholders.

Outcome evaluation sample questions

Relevance/Coherence

- To what extent is the initiative in line with the UNDP mandate, national priorities and the requirements of targeting women, men and vulnerable groups?
- To what extent is UNDP support relevant to the achievement of the SDGs in the country?
- To what extent did UNDP adopt gender-sensitive, human rights-based and conflict-sensitive approaches?
- To what extent is UNDP engagement a reflection of strategic considerations, including the role of UNDP in a particular development context and its comparative advantage?
- To what extent was the method of delivery selected by UNDP appropriate to the development context?
- To what extent was the theory of change presented in the outcome model a relevant and appropriate vision on which to base the initiatives?

Effectiveness

- To what extent has progress been made towards outcome achievement? What has been the UNDP contribution to the observed change?
- What have been the key results and changes attained for men, women and vulnerable groups?
- How has delivery of country programme outputs led to outcome-level progress?
- Have there been any unexpected outcome-level results achieved beyond the planned outcome?
- To what extent has UNDP improved the capacities of national implementing partners to advocate on environmental issues, including climate change issues and disaster risk reduction?
- To what extent has UNDP partnered with civil society and local communities to promote environmental and disaster risk awareness in the country?
- To what extent have the results at the outcome and output levels generated results for gender equality and the empowerment of women?
- To what extent have marginalized groups benefited?
- To what extent have triangular and South-South cooperation and knowledge management contributed to the results attained?

- Which programme areas are the most relevant and strategic for UNDP to scale up or consider going forward?

Efficiency

- To what extent have the programme or project outputs resulted from economic use of resources?
- To what extent were resources used to address inequalities and gender issues?
- To what extent were quality country programme outputs delivered on time?
- To what extent were partnership modalities conducive to the delivery of country programme outputs?
- To what extent did monitoring systems provide management with a stream of data, disaggregated by sex, that allowed it to learn and adjust implementation accordingly?
- To what extent did UNDP promote gender equality, the empowerment of women, human rights and human development in the delivery of country programme outputs?
- To what extent have UNDP practices, policies, processes and decision-making capabilities affected the achievement of the country programme outcomes?
- To what extent did UNDP engage or coordinate with different beneficiaries (men and women), implementing partners, other United Nations agencies and national counterparts to achieve outcome-level results?

Sustainability

- To what extent did UNDP establish mechanisms to ensure the sustainability for female and male beneficiaries of the country programme outcomes?
- To what extent do national partners have the institutional capacities, including sustainability strategies, in place to sustain the outcome-level results?
- To what extent are policy and regulatory frameworks in place that will support the continuation of benefits for men and women in the future?
- To what extent have partners committed to providing continuing support (financial, female and male staff, etc.)?
- To what extent do mechanisms, procedures and policies exist to carry forward the results attained on gender equality, empowerment of women, human rights, and human development by primary stakeholders?
- To what extent do partnerships exist with other national institutions, NGOs, United Nations agencies, the private sector and development partners to sustain the attained results?

Project evaluation sample questions:

Relevance/ Coherence

- To what extent was the project in line with national development priorities, country programme outputs and outcomes, the UNDP Strategic Plan, and the SDGs?
- To what extent does the project contribute to the theory of change for the relevant country programme outcome?
- To what extent were lessons learned from other relevant projects considered in the design?
- To what extent were perspectives of men and women who could affect the outcomes, and those who could contribute information or other resources to the attainment of stated results, taken into account during project design processes?
- To what extent does the project contribute to gender equality, the empowerment of women and the human rights-based approach?

- To what extent has the project been appropriately responsive to political, legal, economic, institutional, etc., changes in the country?

Effectiveness

- To what extent did the project contribute to the country programme outcomes and outputs, the SDGs, the UNDP Strategic Plan, and national development priorities?
- To what extent were the project outputs achieved, considering men, women, and vulnerable groups?
- What factors have contributed to achieving, or not, intended country programme outputs and outcomes?
- To what extent has the UNDP partnership strategy been appropriate and effective?
- What factors contributed to effectiveness or ineffectiveness?
- In which areas does the project have the greatest achievements? Why and what have been the supporting factors? How can the project build on or expand these achievements?
- In which areas does the project have the fewest achievements? What have been the constraining factors and why? How can or could they be overcome?
- What, if any, alternative strategies would have been more effective in achieving the project objectives?
- Are the project objectives and outputs clear, practical and feasible within its frame? Do they clearly address women, men and vulnerable groups?
- To what extent have different stakeholders been involved in project implementation?
- To what extent are project management and implementation participatory, and is this participation of men, women and vulnerable groups contributing towards achievement of the project objectives?
- To what extent has the project been appropriately responsive to the needs of the national constituents (men, women, other groups) and changing partner priorities?
- To what extent has the project contributed to gender equality, the empowerment of women and the realization of human rights?

Efficiency

- To what extent was the project management structure as outlined in the project document efficient in generating the expected results?
- To what extent were resources used to address inequalities in general, and gender issues in particular?
- To what extent have the UNDP project implementation strategy and execution been efficient and cost-effective?
- To what extent has there been an economical use of financial and human resources? Have resources (funds, male and female staff, time, expertise, etc.) been allocated strategically to achieve outcomes?
- To what extent have resources been used efficiently? Have activities supporting the strategy been cost-effective?
- To what extent have project funds and activities been delivered in a timely manner?
- To what extent do the M&E systems utilized by UNDP ensure effective and efficient project management?

Sustainability

- Are there any financial risks that may jeopardize the sustainability of project outputs affecting women, men and vulnerable groups?

- To what extent will targeted men, women and vulnerable people benefit from the project interventions in the long-term?
- To what extent will financial and economic resources be available to sustain the benefits achieved by the project?
- Are there any social or political risks that may jeopardize sustainability of project outputs and the project contributions to country programme outputs and outcomes?
- Do the legal frameworks, policies and governance structures and processes within which the project operates pose risks that may jeopardize sustainability of project benefits?
- To what extent did UNDP actions pose an environmental threat to the sustainability of project outputs, possibly affecting project beneficiaries (men and women) in a negative way? What is the chance that the level of stakeholder ownership will be sufficient to allow for the project benefits to be sustained?
- To what extent do mechanisms, procedures and policies exist to allow primary stakeholders to carry forward the results attained on gender equality, empowerment of women, human rights and human development?
- To what extent do stakeholders (men, women, vulnerable groups) support the project's long-term objectives?
- To what extent are lessons learned documented by the project team on a continual basis and shared with appropriate parties who could learn from the project?
- To what extent do UNDP interventions have well-designed and well-planned exit strategies which include a gender dimension?
- What could be done to strengthen exit strategies and sustainability in order to support female and male project beneficiaries as well as marginalized groups?

Sample evaluation questions on cross-cutting issues

Human rights

- To what extent have poor, indigenous and physically challenged, women, men and other disadvantaged and marginalized groups benefited from the work of UNDP in the country?

Gender equality

All evaluation criteria and evaluation questions applied need to be checked to see if there are any further gender dimensions attached to them, in addition to the stated gender equality questions.

- To what extent have gender equality and the empowerment of women been addressed in the design, implementation and monitoring of the project?
- Is the gender marker assigned to this project representative of reality?
- To what extent has the project promoted positive changes in gender equality and the empowerment of women? Did any unintended effects emerge for women, men or vulnerable groups?

Disability

- Were persons with disabilities consulted and meaningfully involved in programme planning and implementation?
- What proportion of the beneficiaries of a programme were persons with disabilities?
- What barriers did persons with disabilities face?
- Was a twin-track approach adopted? ¹²³

¹²³ The twin-track approach combines mainstream programmes and projects that are inclusive of persons with disabilities as well as programmes and projects that are *targeted* towards persons with disabilities. It is an essential element of any strategy

4. Methodology

The TOR may **suggest** an overall approach and method for conducting the evaluation, as well as data sources and tools that will likely yield the most reliable and valid answers to the evaluation questions within the limits of resources. However, final decisions about the specific design and methods for the evaluation should emerge from consultations with the programme unit, the evaluators and key stakeholders about what is appropriate and feasible to meet the evaluation purpose and objectives and answer the evaluation questions, given limitations of budget, time and data.

Evaluation should employ a combination of qualitative and quantitative evaluation methods and instruments. The evaluator is expected to follow a participatory and consultative approach that ensures close engagement with the evaluation managers, implementing partners and male and female direct beneficiaries. Suggested methodological tools and approaches may include:

- **Document review.** This would include a review of all relevant documentation, inter alia
 - Project document (contribution agreement).
 - Theory of change and results framework.
 - Programme and project quality assurance reports.
 - Annual workplans.
 - Activity designs.
 - Consolidated quarterly and annual reports.
 - Results-oriented monitoring report.
 - Highlights of project board meetings.
 - Technical/financial monitoring reports.
- **Interviews and meetings** with key stakeholders (men and women) such as key government counterparts, donor community members, representatives of key civil society organizations, United Nations country team (UNCT) members and implementing partners:
 - **Semi-structured interviews**, based on questions designed for different stakeholders based on evaluation questions around relevance, coherence, effectiveness, efficiency, and sustainability.
 - Key informant and **focus group discussions** with men and women, beneficiaries and stakeholders.
 - All interviews with men and women should be undertaken in full confidence and anonymity. The final evaluation report should not assign specific comments to individuals.
- **Surveys and questionnaires** including male and female participants in development programmes, UNCT members and/or surveys and questionnaires to other stakeholders at strategic and programmatic levels.
- **Field visits** and on-site validation of key tangible outputs and interventions.
- **Other methods** such as outcome mapping, observational visits, group discussions, etc.
- **Data review and analysis** of monitoring and other data sources and methods. To ensure maximum validity, reliability of data (quality) and promote use, the evaluation team will ensure triangulation of the various data sources.
- **Gender and human rights lens.** All evaluation products need to address gender, disability, and human right issues.

that seeks to mainstream disability inclusion successfully. Also, see chapter 9 of the Technical Notes. Entity Accountability Framework. United Nations Disability and Inclusion Strategy: <https://www.un.org/en/disabilitystrategy/resources>

The final methodological approach including interview schedule, field visits and data to be used in the evaluation should be clearly outlined in the inception report and fully discussed and agreed between UNDP, key stakeholders and the evaluators.

5. Evaluation products (deliverables)

The TOR should clearly outline the outputs UNDP expects from the evaluation team, with a detailed timeline and schedule for completion of the evaluation products. Where relevant, the TOR should also detail the length of specific products (number of pages). These products could include:

- **Evaluation inception report (10-15 pages).** The inception report should be carried out following and based on preliminary discussions with UNDP after the desk review and should be produced before the evaluation starts (before any formal evaluation interviews, survey distribution or field visits) and prior to the country visit in the case of international evaluators.
- **Evaluation debriefings.** Immediately following an evaluation, UNDP may ask for a preliminary debriefing and findings.
- **Draft evaluation report (within an agreed length).** A length of 40 to 60 pages including executive summary is suggested.
- **Evaluation report audit trail.** The programme unit and key stakeholders in the evaluation should review the draft evaluation report and provide an amalgamated set of comments to the evaluator within an agreed period of time, as outlined in these guidelines. Comments and changes by the evaluator in response to the draft report should be retained by the evaluator to show how they have addressed comments.
- **Final evaluation report.**
- **Presentations to stakeholders and/ or evaluation reference group** (if required).
- **Evaluation brief and other knowledge products** or participation in knowledge-sharing events, if relevant to maximize use.

6. Evaluation team composition and required competencies

This section details the specific skills, competencies and characteristics required of the evaluator / individual evaluators in the evaluation team, and the expected structure and composition of the evaluation team, including roles and responsibilities of team members. This may include:

- **Required qualifications:** education, length of experience in conducting/ managing evaluations, relevant knowledge, and specific country/regional experience.
- **Technical competencies:** team leadership skills and experience, technical knowledge in UNDP thematic areas, with specifics depending on the focus of the evaluation, data analysis and report writing etc.
- **Technical knowledge and experience:** Gender expertise/competencies in the evaluation team are a must. At least one evaluation team member or reference group member needs to have knowledge and/or experience of disability inclusion. Technical knowledge and experience in other cross-cutting areas such equality, disability issues, rights-based approach, and capacity development.
- **Language skills required.**

The section also should specify the type of evidence (resumes, work samples, references) that will be expected to support claims of knowledge, skills and experience.

The TOR should explicitly demand evaluators' independence from any organizations that have been involved in designing, executing, or advising any aspect of the intervention that is the subject of the evaluation.¹²⁴

7. Evaluation ethics

The TOR should include an explicit statement that evaluations in UNDP will be conducted in accordance with the principles outlined in the UNEG 'Ethical Guidelines for Evaluation'.¹²⁵

Standard text includes:

"This evaluation will be conducted in accordance with the principles outlined in the UNEG 'Ethical Guidelines for Evaluation'. The consultant must safeguard the rights and confidentiality of information providers, interviewees, and stakeholders through measures to ensure compliance with legal and other relevant codes governing collection of data and reporting on data. The consultant must also ensure security of collected information before and after the evaluation and protocols to ensure anonymity and confidentiality of sources of information where that is expected. The information knowledge and data gathered in the evaluation process must also be solely used for the evaluation and not for other uses with the express authorization of UNDP and partners."

8. Implementation arrangements

This section describes the organization and management structure for the evaluation and defines the roles, key responsibilities and lines of authority of all parties involved in the evaluation process. Implementation arrangements are intended to clarify expectations, eliminate ambiguities, and facilitate an efficient and effective evaluation process.

The section should describe the specific roles and responsibilities of the evaluators, including those of the members of the team, the evaluation manager, the commissioning programme unit and key stakeholders. The composition and expected roles and responsibilities of the advisory panel members or other quality assurance entities and their working arrangements should also be made explicit. The feedback mechanism regarding different evaluation products need to be outlined.

In case of a joint evaluation, the roles and responsibilities of participating agencies should be clarified. Issues to consider include: lines of authority; lines of and processes for approval; logistical considerations, such as how office space, supplies, equipment and materials will be provided; and processes and responsibility for approving deliverables.

9. Time frame for the evaluation process

This section lists and describes all tasks and deliverables for which evaluators, or the evaluation team, will be responsible and accountable, as well as those involving the commissioning office (e.g. workplan,

¹²⁴ For this reason, UNDP staff members based in other country offices, regional centres and headquarters units should not be part of the evaluation team.

¹²⁵ UNEG, 'Ethical Guidelines for Evaluation', 2020. Access at: <http://www.unevaluation.org/document/detail/2866>

agreements, briefings, draft report, final report). This must indicate for each the due date or time frame, as well as who is responsible for its completion. At a minimum, the time breakdown for the following activities should be included:

- Desk review.
- Briefings of evaluators.
- Finalizing the evaluation design and methods and preparing the detailed inception report.
- In-country data collection and analysis (visits to the field, interviews, questionnaires).
- Preparing the draft report.
- Stakeholder meeting and review of the draft report (for quality assurance).
- Incorporating comments and finalizing the evaluation report.

In addition, the evaluators may be expected to support UNDP efforts in knowledge sharing and dissemination.

Required formats for the inception reports, evaluation reports and other deliverables should be included in the annexes of the TOR for the evaluation being commissioned. This section should also state the number of working days to be used by each member of the evaluation team and the period during which they will be engaged (e.g., 30 working days over a period of three months).

Example of working day allocation and schedule for an evaluation (outcome evaluation)

ACTIVITY	ESTIMATED # OF DAYS	DATE OF COMPLETION	PLACE	RESPONSIBLE PARTY
Phase One: Desk review and inception report				
Meeting briefing with UNDP (programme managers and project staff as needed)	-	At the time of contract signing 1 June 2018	UNDP or remote	Evaluation manager and commissioner
Sharing of the relevant documentation with the evaluation team	-	At the time of contract signing 1 June 2018	Via email	Evaluation manager and commissioner
Desk review, Evaluation design, methodology and updated workplan including the list of stakeholders to be interviewed	7 days	Within two weeks of contract signing 1 to 15 June 2018	Home-based	Evaluation Team
Submission of the inception report (15 pages maximum)	-	Within two weeks of contract signing 15 June 2018		Evaluation team
Comments and approval of inception report	-	Within one week of submission of the inception report 22 June 2018	UNDP	Evaluation manager
Phase Two: Data-collection mission				
Consultations and field visits, in-depth interviews, and focus groups	15 days	Within four weeks of contract signing 1 to 21 July 2018	In country With field visits	UNDP to organize with local project partners, project staff, local authorities, NGOs, etc.
Debriefing to UNDP and key stakeholders	1 day	21 July 2018	In country	Evaluation team
Phase Three: Evaluation report writing				
Preparation of draft evaluation report (50 pages maximum excluding annexes), executive summary (4-5 pages)	7 days	Within three weeks of the completion of the field mission 21 July to 15 August	Home-based	Evaluation team
Draft report submission	-	15 August		Evaluation team
Consolidated UNDP and stakeholder comments to the draft report	-	Within two weeks of submission of the draft evaluation report 29 August 2018	UNDP	Evaluation manager and evaluation reference group
Debriefing with UNDP	1 day	Within one week of receipt of comments 4 September 2018	Remotely UNDP	UNDP, evaluation reference group, stakeholder, and evaluation team
Finalization of the evaluation report incorporating additions and comments provided by project staff and UNDP country office	4 days	Within one week of final debriefing 11 September 2018	Home-based	Evaluation team
Submission of the final evaluation report to UNDP country office (50 pages maximum excluding executive summary and annexes)	-	Within one week of final debriefing 11 September 2018	Home-based	Evaluation team
Estimated total days for the evaluation	35			

This is an illustrative example and individual evaluations will have their own requirements based on the nature and complexity of outcomes or projects, budget available, size of the evaluation team and deadline for completion, sharing or inclusion in other processes. Complex and larger programme and project evaluations often require more than 30 days.

The evaluation scope, number of days and budgets must be realistic and balanced, otherwise it could jeopardize the credibility and hence the utility of the evaluation.

10. Application submission process and criteria for selection

As required by the programme unit.

11. TOR annexes

Annexes can be used to provide additional detail about evaluation background and requirements to facilitate the work of evaluators. Some examples include:

- **Intervention results framework and theory of change.** Provides more detailed information on the intervention being evaluated.
- **Key stakeholders and partners.** A list of key stakeholders and other individuals who should be consulted, together with an indication of their affiliation and relevance for the evaluation and their contact information. This annex can also suggest sites to be visited.
- **Documents to be consulted.** A list of important documents and web pages that the evaluators should read at the outset of the evaluation and before finalizing the evaluation design and inception report. This should be limited to the critical information that the evaluation team needs. Data sources and documents may include:
 - Relevant national strategy documents.
 - Strategic and other planning documents (e.g., programme and project documents).
 - Monitoring plans and indicators.
 - Partnership arrangements (e.g., agreements of cooperation with governments or partners).
 - Previous evaluations and assessments.
 - UNDP evaluation policy, UNEG norms and standards and other policy documents.
- **Evaluation matrix** (suggested as a deliverable to be included in the inception report). The evaluation matrix is a tool that evaluators create as a map and reference in planning and conducting an evaluation. It also serves as a useful tool for summarizing and visually presenting the evaluation design and methodology for discussions with stakeholders. It details evaluation questions that the evaluation will answer, data sources, data collection and analysis tools or methods appropriate for each data source, and the standard or measure by which each question will be evaluated. Table 5 provides a sample evaluation matrix template.

Table 7. Sample evaluation matrix

Relevant evaluation criteria	Key questions	Specific sub-questions	Data sources	Data collection methods/tools	Indicators/success standards	Methods for data analysis

- **Schedule of tasks, milestones, and deliverables.** Based on the time frame specified in the TOR, the evaluators present the detailed schedule.
- **Required format for the evaluation report.** The final report must include, but not necessarily be limited to, the elements outlined in the template for evaluation reports (see annex 4 below).
- **Dispute and wrongdoing resolution process and contact details** (annex 3)
- **Pledge of ethical conduct in evaluation.** UNDP programme units should request each member of the evaluation team to read carefully, understand and sign the 'Pledge of Ethical Conduct in Evaluation of the United Nations system'.¹²⁶

¹²⁶

<http://www.unevaluation.org/document/detail/2866#:~:text=The%20UNEG%20Ethical%20Guidelines%20for%20Evaluation%20were%20first%20published%20in%202008.&text=This%20document%20aims%20to%20support,day%20to%20day%20evaluation%20practice.>

Annex 2. Summary of common data-collection methods/sources used in UNDP evaluations¹²⁷

METHOD/SOURCE	DESCRIPTION	ADVANTAGES	CHALLENGES
UNDP monitoring systems	Uses performance indicators to measure progress, particularly actual results against expected results	<ul style="list-style-type: none"> Can be a reliable, cost-efficient, objective method to assess progress of outputs and outcomes 	<ul style="list-style-type: none"> Dependent upon viable monitoring systems that have established baseline indicators and targets and have collected reliable data in relation to targets over time, as well as data relating to outcome indicators
Reports and documents	Existing documentation, including quantitative and descriptive information about the initiative, its outputs and outcomes, such as documentation from capacity development activities, donor reports and other evidentiary evidence	<ul style="list-style-type: none"> Cost-efficient 	<ul style="list-style-type: none"> Documentary evidence can be difficult to code and analyse in response to questions Difficult to verify reliability and validity of data
Questionnaires	Provides a standardized approach to obtaining information on a wide range of topics from a large number or diversity of stakeholders (usually employing sampling techniques) to obtain information on their attitudes, beliefs, opinions, perceptions, level of satisfaction, etc. concerning the operations, inputs, outputs, and contextual factors of a UNDP initiative	<ul style="list-style-type: none"> Good for gathering descriptive data on a wide range of topics quickly at relatively low cost Easy to analyse Gives anonymity to respondents 	<ul style="list-style-type: none"> Self-reporting may lead to biased reporting Data may provide a general picture but may lack depth May not provide adequate information on context Subject to sampling bias
Interviews	Solicit person-to-person responses to pre-determined questions designed to obtain in-depth information about a	<ul style="list-style-type: none"> Facilitates fuller coverage, range, and depth of information of a topic 	<ul style="list-style-type: none"> Can be time-consuming Can be difficult to analyse Can be costly Potential for Interviewer to bias client's responses

¹²⁷ Methods described are illustrative and not exhaustive of the types of methods applicable to the UNDP evaluation context.

	person’s impressions or experiences, or to learn more about their answers to questionnaires or surveys		
On-site observation	Entails use of a detailed observation form to record accurate information on site about how a programme operates (ongoing activities, processes, discussions, social interactions, and observable results as directly observed, during the course of an initiative)	<ul style="list-style-type: none"> ▪ Can see operations of a programme as they are occurring ▪ Can adapt to events as they occur 	<ul style="list-style-type: none"> ▪ Can be difficult to categorize or interpret observed behaviours ▪ Can be expensive ▪ Subject to (site) selection bias
Group interviews	A small group (six to eight people) is interviewed together to explore in-depth stakeholder opinions, similar or divergent points of view, or judgements about a development initiative or policy, to collect information around tangible and non-tangible changes resulting from an initiative	<ul style="list-style-type: none"> ▪ Quick, reliable way to obtain common impressions from diverse stakeholders ▪ Efficient way to obtain a high degree of range and depth of information in a short time 	<ul style="list-style-type: none"> ▪ Can be hard to analyse responses ▪ Requires trained facilitator ▪ May be difficult to schedule
Key informants	Qualitative in-depth interviews, often one-on-one, with a wide range of stakeholders who have first-hand knowledge of the initiative’s operations and context. These community experts can provide specific knowledge, and understanding of problems and recommend solutions	<ul style="list-style-type: none"> ▪ Can provide insight on the nature of problems and give recommendations for solutions ▪ Can provide different perspectives on a single issue or on several issues 	<ul style="list-style-type: none"> ▪ Subject to sampling bias ▪ Must have some means to verify or corroborate information
Expert panels	A peer review, or reference group, composed of external experts to provide input on technical or other substance topics covered by the evaluation	<ul style="list-style-type: none"> ▪ Adds credibility ▪ Can serve as added (expert) source of information that can provide greater depth ▪ Can verify or substantiate information and results in topic area 	<ul style="list-style-type: none"> ▪ Cost of consultancy and related expenses if any ▪ Must ensure impartiality and that there are no conflicts of interest

<p>Case studies</p>	<p>Involves comprehensive examination through cross-comparison of cases to obtain in-depth information with the goal to fully understand the operational dynamics, activities, outputs, outcomes and interactions of a development project or programme</p>	<ul style="list-style-type: none"> ▪ Useful to fully explore factors that contribute to outputs and outcomes 	<ul style="list-style-type: none"> ▪ Requires considerable time and resources not usually available for commissioned evaluations ▪ Can be difficult to analyse
<p>Remote/ virtual engagement</p>	<p>In times of crises, access challenges or other inconveniences remote/virtual tools could be feasible options such as Zoom, Skype, WhatsApp, telephone, and others.</p> <p>Additionally, new documentation could be anticipated, if applicable: GIS satellite images, social media analysis, or other big data information analysis.</p>	<ul style="list-style-type: none"> ▪ Can be cheap but costs for registration and different packages need to be considered ▪ Reduces travel costs ▪ Reduces the carbon footprints of individuals and organizations 	<ul style="list-style-type: none"> ▪ Requires a stable internet connection and access to technology (computers, mobile phones, internet, etc.) ▪ Requires specific IT, communication and facilitation skills ▪ Requires special experiences regarding data gathering, data analysis and data interpretation skills especially for new documentation as stated ▪ Meeting virtually is different than meeting in person (loss of certain communicational aspects in human psychology) ▪ Not possible to make observations or meet people coincidentally or more informally. ▪ Certain groups and individuals may not have the opportunity to be involved ▪ More time for conducting an evaluation may need to be envisaged

Annex 3. UNDP Evaluation dispute resolution process

Dispute settlement

Should you or a member of the evaluation team feel unduly pressured to change the findings or conclusions of an evaluation you have been contracted to undertake you are freely able to raise your concerns with the management within UNDP.

Please send your concerns to the Deputy Director of the Region who will ensure a timely response. Please also include the Independent Evaluation Office, in your correspondence (evaluation.office@undp.org).

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 is the principal channel to receive allegations.¹²⁸

Anyone with information regarding fraud 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.

When reporting to the Investigations Hotline, people are encouraged to be as specific as possible, including the basic details of who, what, where, when and how any of these incidents occurred. Specific information will allow OAI to properly investigate the alleged wrongdoing.

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 different ways:

[ONLINE REFERRAL FORM](#) *(You will be redirected to an independent third-party site)*

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

EMAIL directly to OAI at: reportmisconduct@undp.org

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

¹²⁸ <https://www.undp.org/accountability/audit/investigations>

Annex 4. UNDP evaluation report template and quality standards

This **evaluation report template** is intended to serve as a guide for preparing meaningful, useful and credible evaluation reports that meet quality standards. It does not prescribe a definitive section-by-section format that all evaluation reports should follow. Rather, it suggests the areas of content that should be included in a quality evaluation report.

The evaluation report should be complete and logically organized. It should be written clearly and be understandable to the intended audience. In a country context, the report should be translated into local languages whenever possible. The report should include the following:

1. **Title and opening pages** should provide the following basic information:
 - Name of the evaluation intervention.
 - Time frame of the evaluation and date of the report.
 - Countries of the evaluation intervention.
 - Names and organizations of evaluators.
 - Name of the organization commissioning the evaluation.
 - Acknowledgements.

2. **Project and evaluation information details** to be included in all final versions of evaluation reports (non-GEF)¹²⁹ on second page (as one page):

Project/outcome Information		
Project/outcome title		
Atlas ID		
Corporate outcome and output		
Country		
Region		
Date project document signed		
Project dates	Start	Planned end
Total committed budget		
Project expenditure at the time of evaluation		
Funding source		
Implementing party¹³⁰		

¹²⁹ GEF evaluations have their own project information template requirements.

¹³⁰ This is the entity that has overall responsibility for implementation of the project (award), effective use of resources and delivery of outputs in the signed project document and workplan.

Evaluation information		
Evaluation type (project/ outcome/thematic/country programme, etc.)		
Final/midterm review/ other		
Period under evaluation	Start	End
Evaluators		
Evaluator email address		
Evaluation dates	Start	Completion

3. **Table of contents**, including boxes, figures, tables, and annexes with page references.
4. **List of acronyms and abbreviations.**
5. **Executive summary (four/ five page maximum).** A stand-alone section of two to three pages that should:
 - Briefly describe the intervention of the evaluation (the project(s), programme(s), policies, or other intervention) that was evaluated.
 - Explain the purpose and objectives of the evaluation, including the audience for the evaluation and the intended uses.
 - Describe key aspects of the evaluation approach and methods.
 - Summarize principle findings, conclusions and recommendations.
6. **Introduction**
 - Explain why the evaluation was conducted (the purpose), why the intervention is being evaluated at this point in time, and why it addressed the questions it did.
 - Identify the primary audience or users of the evaluation, what they wanted to learn from the evaluation and why, and how they are expected to use the evaluation results.
 - Identify the intervention being evaluated (the project(s) programme(s) policies or other intervention).
 - Acquaint the reader with the structure and contents of the report and how the information contained in the report will meet the purposes of the evaluation and satisfy the information needs of the intended users.
7. **Description of the intervention** provides the basis for report users to understand the logic and assess the merits of the evaluation methodology and understand the applicability of the evaluation results. The description needs to provide sufficient detail for the report user to derive meaning from the evaluation. It should:
 - Describe **what is being evaluated, who seeks to benefit** and the **problem or issue** it seeks to address.

- Explain the **expected results model or results framework, implementation strategies** and the key **assumptions** underlying the strategy / theory of change.
 - Link the intervention to **national priorities**, UNSDCF priorities, corporate multi-year funding frameworks or Strategic Plan goals, or other **programme or country-specific plans and goals**.
 - Identify the **phase** in the implementation of the intervention and any **significant changes** (e.g., plans, strategies, logical frameworks, theory of change) that have occurred over time, and explain the implications of those changes for the evaluation.
 - Identify and describe the **key partners** involved in the implementation and their roles.
 - Include data and an analysis of **specific social groups** affected. Identify **relevant cross-cutting issues** addressed through the intervention, i.e., gender equality, human rights, vulnerable/ marginalized groups, leaving no one behind.
 - Describe the **scale of the intervention**, such as the number of components (e.g., phases of a project) and the size of the target population (men and women) for each component.
 - Indicate the **total resources**, including human resources and budgets.
 - Describe the context of the **social, political, economic, and institutional factors**, and the **geographical landscape** within which the intervention operates, and explain the challenges and opportunities those factors present for its implementation and outcomes.
 - Point out **design weaknesses** (e.g., intervention logic, theory of change) or other **implementation constraints** (e.g., resource limitations).
8. **Evaluation scope and objectives.** The report should provide a clear explanation of the evaluation’s scope, primary objectives and main questions.
- **Evaluation scope.** The report should define the parameters of the evaluation, for example, the time period, the segments of the target population and geographic area included, and which components, outputs or outcomes were or were not assessed.
 - **Evaluation objectives.** The report should spell out the types of decisions the evaluation will feed into, the issues to be considered in making those decisions and what the evaluation will need to achieve to contribute to those decisions.
 - **Evaluation criteria.** The report should define the evaluation criteria or performance standards used¹³¹ and explain the rationale for selecting those particular criteria.
 - **Evaluation questions.** The report should detail the main evaluation questions addressed by the evaluation and explain how the answers to those questions address the information needs of users.
9. **Evaluation approach and methods.**¹³² The evaluation report should describe in detail the selected methodological approaches, methods and analysis; the rationale for their selection; and how, within the time and money constraints, the approaches and methods employed yielded data that helped to answer the evaluation questions and achieved the evaluation purposes. **The report should specify how gender equality, disability, vulnerability and social inclusion were addressed in the methodology, including how data collection and analysis methods integrated gender considerations, use of disaggregated**

¹³¹ The evaluation criteria most commonly applied to UNDP evaluations are the OECD-DAC criteria of relevance, coherence, efficiency, effectiveness and sustainability.

¹³² All aspects of the described methodology need to receive full treatment in the report. Some of the more detailed technical information may be contained in annexes to the report.

data and outreach to diverse stakeholder groups. The description should help the report users judge the merits of the methods used in the evaluation and the credibility of the findings, conclusions and recommendations. The description of methodology should include discussion of each of the following:

- **Evaluation approach.**
- **Data sources:** the sources of information (documents reviewed and stakeholders met) as well as the rationale for their selection and how the information obtained addressed the evaluation questions.
- **Sample and sampling frame.** If a sample was used, describe the sample size and characteristics, the sample selection criteria; the process for selecting the sample (e.g. random, purposive); if applicable, how comparison and treatment groups were assigned; and the extent to which the sample is representative of the entire target population, including discussion of the limitations of sample for generalizing results.
- **Data collection procedures and instruments:** methods or procedures used to collect data, including discussion of data-collection instruments (e.g., interview protocols), their appropriateness for the data source, and evidence of their reliability and validity, as well as gender-responsiveness.
- **Performance standards:**¹³³ the standard or measure that will be used to evaluate performance relative to the evaluation questions (e.g. national or regional indicators, rating scales).
- **Stakeholder participation:** who participated and how the level of involvement of men and women contributed to the credibility of the evaluation and the results.
- **Ethical considerations:** including the measures taken to protect the rights and confidentiality of informants (see UNEG ‘Ethical Guidelines for Evaluators’ for more information).¹³⁴
- **Background information on evaluators:** the composition of the evaluation team, the background and skills of team members, and the appropriateness of the technical skill mix, gender balance and geographical representation for the evaluation.
- **Major limitations of the methodology** should be identified and openly discussed, as well as any steps taken to mitigate them.

10. **Data analysis.** The report should describe the procedures used to analyse the data collected to answer the evaluation questions. It should detail the various steps and stages of analysis that were carried out, including the steps to confirm the accuracy of data and the results for different stakeholder groups (men and women, different social groups, etc.). The report should also discuss the appropriateness of the analyses to the evaluation questions. Potential weaknesses in the data analysis and gaps or limitations of the data should be discussed, including their possible influence on the way findings may be interpreted and conclusions drawn.

11. **Findings** should be presented as statements of fact that are based on analysis of the data. They should be structured around the evaluation questions so that report users can readily make the connection between what was asked and what was found. Variances between planned and actual results should be explained, as well as factors affecting the achievement

¹³³ A summary matrix displaying, for each of the evaluation questions, the data sources, data collection tools or methods and the standard or measure by which each question was evaluated. This is a good illustrative tool to simplify the logic of the methodology for the report reader.

¹³⁴ UNEG, 2020, Ethical Guidelines for Evaluation: <http://www.unevaluation.org/document/detail/2866>

of intended results. Assumptions or risks in the project or programme design that subsequently affected implementation should be discussed. Findings should reflect gender equality and women's empowerment, disability and other cross-cutting issues, as well as possible unanticipated effects.

12. **Conclusions** should be comprehensive and balanced and highlight the strengths, weaknesses and outcomes of the intervention. They should be well substantiated by the evidence and logically connected to evaluation findings. They should respond to key evaluation questions and provide insights into the identification of and/or solutions to important problems or issues pertinent to the decision-making of intended users, including issues in relation to gender equality and women's empowerment as well as to disability and other cross-cutting issues.
13. **Recommendations.** The report should provide a reasonable number of practical, actionable and feasible recommendations directed to the intended users of the report about what actions to take or decisions to make. The recommendations should be specifically supported by the evidence and linked to the findings and conclusions around key questions addressed by the evaluation. They should address sustainability of the initiative and comment on the adequacy of the project exit strategy, if applicable. Recommendations should also provide specific advice for future or similar projects or programming. Recommendations should address any gender equality and women's empowerment issues and priorities for action to improve these aspects. Recommendations regarding disability and other cross-cutting issues also need to be addressed.
14. **Lessons learned.** As appropriate and/or if requested in the TOR, the report should include discussion of lessons learned from the evaluation, that is, new knowledge gained from the particular circumstance (intervention, context, outcomes, even evaluation methods) that are applicable to a similar context. Lessons should be concise and based on specific evidence presented in the report. Gender equality and women's empowerment, disability and other cross-cutting issues should also be considered.
15. **Report annexes.** Suggested annexes should include the following to provide the report user with supplemental background and methodological details that enhance the credibility of the report:
 - TOR for the evaluation.
 - Additional methodology-related documentation, such as the evaluation matrix and data-collection instruments (questionnaires, interview guides, observation protocols, etc.) as appropriate.
 - List of individuals or groups interviewed or consulted, and sites visited. This can be omitted in the interest of confidentiality if agreed by the evaluation team and UNDP.
 - List of supporting documents reviewed.
 - Project or programme results model or results framework.
 - Summary tables of findings, such as tables displaying progress towards outputs, targets, and goals relative to established indicators.
 - Pledge of ethical conduct in evaluation signed by evaluators.

Annex 5. Management response template

UNDP management response template

[Name of the Evaluation] Date:

Prepared by:

Position:

Unit/Bureau:

Cleared by:

Position:

Unit/Bureau:

Input into and update in ERC:

Position:

Unit/Bureau:

Evaluation recommendation 1.				
Management response:				
Key action(s)	Completion date	Responsible unit(s)	Tracking*	
			Comments	Status (initiated, completed or no due date)
1.1				
1.2				
1.3				
Evaluation recommendation 2.				
Management response:				
Key action(s)	Completion date	Responsible unit(s)	Tracking	
			Comments	Status (initiated, completed or no due date)
2.1				
2.2				
2.3				
Evaluation recommendation 3.				
Management response:				
Key action(s)	Completion date	Responsible unit(s)	Tracking	
			Comments	Status (initiated, completed or no due date)
3.1				
3.2				

* Status of implementation is tracked electronically in the ERC database.



SECTION 5

ROLES AND RESPONSIBILITIES IN DECENTRALIZED EVALUATION

5. ROLES AND RESPONSIBILITIES IN DECENTRALIZED EVALUATIONS

Section 5 summarizes the roles and responsibilities in the development and implementation of evaluation plans and in the implementation, dissemination and use of decentralized evaluations.

The table below summarizes roles and responsibilities in the development and oversight of evaluation plans and in the implementation, dissemination and use of decentralized evaluations.

ROLE	PERSON/ORGANIZATION	RESPONSIBILITIES
EVALUATION COMMISSIONER/ OWNER	<p>The agency or entity calling for the evaluation to be carried out. Within UNDP, responsibility for decentralized evaluations rests ultimately with the global, regional and country office senior management, who “own” the evaluation plan for their programme, i.e.</p> <ul style="list-style-type: none"> • Bureau directors • Resident representatives 	<ul style="list-style-type: none"> ▪ Lead and ensure the development of a comprehensive, representative, strategic and costed evaluation plan ▪ Ensure the timely implementation of the evaluation plan ▪ Promote joint evaluation work with the United Nations system and other partners ▪ Ensure evaluability of UNDP initiatives, that clear and comprehensive results frameworks (theory of change) are in place ▪ Ensure that all required data and relevant documentation are made available ▪ Appoint the evaluation manager ▪ Safeguard the independence of the evaluation exercise and ensure quality of evaluations ▪ Ensure that all steps in the evaluation process are applied as defined in the UNDP evaluation guidelines ▪ Ensure that gender equality and woman’s empowerment and other cross-cutting issues are considered in all steps of the evaluation process ▪ Ensure management responses are prepared for all evaluations with time-bound key actions for their implementation ▪ Accountable for the quality and approval of final terms of reference (TORs), final evaluation reports and management responses before final submission to the ERC
EVALUATION MANAGER	Monitoring and evaluation (M&E) specialist or officer	<ul style="list-style-type: none"> ▪ Lead the evaluation process and participate in all of its stages - evaluability assessment, preparation, implementation, management and use

(not the programme/project manager)

In cases where there is no dedicated M&E specialist/officer in place, the programme unit **must ensure that the evaluation manager is not the programme/ project manager.**

Regional evaluation focal points can provide additional support in case of limited evaluation capacity at the country office level.

Some of the listed responsibilities can be performed by the M&E focal point

- Establish and lead the **evaluation reference group**, where applicable
- Lead the **evaluability assessment**
- Lead the development of the **TOR**, adhering to the UNDP TOR template, and ensure the inclusion of gender equality and women’s empowerment and other cross-cutting issues
- Prepare relevant supporting documents for the tender of the evaluations / recruitment of evaluators and share with operations teams
- Participate in the selection and recruitment of external evaluators
- Safeguard the independence of evaluations
- With the support of the project or programme officer, provide the evaluators with administrative support and required data and documentation
- Organize the kick-off meeting to introduce the evaluators to the evaluation reference group, where applicable, and discuss the evaluation assignment
- Liaise with the programme/ project manager(s) throughout the evaluation process
- Connect the evaluators with the wider programme unit, senior management and key evaluation stakeholders, and ensure a fully inclusive and transparent approach to the evaluation
- Circulate, review and approve **inception reports** including evaluation questions and methodologies
- Ensure that gender equality and women’s empowerment and other cross-cutting issues are considered in inception reports, including a gender-responsive methodology
- Circulate, review and comment on **draft evaluation reports** (according to the TOR and inception report)
- Ensure that gender equality and women’s empowerment and other cross-cutting issues are considered in draft evaluation reports and ensure that all and respective evaluation questions are answered, and relevant data, disaggregated by sex, is presented, analysed and interpreted
- Collect and consolidate comments on draft evaluation reports in one feedback document (audit trail) and share with the evaluation team for finalization of the evaluation report Review **final evaluation report** to ensure compliance to the UNDP report template

		<p>and quality assurance and seek final approval of the commissioner of the evaluation</p> <ul style="list-style-type: none"> ▪ Contribute to the development of management responses and key actions to all recommendations addressed to UNDP ▪ Ensure evaluation TOR, final evaluation reports, management responses, lessons learned, and other relevant information are publicly available through the Evaluation Resource Center (ERC) within the specified time frame ▪ Facilitate, monitor and report on a quarterly basis on the implementation of management responses and key actions ▪ Facilitate knowledge sharing and use of findings in programming and decision-making, e.g. by organizing separate meetings and events or integrating findings and recommendations into regular meetings ▪ Encourage the preparation of separate “Evaluation Briefs” together with the programme / project officer and communication officer for wider distribution and learning
<p>PROGRAMME/ PROJECT MANAGER</p>	<p>UNDP manager of the programme/ outcome/ project under evaluation</p> <p>Typically:</p> <ul style="list-style-type: none"> • Senior managers for country programmes • Global programme/ project managers for global programmes • Regional programme/ project managers • Programme officers (programme team leaders, programme analysts) 	<ul style="list-style-type: none"> ▪ Participate and involve relevant stakeholders in the development of the evaluation plan ▪ Ensure evaluability of UNDP initiatives in a given project, programme thematic or results area ▪ Support the evaluability assessment of projects or programmes with planned evaluations ▪ Support the establishment of the evaluation reference group with key project partners where needed and participate in calls/ meetings on request ▪ Provide inputs/ advice to the evaluation manager and evaluation reference group on the detail and scope of the TOR for the evaluation and how the findings will be used ▪ Ensure and safeguard the independence of evaluations ▪ Provide the evaluation manager with all required data (e.g. relevant monitoring data) and documentation (reports, minutes, reviews, studies, etc.), contacts/ stakeholder list etc. ▪ Ensure that data and documentation in general, but in particular related to gender equality and women’s empowerment and

		<p>other cross-cutting issues, are made available to the evaluation manager</p> <ul style="list-style-type: none"> ▪ Provide comments and clarification on the TOR, inception report and draft evaluation reports ▪ Respond to evaluation recommendations by providing management responses and key actions to all recommendations addressed to UNDP ▪ Ensure dissemination of the evaluation report to all the stakeholders including the project board ▪ Implement relevant key actions on evaluation recommendations
<p>EVALUATION SUPPORT, QUALITY ASSURANCE AND ERC PORTAL MANAGEMENT</p>	<p>M&E specialist or officer</p> <p>M&E focal points</p> <p>(Global, regional and country office)</p>	<ul style="list-style-type: none"> ▪ Upload the evaluation plan to the ERC, manage required changes to the evaluation plan and get approval for changes from the regional evaluation focal point ▪ Report to management on compliance with the evaluation plan, completion of management responses and key actions and quality assessment results ▪ Support the evaluation manager, if different, in all the steps of the evaluation process ▪ Participate in the evaluation reference group, if applicable ▪ Review and support the approval of the evaluation TOR, ensuring they meet UNDP guidance requirements, including gender equality and women’s empowerment and other cross-cutting issues ▪ Participate in the selection/ recruitment of external evaluators, if applicable ▪ Review and support the approval of the evaluation inception report, ensuring it meets UNDP requirements, including gender equality and women’s empowerment and other cross-cutting issues ▪ Review and support the approval of the draft and final evaluation reports, and also ensure that gender equality and women’s empowerment and other cross-cutting issues have been included ▪ Review the management responses and key actions ▪ Ensure evaluation TOR, final evaluation reports, management responses, lessons learned and other relevant information are publicly available through the ERC within the specified time frame ▪ Support the evaluation manager, if different, in the facilitation of knowledge sharing and

		<p>use of findings in programming and decision-making</p> <ul style="list-style-type: none"> Facilitate, monitor and report on a quarterly basis on the implementation of management responses and key actions The regional M&E officer ensures the management response tracking through the ERC and supports M&E capacity development and knowledge sharing. They close the management response process when all planned actions have been completed or after five years
EVALUATION REFERENCE GROUP (OPTIONAL)	<p>Key project/ outcome stakeholders, government partners and/ or donors including representatives from project management boards</p> <p>Should also include persons who have some technical expertise and experience in evaluation design, implementation and quality assurance. The evaluation reference group should not include representatives from the project under evaluation, to avoid conflict of interest</p>	<ul style="list-style-type: none"> Perform advisory role throughout the evaluation process providing inputs into and review of TOR, inception reports and draft evaluation reports Ensure that gender equality and women’s empowerment and other cross-cutting issues are considered in all steps of the evaluation process Ensure that the United Nations Evaluation Group (UNEG) evaluation standards are adhered to, including safeguarding of transparency and independence Provide advice on the evaluation relevance, the appropriateness of evaluation questions and methodology, and the extent to which conclusions are credible, considering the evidence presented, and recommendations action-oriented Support and provide input to the development of the management responses and key actions
EVALUATION PARTNERS	Government partners, stakeholders, donors	<ul style="list-style-type: none"> Participate in the review of key evaluation deliverables, including the TOR, inception report, and successive versions of the draft evaluation report Ensure that data and documentation in general, but in particular relating to gender equality and women’s empowerment and other cross-cutting issues, are made available to the evaluation manager Where donors are leading the evaluation process, UNDP should ensure the inclusion of key elements required by the programme unit including gender equality and women’s empowerment and other cross-cutting issues Where UNDP is leading the process, it should be flexible to suggestions and requirements from the donors, but also retain the UNDP

		<p>requirements for the TOR, evaluation implementation and evaluation reports</p> <ul style="list-style-type: none"> ▪ Evaluations of donor-funded projects should be uploaded to the ERC along with recommendations, management responses and key actions that relate to UNDP
INDEPENDENT EVALUATORS	External evaluation experts and/ or firms	<ul style="list-style-type: none"> ▪ Fulfil the contractual arrangements under the TOR ▪ Develop the evaluation inception report, including an evaluation matrix and a gender-responsive methodology, in line with the TOR, UNEG norms and standards and ethical guidelines ▪ Conduct data collection and field visits according to the TOR and inception report ▪ Produce draft reports adhering to UNDP evaluation templates, and brief the evaluation manager, programme/ project managers and stakeholders on the progress and key findings and recommendations ▪ Consider gender equality and women’s empowerment and other cross-cutting issues, check if all and respective evaluation questions are answered, and relevant data, disaggregated by sex, is presented, analysed and interpreted ▪ Finalize the evaluation report, incorporating comments and questions from the feedback/ audit trail. Record own feedback in the audit trail
REGIONAL EVALUATION TECHNICAL SUPPORT AND QUALITY ASSURANCE	Regional M&E evaluation focal points, specialists and advisers	<ul style="list-style-type: none"> ▪ Support the evaluation process and ensure compliance with corporate standards ▪ Oversee and support evaluation planning and the uploading, implementation and adjustment of evaluation plans in ERC ▪ Review evaluation plan adjustment requests (date changes, additions, deletions), ensuring the rationale for the changes is appropriate, and approve through the ERC ▪ Provide technical support to country offices including: advice on the development of TORs; recruitment of evaluators and maintaining evaluator rosters; implementation of evaluations; and

		<p>finalization of evaluations, management responses and key actions</p> <ul style="list-style-type: none"> ▪ Ensure that gender equality and women’s empowerment and other cross-cutting issues are considered in all steps of the evaluation process ▪ Ensure management response tracking through ERC and support M&E capacity development and knowledge sharing ▪ Resolve disputes when issues arise in the implementation of evaluations
GLOBAL EVALUATION QUALITY ASSURANCE	Bureau for Policy and Programme Support	<ul style="list-style-type: none"> ▪ Provide global oversight of decentralized evaluation and guidance ▪ Coordinate communication between UNDP management and the IEO ▪ Advise country offices and regional bureaux on the decentralized evaluation function for UNDP ▪ Support M&E focal points to ensure that evaluation plans are properly implemented ▪ Monitor the implementation of management responses for independent and decentralized evaluations ▪ In cooperation with IEO, provide guidance to UNDP implementing units on the use of evaluation findings and lessons learned
GLOBAL AND REGIONAL EVALUATION GUIDANCE AND OVERSIGHT	Independent Evaluation Office	<ul style="list-style-type: none"> ▪ Provide norms, standards, guidelines and tools to support the quality enhancement of evaluations ▪ Oversee and report on decentralized evaluation implementation and adherence to evaluation plans ▪ Implement and report on the annual quality assessment of all decentralized evaluations through an independent quality assessment review panel ▪ Manage and maintain the ERC and provide support for technical issues in use of the ERC ▪ Provide guidance on decentralized evaluations through the regional evaluation focal points

5.1. Task Sheet: UNDP Evaluation Manager roles and responsibilities

The role and responsibility of the **UNDP Evaluation Manager** is to lead and participate in the entire evaluation process. Details for all four stages/ steps of an evaluation process can be found in Section 4 of the UNDP Evaluation Guidelines. They are:

- (a) Evaluability assessment
- (b) Evaluation preparation
- (c) Evaluation management
- (d) Use of the evaluation

Overall, together with the evaluation commissioner, the evaluation manager safeguards the independence of the evaluation.

Evaluability Assessment (Step 1)

Under the guidance of the evaluation commissioner, the evaluation manager and/ or M&E focal point lead and undertake the **evaluability assessment** of the intervention to be evaluated (“evaluand”) in collaboration with the UNDP programme units and national stakeholders.

Evaluation preparation (Step 2)

- 1 Establish and lead the **evaluation reference group**, where applicable.
- 2 Lead the **development of the TOR**, ensuring a participatory process, adherence to the UNDP TOR template, and the inclusion of gender equality and women’s empowerment and other cross-cutting issues.
- 3 With the help of the project or programme manager, **compile basic documentation** that will be provided to the evaluation team.
- 4 Prepare **relevant supporting documents** for the tender of the evaluation/ recruitment of evaluators and share with the operations teams.
- 5 Participate in the **selection/ recruitment of external evaluators** together with the reference group, where applicable.

Evaluation management (Step 3)

- 1 Provide the evaluators with **administrative support** and required data and documentation.
- 2 Organize a **kick-off meeting** to introduce the evaluators to the evaluation reference group, if applicable, and discuss the evaluation assignment.
- 3 Liaise with the programme/ project manager(s) throughout the evaluation process and **connect the evaluators** with the wider programme unit, senior management and key evaluation stakeholders, to ensure a fully inclusive and transparent approach to the evaluation.
- 4 Circulate, review and approve the **inception report** including evaluation questions and methodologies. Ensure that gender equality and women’s empowerment and other cross-cutting issues are considered in the inception report, including a gender-responsive methodology.

- 5 Circulate, review and comment on the **draft evaluation report** and ensure compatibility with the TOR and inception report. Ensure that gender equality and women's empowerment and other cross-cutting issues are considered, all evaluation questions are answered, and relevant data, disaggregated by sex, is presented, analysed and interpreted.
- 6 Collect and consolidate **comments on draft evaluation reports** in one feedback document (audit trail) and share them with the evaluation team for finalization of the evaluation report.
- 7 Review the **final evaluation report**, ensuring compliance to the UNDP report template and quality assurance, and seek final approval of the commissioner of the evaluation.

Use of the evaluation (step 4)

- 1 Contribute to the **development of the management response** and key actions to all recommendations addressed to UNDP.
- 2 Ensure that the evaluation TOR, final evaluation report, management response, lessons learned, and other relevant information are **publicly available** through the ERC within the specified time frame.
- 3 Facilitate, monitor and report on a quarterly basis on the **implementation of the management responses** and key actions.
- 4 **Facilitate knowledge sharing** and use of findings in programming and decision-making, e.g. organizing separate meetings and events or integrating findings and recommendations in regular meetings.
- 5 Together with the programme/ project officer and the communication officer, encourage the **preparation of other evaluation "products"**, such as a separate evaluation brief, for example, for wider distribution of evaluation findings and learning.

5.2. Task Sheet: UNDP M&E focal point roles and responsibilities

If the M&E focal point is not the evaluation manager, then they will be engaged in the evaluation process as described below. If the M&E focal point *is* the evaluation manager, then the assignments from the Evaluation Manager task sheet are applicable.

M&E focal points are appointed at global, regional, and country office level. Overall, the M&E focal point supports the evaluation manager in all steps of the evaluation process and ensures compliance with corporate standards. In case of possible disputes, the regional M&E focal point can be contacted to offer resolutions.

Evaluability assessment (Step 1)

- 1 Support the evaluation manager to undertake the evaluability assessment of the intervention to be evaluated (“evaluand”) in collaboration with the UNDP programme units and national stakeholders.

Evaluation preparation (Step 2)

- 2 Upload the evaluation plan to the ERC, manage required changes to the evaluation plan and get approval for changes from the regional M&E focal point.
- 3 The regional evaluation officer reviews the evaluation plan adjustment requests (date changes, additions, deletions), ensuring the rationale for the changes is appropriate, and approves through the ERC.
- 4 Report to management on corporate compliance with the evaluation plan.
- 5 Participate in the evaluation reference group, if applicable.
- 6 Review and support approval of the evaluation TOR, ensuring they meet UNDP guidance requirements, including gender equality and women’s empowerment and other cross-cutting issues.
- 7 Participate in the selection/ recruitment of external evaluators, if applicable.
- 8 The regional M&E officer provides technical support to country offices including advice on the development of TORs (including gender equality and women’s empowerment and other cross-cutting issues), recruitment of evaluators and maintaining evaluator rosters.

Evaluation management (Step 3):

- 1 Review and support the approval of the **evaluation inception report**, ensuring it meets UNDP requirements, including gender equality and women’s empowerment and other cross-cutting issues.
- 2 Review and support the approval of the **draft and final evaluation reports**, including gender equality and women’s empowerment and other cross-cutting issues, seeking support from the regional evaluation officer if requested.
- 3 Ensure that the evaluation TOR, final evaluation reports, management responses, lessons learned, and other relevant information are publicly available through the ERC within the specified time frame.

Use of the evaluation (Step 4)

- 1 Review the management response, key actions and quality assessment results and ensure completion of this step together with the regional M&E focal point, if requested.
- 2 Support the Evaluation Manager in the facilitation of knowledge sharing and use of findings in programming and decision-making.
- 3 Facilitate, monitor and report on a quarterly basis on the implementation of management responses and key actions. The regional evaluation officer ensures management response tracking through the ERC and supports M&E capacity development and knowledge sharing. They close the management response process when all planned actions have been completed or after five years.



SECTION 6

EVALUATION QUALITY ASSESSMENT



CONTENTS

6. QUALITY ASSESSMENT OF DECENTRALIZED EVALUATIONS	1
6.1 Purpose and scope	1
6.2 Quality assessment process	2
6.3 Roles and responsibilities.....	3
6.4 Quality assessment review pool	4
6.5 Quality assessment reporting	4
6.6 Quality assessment sections and weighting	5
6.7 Quality assessment question ratings	6
6.8 Quality assessment tool.....	6
6.9 Supporting documentation.....	6
6.10 Quality assessment questions.....	7
6.10.1 Terms of Reference and Design	7
6.10.2 Report structure, methodology and data sources.....	8
6.10.3 Cross-cutting issues.....	10
6.10.4 Findings, conclusions and recommendations	12
6.10.5 Management Response	13
6.10.6 General findings and lessons learned.	13
6.10.7 Validation of the ratings given by GEF terminal evaluations.....	14
6.10.8 Summary quality assessment result	16
6.11 UN-SWAP evaluation performance indicator and assessment.....	17
6.11.1 What is the UN-SWAP evaluation performance indicator?.....	17
6.11.2 The evaluation performance indicator criteria and scorecard	17
6.11.3 The assessment process.....	18
6.11.4 Reporting.....	18
Figure 1. Quality assessment process	3
Figure 2. ERC quality assessment summary report by region	4
Figure 3. Quality assessment rating scale.....	5

6. QUALITY ASSESSMENT OF DECENTRALIZED EVALUATIONS¹

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

High quality evaluations are critical for results-based management, knowledge generation, and accountability to programme partners. One of the requirements of the UNDP Evaluation Policy is that programme units—headquarters bureaux, regional bureaux and country offices—ensure that evaluations inform programme management and contribute to development results.² There is therefore increased emphasis to strengthen support for decentralized evaluations (those carried out by programme units) in order to improve their compliance with the Evaluation Policy, improve the quality of evaluations and increase the use of evaluations by policymakers and stakeholders.

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

6.1 Purpose and scope

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

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

- Improving the quality of evaluative evidence to better manage contributions to development results.
- Supporting accountability by providing an independent assessment of the quality of decentralized evaluation reports to the UNDP Executive Board and management.
- Strengthening consistency in evaluation reporting and quality across projects.

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

² UNDP Evaluation Policy, <http://web.undp.org/evaluation/policy.shtml>

- Supporting bureau oversight functions by providing concurrent feedback through detailed analysis of the quality of evaluation reports, with recommendations for their improvement.
- Contributing to corporate lessons learned by drawing from good evaluations in the annual report on evaluation.

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

Quality assessments are carried out for all decentralized evaluations conducted by UNDP, as well as the United Nations Capital Development Fund and United Nations Volunteers programme, outcome, project and programme evaluations and thematic evaluations. Feedback from IEO can be used by programme units and country offices to make adjustments that will strengthen areas of the evaluative evidence and the report, as well as adjust the management and implementation of evaluations to ensure usable findings and recommendations and the overall utility of decentralized evaluation reports. The quality assessment questions are in line with and reflect the UNDP quality standards for programming.³

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

6.2 Quality assessment process

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

1. **Posting evaluations to the Evaluation Resource Center (ERC)**⁴
 - The programme unit posts an electronic and printable copy of the terms of reference (TOR) for an evaluation and the final evaluation report on the ERC within two weeks of completion.
 - Only final documents should be uploaded. Drafts should not be uploaded as the ERC is a public website.
 - The management response and key actions should be uploaded within six weeks of completion of the report.
2. **Verification**
 - The IEO will verify if a report posted on the ERC is part of the programme unit evaluation plan and whether it is the final document.
 - If a TOR or evaluation appear to be in draft and not final versions, or if supporting annexes are not uploaded, IEO will contact the country office and regional office to ensure that the correct documents are uploaded.⁵

³ Access at:

https://popp.undp.org/_layouts/15/WopiFrame.aspx?sourcedoc=/UNDP_POPP_DOCUMENT_LIBRARY/Public/PPM_Programming%20Standards_Quality%20Standards%20for%20Programming.docx&action=default

⁴ Access at: <http://erc.undp.org>

⁵ The ERC is a public website and therefore all documents should be final and of high quality. The quality assessment ratings are available only to UNDP.

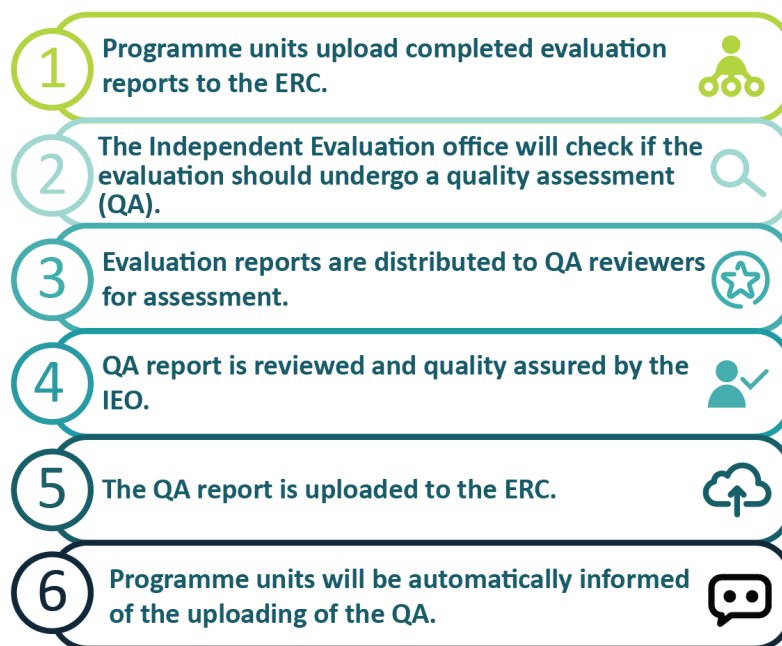
3. Quality assessment

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

4. Feedback

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

Figure 1. Quality assessment process



6.3 Roles and responsibilities

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

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

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

6.4 Quality assessment review pool

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

6.5 Quality assessment reporting

The Quality Assessment process is undertaken at the end of each quarter, April for Q1, July for Q2, October for Q3, and January for Q4. In high completion periods, normally November through to January, the quality assessment process is undertaken more frequently. A quality assessment report for an individual evaluation is available as soon as the IEO performs quality assurance checks on the assessment, normally within two weeks of completion and submission of the quality assessment report. Results at the global, regional and country office levels are available through the ERC.

Figure 2. ERC quality assessment summary report by region



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

6.6 Quality assessment sections and weighting

The key parameters of a quality assessment draw on the basic quality requirements for acceptable evaluation reports as outlined in the Evaluation Guidelines. Overall, the quality assessment process includes four weighted sections and 39 questions. Questions may be left unrated by reviewers where they find them not relevant due to the direction of the TOR or the context of the intervention under evaluation. Quality assessment sections include:

- **Terms of reference and design:** Five questions weighted 15 percent.
 - Do the terms of reference appropriately and clearly outline the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?
- **Report structure, methodology and data sources:** 8 questions weighted 30 percent.
 - Does the evaluation report appropriately and clearly address the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?
- **Cross-cutting issues:** Five questions weighted 15 percent.
 - Does the evaluation report address LNOB and other cross cutting issues where relevant?
- **Findings, conclusions and recommendations:** Eleven questions weighted 40 percent.
 - Does the report clearly and concisely outline and support its findings, conclusions and recommendations?
- **Management response:** two questions not included in the overall QA scoring.
 - Is the management response clear, realistic and actionable?

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

Figure 3. Quality assessment rating scale

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

6.7 Quality assessment question ratings

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

6.8 Quality assessment tool

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

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

6.9 Supporting documentation

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

For UNDP projects the documentation includes:

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

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

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

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

6.10 Quality assessment questions

6.10.1 Terms of Reference and Design

Do the terms of reference appropriately and clearly outline the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines? (Section 4.3.2)	
1.1	<p>Does the TOR clearly outline the purpose, scope, and objectives of the evaluation in a logical and realistic manner? (Section 4, page 14-15-16 and Annex 1, page 39-40)</p> <ul style="list-style-type: none"> ▪ Does it follow the proposed structure detailed in the UNDP evaluation guidelines? ▪ Does it provide details on evaluation context? ▪ Does it include information regarding the results framework in the main text or annexes?
1.2	<p>Does the TOR clearly detail timescales and allocation of days for the evaluation? (Section 4, page 15, Annex 1, page 49)</p> <ul style="list-style-type: none"> ▪ Is the allocation of days detailed for each step of the evaluation process and adequate given the scope of the evaluation? ▪ Does it specify details (including the composition) of the evaluation team and roles and responsibilities of individual team members?
1.3	<p>Does the TOR outline the involvement of evaluation partners, stakeholders and feedback mechanism? (Section 4, page 15, Annex 1, page 47)</p> <ul style="list-style-type: none"> ▪ Does the ToR identify which partners will be involved in the evaluation and how? ▪ Is the audit trail-feedback process, corresponding/reporting parties, and timeline outlined?
1.4	<p>Is there an outline of the evaluation approach and methodology in the TOR? (Section 4, page 14-15 and Annex 1, page 45)</p> <ul style="list-style-type: none"> ▪ Are the number of key evaluation questions appropriate given the scope of the evaluation? ▪ Is there guidance for a general methodological approach? ▪ Are available data sources as well as data requirements of the evaluation explained?
1.5	<p>Does the TOR include a requirement for a “Leave no One Behind”, gender responsive, and disability inclusive evaluation? (Section 4, page 16-17, and page 20, Annex 1, page 44)</p> <ul style="list-style-type: none"> ▪ Does the ToR detail LNOB, gender, and disability inclusive specific questions, supplementing or integrated under the key evaluation questions? ▪ Does the TOR require the evaluation to explain how the tools, methodologies, and data analysis adopted will capture the LNOB, gender and other cross cutting issues?

6.10.2 Report structure, methodology and data sources

Does the evaluation report appropriately and clearly address the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?	
STRUCTURE	
2.1	<p>Is the evaluation report well balanced and structured?</p> <ul style="list-style-type: none"> Does it follow the proposed evaluation report structure detailed in the UNDP Evaluation guidelines (section 4, page 29-30 and Annex 4) Does it cover all evaluation criteria as requested in the ToR? Is the length of the main report reasonable for the requirements of the evaluation? Does the report include all annexes/appendices? (TOR for the evaluation, Evaluation matrix and data collection instruments, List of individuals or groups interviewed or consulted, and sites visited, List of supporting documents reviewed)
2.2	<p>Does the report have a high-quality Executive summary that captures the main elements of the report? (Section 4, Annex 4, page 57)</p> <ul style="list-style-type: none"> Does it include a brief overview of the purpose and objective- why the evaluation is being undertaken? Does it include a summary of the evaluation scope and main areas of inquiry - what is being evaluated? Does it include a summary of key evaluation findings? Does it include a summary of concluding statements and proposed recommendations?
2.3	<p>Is the nature and extent of partners' and stakeholder's role and involvement in project implementation explained adequately?</p>
METHODOLOGY	
2.4	<p>Is there an examination of intervention`s ToC/ results pathway or an attempt to reconstruct the ToC/ results pathway?</p>
2.5	<p>Is the evaluation's methodological approach clearly outlined? (Section 4 Annex 4, page 58-59)</p> <ul style="list-style-type: none"> Did the evaluator explain the general evaluation approach/framework used to answer Evaluation Questions? (i.e., if impact/attribution is to be uncovered, impact assessment methodology; if contribution is to be uncovered other frameworks such as outcome harvesting, contribution analysis, process tracing etc...) Did the evaluator make a distinction between Methodology adopted for the evaluation and data collection tools (which is covered under question 5)? Are any changes from the approach outlined in the TOR detailed?
DATA COLLECTION and ANALYSIS	
2.6	<p>Is the data collection and analysis strategy clearly defined? (Section 4 Annex 4, page 59)</p> <ul style="list-style-type: none"> Are data sources being clearly detailed? Are data analysis approaches (including triangulation methods) explained? Are data collection methods (including sampling strategy) detailed?

	<i>Note: The evaluator should explain if he/she relied on primary data – their collection or secondary data (the sources from the project documents) whether he/she did surveys, case studies, interviews, focus groups, group discussions, employed geospatial data collection methods etc.</i>
2.7	<p>Is the data collection and analysis adequate and sufficiently diverse for the scope of the evaluation?</p> <ul style="list-style-type: none"> ▪ Are a comprehensive set of data sources included where appropriate? ▪ Is there a clear presentation of data analysis and references to data sources within the report? ▪ Did the evaluator explain how project monitoring data was used and verified?
2.8	<p>Are limitations in implementation of the evaluation clearly explained as well as mitigation measures?</p> <ul style="list-style-type: none"> ▪ Are constraints in access to data (including interviewees and quantitative data sources) explained? ▪ Are explanations provided as to how constraints in access to data were addressed?

6.10.3 Cross-cutting issues

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

3.1	<p>If requested in the ToR, does the evaluation adequately include and analyse cross cutting issues such as:</p> <ul style="list-style-type: none"> ▪ Leave no one behind (LNOB) https://erc.undp.org/methods-center/methods/assessing-crossing-cutting-themes/assessing-leaving-no-one-behind ▪ Social and Environmental standards (this is compulsory for GEF evaluations) https://erc.undp.org/methods-center/methods/assessing-crossing-cutting-themes/ses
3.2	<p>Are gender equality and empowerment of women (GEEW) integrated in the evaluation scope of analysis and evaluation criteria and questions designed in a way that ensures GEEW related data will be collected?</p> <ul style="list-style-type: none"> ▪ Does the evaluation include an objective specifically addressing gender equality and/or human rights issues and/or was gender mainstreamed in other objectives? ▪ Was a stand-alone evaluation criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? ▪ Were one or several dedicated gender equality and empowerment of women evaluation questions integrated into the evaluation? <p><i>Section 6.10.3, page 11: https://erc.undp.org/pdf/UNDP_Evaluation_Guidelines.pdf</i> UN SWAP Eval Performance Indicator Scorecard: http://www.unevaluation.org/document/download/2149</p>
3.3	<p>Were gender-responsive methodologies, methods and tools, and data analysis techniques selected?</p> <ul style="list-style-type: none"> ▪ Does the evaluation specify how gender issues are addressed in the methodology, including how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? ▪ Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating gender equality and empowerment of women considerations? ▪ Were diverse range of data sources and processes employed (i.e., triangulation, validation) to guarantee inclusion, accuracy and credibility? ▪ Do the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? <p><i>Section 6.10.3, page 11: https://erc.undp.org/pdf/UNDP_Evaluation_Guidelines.pdf</i> UN SWAP Eval Performance Indicator Scorecard: http://www.unevaluation.org/document/download/2149</p>
3.4	<p>Do the evaluation findings, conclusions and recommendation reflect a gender analysis?</p> <ul style="list-style-type: none"> ▪ Does the evaluation have a background section that includes analysis of specific social groups affected and/ or spelling out the relevant instruments or policies related to gender equality and human rights? ▪ Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/ or disaggregates quantitative data by sex, where applicable?

- Are unanticipated effects of the intervention on gender equality and human rights are described?
- Does the evaluation report provide specific recommendations addressing issues of gender equality and empowerment of women, and priorities for action to improve gender equality and empowerment of women or the intervention or future initiatives in this area?

Section 6.10.3, page 11: https://erc.undp.org/pdf/UNDP_Evaluation_Guidelines.pdf
 UN SWAP Eval Performance Indicator Scorecard:
<http://www.unevaluation.org/document/download/2149>

3.5 Does the evaluation include considerations of disability inclusion in the intervention?

- Do the evaluation questions cover different aspects of disability inclusion?
- Do evaluation findings and analysis provide data and evidence on disability inclusion?
- Do the conclusions and/ or recommendations of the evaluation reflect their findings on disability inclusion?

Note, this is a new requirement from 2019. Only review this IF the TOR requests inclusion or the report includes an analysis of disability inclusion issues. If there is no request in the TOR mark as N/A

Section 6.10.3, page 11: https://erc.undp.org/pdf/UNDP_Evaluation_Guidelines.pdf
 The UN disability inclusion strategy requires the inclusion of an analysis of the inclusion of disability in evaluations. The strategy and guidance can be found here <https://www.un.org/en/content/disabilitystrategy/>

6.10.4 Findings, conclusions and recommendations

Does the report clearly and concisely outline and support its findings, conclusions and recommendations?	
FINDINGS AND CONCLUSIONS	
4.1	<p>Does the evaluation report contain a concise and logically articulated set of findings? (<i>Section 4, 4.4.5, p.30; Annex 4, Evaluation template and quality standards, pp.59-60; Section 6, 6.10.4, p.12</i>)</p> <ul style="list-style-type: none"> ▪ Do the findings address the evaluation criteria and evaluation questions? ▪ Are the findings detailed and supported by evidence?
4.2	<p>Does the Evaluation assess the RELEVANCE of the intervention, if required by the TOR? (<i>As a separate Relevance section or as a discussion within the findings and conclusions</i>)</p> <p><i>Including, but not limited to:</i></p> <ul style="list-style-type: none"> ▪ Government priorities and strategies ▪ UNDP Country Programme Document outcomes and outputs
4.3	<p>Does the evaluation assess the COHERENCE of the intervention, if required by the TOR? (<i>As a separate Coherence section or as a discussion within the findings and conclusions</i>)</p> <p><i>Including, but not limited to:</i></p> <ul style="list-style-type: none"> ▪ <i>Internal coherence: the synergies and interlinkages between the intervention and other interventions carried out by UNDP.</i> ▪ <i>External coherence: the consistency of the intervention with other actors' interventions in the same context. (Complementarity, harmonization and co-ordination with others, and the extent to which the intervention is adding value while avoiding duplication of effort).</i>
4.4	<p>Does the Evaluation assess the EFFECTIVENESS of the intervention, if required by the TOR? (<i>As a separate Effectiveness section or as a discussion within the findings and conclusions</i>)</p>
4.5	<p>Does the Evaluation clearly assess the EFFICIENCY of the intervention, if required by the TOR? (<i>As a separate Efficiency section or as a discussion within the findings and conclusions</i>)</p> <p><i>Including but not limited to:</i></p> <ul style="list-style-type: none"> ▪ How well did the intervention use and allocate its human and financial resources? ▪ Are variances between planned and actual expenditures across results/outputs assessed and explained?
4.6	<p>Does the Evaluation clearly assess SUSTAINABILITY of the intervention, if required by the TOR? (<i>As a separate Sustainability section or as a discussion within the findings and conclusions</i>)</p>
4.7	<p>Does the evaluation include an assessment of the results framework, its reporting and overall quality? (<i>Section 6, p.10</i>)</p> <ul style="list-style-type: none"> ▪ Is there an assessment of the quality of results framework, its indicators and level of disaggregation by sex and other vulnerable groups? ▪ Are all indicators in the Results Framework assessed and analysed by the evaluator with final achievements noted?
4.8	<p>Does the evaluation include an assessment of the monitoring system of the intervention and its overall quality? (<i>Section 6, p.10</i>)</p>
4.9	<ul style="list-style-type: none"> ▪ Is there a discussion and analysis of risk management and the quality of adaption/mitigation measures in the evaluation report including Social and Environmental Safeguards?

4.10	<p>Does the evaluation report contain a concise and logically articulated set of conclusions which are standalone in nature? (Section 4, 4.4.5, p.30; Annex 4, UNDP evaluation report template and quality standards, p.60; Section 6, 6.10.4, p.12)</p> <ul style="list-style-type: none"> ▪ Do conclusions build on the findings? ▪ Are they clear, concise and well substantiated? ▪ Do they address the objectives of the evaluation?
------	---

RECOMMENDATIONS

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

6.10.5 Management Response

MANAGEMENT RESPONSE

Is the management response clear, realistic and actionable?

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

6.10.6 General findings and lessons learned.

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

GENERAL FINDINGS AND LESSONS LEARNED

To be completed in narrative format and not graded. This area is for the reviewer to give further thoughts and considerations of the report.

6.1	<p>Lessons for Evaluations and Evaluators: Overall thoughts on the quality of the evaluation report</p>
-----	--

6.2

Were there any innovative aspects of the evaluation report that you can highlight? Yes/No

- If it is a yes, please explain in the comments.

6.10.7 Validation of the ratings given by GEF terminal evaluations

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

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

GEF Evaluation Ratings Validation Table

		UNDP IEO quality assessment rating		GEF terminal evaluation rating		Comments and/or justification for rating/ score adjustment	Suggestions for improvement
		Rating	Score	Rating	Score		
Assessment of outcomes							
Project focus	Indicate what the terminal evaluation has rated for project effectiveness, efficiency and relevance, and based on the available documentation, indicate and justify your rating. Provide your rating also in cases where the terminal evaluation has not included one.						
1	Effectiveness						
2	Efficiency						
3	Relevance						
4	Overall project outcome						
Sustainability							
Project focus	Indicate what the terminal evaluation has rated for sustainability and based on the available documentation indicate and justify your						

	rating. Provide your rating also in cases where the terminal evaluation has not included one.						
5	Financial sustainability						
6	Socio-political sustainability						
7	Institutional framework and governance sustainability						
8	Environmental sustainability						
9	Overall likelihood of sustainability						

Monitoring and evaluation

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

Implementation and execution

Project focus	Indicate what the terminal evaluation has rated for the performance of UNDP as the project implementing agency and based on the available documentation indicate and justify your rating. Provide your rating also in cases where the terminal evaluation has not included one.						
13	Quality of UNDP implementation/ oversight						
14	Quality of Implementing partner execution						
15	Overall quality of implementation and execution						

Overall project performance

Project focus	Does the terminal evaluation include a summary assessment and overall rating of the project results? Indicate the terminal evaluation rating and then indicate whether, based on the available documentation, you think a different rating of overall project results would be more appropriate.		
----------------------	---	--	--

16	Provide justification for any agreement or adjustment to ratings.							
----	---	--	--	--	--	--	--	--

6.10.8 Summary quality assessment result

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

QUALITY ASSESSMENT CRITERIA	Rating						Weighted score
	HS	S	M S	M U	U	H U	
1. Evaluation structure and design							
Do the terms of reference appropriately and clearly outline the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?							
2. Evaluation report and methodology							
Does the evaluation report appropriately and clearly address the purpose, objectives, criteria, methodology and key questions for the evaluation following the UNDP Evaluation guidelines?							
3. Cross-cutting and gender issues							
Does the evaluation report address LNOB and other cross cutting issues where relevant?							
4. Evaluation results, findings, conclusions and recommendations							
Does the report clearly and concisely outline and support its findings, conclusion and recommendations?							

6.11 UN-SWAP evaluation performance indicator and assessment

The United Nations System-wide Policy on Gender Equality and the Empowerment of Women (GEEW) was endorsed by the Chief Executive's Board for Coordination in October 2006 as a means of furthering the goal of gender equality and women's empowerment in the policies and programmes of the United Nations system. In 2012, the United Nations agreed on the System-wide Action Plan (UN-SWAP) on Gender Equality and the Empowerment of Women to implement the aforementioned gender policy. The UN-SWAP on Gender Equality and the Empowerment of Women assigned common performance standards for the gender-related work of all United Nations entities, including evaluation and includes an evaluation performance indicator (EPI). In 2020, the UN-SWAP was updated. The UN-SWAP EPI reporting follows the [UN-SWAP EPI Technical Notes](#) published by the United Nations Evaluation Group (UNEG).⁶

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

6.11.1 What is the UN-SWAP evaluation performance indicator?

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

6.11.2 The evaluation performance indicator criteria and scorecard

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

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

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

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

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

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

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

- **0 = Not at all integrated.** Applies when none of the elements under a criterion are met.
- **1 = Partially integrated.** Applies when some minimal elements have been met but further progress is needed, and remedial action is required to meet the standard.

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

- **2 = Satisfactorily integrated.** Applies when a satisfactory level has been reached and many of the elements have been met but improvement could still be made.
- **3 = Fully integrated.** Applies when all the elements under a criterion have been met, used and fully integrated in the evaluation and no remedial action is required.

The annex to the [UN-SWAP EPI technical note](#)⁷ sets out guiding questions for the assessment against each criterion. After reviewing the individual evaluation report for each criterion, a score is assigned to the report as follows:

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

6.11.3 The assessment process

The UNDP IEO is the focal point for the EPI. Before 2020, the IEO contracts an external expert to conduct the assessment of a set of evaluations, including all independent evaluations and a sample of decentralized evaluations that were finalized in the period being reported (January-December of each year).

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

Box 1: Sample evaluations that have met EPI requirements.

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

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

6.11.4 Reporting

The IEO prepares a final synthesis report, which is uploaded to the UN-SWAP on Gender Equality and the Empowerment of Women reporting portal. UN-Women analyses all UN-SWAP performance indicators, including for evaluation, and an aggregated report is presented every year through the

⁷ Access at: <http://www.unevaluation.org/document/detail/1452>



report of the Secretary-General to the Economic and Social Council on mainstreaming a gender perspective into all policies and programmes in the United Nations system.⁸

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



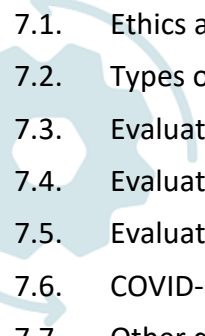
SECTION 7

FREQUENTLY ASKED QUESTIONS (FAQs)






CONTENTS



7.1.	Ethics and conflict of interest	1
7.2.	Types of evaluation	2
7.3.	Evaluation plans	5
7.4.	Evaluation implementation	7
7.5.	Evaluation quality assessment	13
7.6.	COVID-19	14
7.7.	Other questions and issues	17



7. FREQUENTLY ASKED QUESTIONS

Section 7: Details some of the more frequently asked questions during the 2019 regional training and webinars throughout 2020. Links to webinar recordings are in the annexes.

7.1. Ethics and conflict of interest

7.1.1. What is the latest UNDP and UNEG ethical guidance for evaluation?

In 2020, the United Nations Evaluation Group (UNEG) updated the [UNEG Ethical Guidelines for Evaluation](#). This document aims to support leaders of United Nations entities and governing bodies, as well as those organizing and conducting evaluations for the United Nations, to ensure that an ethical lens informs day-to-day evaluation practice. This document provides:

- Four ethical principles for evaluation;
- Tailored guidelines for entity leaders and governing bodies, evaluation organizers and evaluation practitioners;
- A [Pledge of Commitment to Ethical Conduct in Evaluation](#) that all those involved in evaluations are required to sign.

7.1.2. Can we engage a former staff member/ project officer to take part in the evaluation of our country programme?

No, according to the Evaluation Guidelines, *“Evaluators should not have worked or contributed to the project/ programme, outcome or UNDAF/UNSDCF under evaluation at any time in any way”*, therefore in this case, it would not be advisable to engage a national consultant who used to work for the country office.

7.1.3. Can we engage the evaluator of a project/ outcome/ country programme to support the development of the subsequent project/ outcome/ country programme?

No, Section 4.3.5 of the Evaluation Guidelines is clear on the issue of conflict of interest due to potential future involvement or employment: *“Programme units must ensure that the evaluators will not be rendering any service (related or unrelated to the subject of the evaluation) to the programme unit of the project or outcome being evaluated in the immediate future. Evaluators should not subsequently be engaged in the implementation of a programme or project that was the subject of their evaluation. Equally, evaluators should not be engaged as designers of next phases of projects that they have evaluated.”*

7.1.4. Some colleagues add the following statement to evaluation terms of reference (TOR) under the implementation arrangements “It is expected that at least one senior member of the project will accompany the evaluator during the meetings in order to facilitate and provide clarifications where necessary and UNDP norms and standards for evaluations will be applied”.

Does this conflict with the evaluation ethics and norms?

UNDP staff should not join evaluators for stakeholder meetings, even when they have the best intentions. This is clearly stated Section 4.4.4 of the Evaluation Guidelines. Equally the presence of UNDP staff goes against the UNEG [Norms and Standards for Evaluation \(Standard 3.2\)](#), which ensures the confidentiality and anonymity of stakeholders being interviewed.

This needs to be clearly expressed to country offices, otherwise the independence and credibility of evaluations is under threat.

7.2. Types of evaluation

UNSDCF evaluations

7.2.1. Should UNSDCF evaluations be included in UNDP evaluation plans, when they are no longer the responsibility of UNDP country offices?

As of 2022, UNSDCF evaluations do not need to be included in the evaluation plans. Final evaluations do not need to be uploaded to the Evaluation Resource Center (ERC). UNDP IEO does not quality assess these evaluations.

Decentralized Country Programme Evaluations (D-CPEs)

7.2.2. Are CPD evaluations mandatory?

New CPDs must be accompanied by an evaluation of the previous country programme which can come in the form of a D-CPE commissioned by the country office, or an Independent Country Programme Evaluation (ICPE) led and implemented by the Independent Evaluation Office (IEO).

A midterm evaluation of the country programme will enable a country office to make course corrections during the implementation of their programme.

Country offices may also consider an outcome evaluation instead of a midterm review. Outcome evaluations often produce more interesting results, because they focus on desired changes. The findings of an outcome evaluation can then be used for planning the next country programme.

If you conduct a midterm review of your country programme at this time, it is essential to reflect changes due to COVID-19, especially if any reprogramming has been done or delays have occurred

in the projects. This is all going to have major impacts on the achievements of the country programme.

7.2.3. How can we find a quality assured international consultant to serve as the team lead for our country programme evaluation, working with a team of three national consultants sourced from an institutional service provider?

One way to look for evaluators is to check the ERC for evaluations that have received satisfactory or highly satisfactory quality assessment scores, particularly on D-CPE or outcome evaluations. The IEO has prepared a step-by-step “how to” note on using the ERC for this purpose.¹

You may also want to check with the Bureau of Policy and Programme Support (BPPS) to access the Express Roster of evaluators, and ask the regional evaluation focal point whether there is a regional roster.

Global Environment Facility (GEF) evaluations

7.2.4. Why are GEF evaluations so dominant in evaluation plans and implementation?

GEF terminal evaluations are mandatory for all mid-size and large GEF projects, and midterm reviews are mandatory for all large projects. All GEF terminal and midterm evaluations have clear budget allocations

7.2.5. What are we supposed to do with terminal/final evaluation recommendations, which are often backward looking, when the project is completed and closed?

You will always have an issue of the timeliness of recommendations, particularly for closing projects. This is especially true for GEF terminal evaluations. However, you will see that a good number of recommendations are also repetitive. Some issues that come up again and again can be very relevant for the design of new programmes. Therefore, though you may think these are not directly related to your work or the completed intervention, the value of that knowledge for designing new projects should not be underestimated.

7.2.6. Are GEF evaluations excluded from the UNDP quality assessment process?

GEF midterm reviews are not quality assessed by IEO. However, we do quality assess GEF terminal evaluations through the same quality assessment process, plus an additional validation of the scoring of the GEF evaluation ratings. Therefore, we provide extra time for the GEF quality assessment, where the external reviewers also review the project documents and other relevant

¹ <https://erc.undp.org/resources/docs/guidance/ERC-Guide-finding-good-evaluator.pdf>

reports (Programme Implementation Reports and midterm reviews). Quality reviewers undertake a validation of whether they agree with the scoring that the evaluator has given, and GEF considers only the revised scores.

There is newly [updated guidance for terminal evaluations](#) available, which is more detailed and outlines roles and responsibilities throughout. It is also aligned with the UNDP Evaluation Guidelines and the GEF evaluation policy. Hopefully, the quality of terminal evaluations will improve with this new detailed guidance.

Outcome and impact evaluations

7.2.7. What can we do to make our reviews and evaluations more useful? How can we move towards more outcome and impact evaluations?

In order to capture higher-level results at outcome and impact level, this needs to be clearly reflected in the TOR. Evaluations should not look at activities and outputs, but should focus more on the links between outputs and the overall outcome and impact goals. Throughout the evaluation management process (inception reports, draft reports) it will be important to remain focused and ensure that those linkages are made.

7.2.8. What are the recommended timing and costs for a complex outcome evaluation that covers several projects and requires an international consultant?

The problem here is that budgets are often set first. Later in the evaluation process, we build our team according to our budget and find out that we do not have sufficient funds for the adequate number of days. It is estimated, depending on the scope of the evaluation, to consider 30 to 50 days for an outcome evaluation. Therefore, for two people with an average daily rate of US\$ 650 per person per day, a standard outcome evaluation would require a budget of around \$50,000.

The ERC gives examples of TORs and budgets for all types of evaluations.

7.2.9. We are planning to carry out three outcome evaluations this year, and also have an ICPE scheduled. Should we still carry out the outcome evaluations or can they be deleted?

There may be a duplication of work and findings if the country office undertakes outcome evaluations at the same time as an ICPE, which will also cover outcomes in some detail. However, this should be discussed with the ICPE lead evaluator, the country office and the regional bureau.

Virtual evaluations

7.2.10. What is a virtual evaluation, and what does it mean in practice?

When we talk about a “virtual” evaluation, we refer to an in-depth desk review coupled with remote interviews.

Evaluators will need to be provided with all available documentation, including quarterly project reports and annual project board meeting notes, not only annual reports and work plans. It is also important that we provide detailed data to the evaluators and explain the rationale behind it.

Remote interviews may be easier with UNDP staff and colleagues than with other stakeholders, due to internet access and connections. In the planning phase of the evaluation, it will be important to agree on which online tools to use, such as Skype, Zoom etc., appropriate to the country context. Of course, interviews can be conducted over the phone as well. Therefore, it is very important that you have updated telephone lists, including mobile phone numbers, for all stakeholders.

It is challenging to conduct a virtual evaluation, and stakeholders need to buy into and participate in the evaluation. There will be challenges such as delays when people are not available, extensions to contracts because interviews are spread across a longer time, and gaps which can affect the usability and the credibility of the evaluation.

Another challenge is the possible differences in evaluation team composition and / or team dynamics between international and national evaluators. The virtual approach may be new for many people. It is very important to clearly specify the different tasks and roles between international and national evaluators in the TOR.

Even though flexibility is important, we must also ensure that our evaluations remain credible. Therefore, evaluation must follow very clear protocols, and record how they are conducted. We have released guidance on how to conduct evaluations remotely.

7.3. Evaluation plans

7.3.1. How should we adjust our evaluation plans during COVID-19?

Given what we currently know about COVID-19, things are not going to go back to normal anytime soon. Therefore, we should be looking at restructuring our evaluation plans for 2021, and reviewing plans for 2022.

Evaluations should not just be shifted from one year to another. We really should be looking at the evaluation plans in their entirety. We should be considering what exactly we want to capture from

our portfolios of programmes and projects to inform our work going forward, whether for ongoing work, or new interventions related to COVID-19. This will take us to new areas regarding how to evaluate COVID-19 interventions, which will be necessary to support our programmes.

7.3.2. Can we institutionalize the annual review of the country programme evaluation plan to make this part of the country office business plan discussion at the beginning of the year?

Yes, we need to get the evaluation plan into country office business plan discussions. In the Evaluation Guidelines we recommend that you review your evaluation plan at the beginning of every year. We also recommend that you undertake a midterm review of your evaluation plan.

We also need to work with regions and see how to get this formalized within the operational processes of the country offices.

7.3.3. Can we combine several planned project evaluations that relate to one portfolio into a single portfolio evaluation?

Yes, you can do that. In such case, it is useful to communicate with any donors and enquire whether the planned project evaluations could be changed to outcome or thematic evaluations. It can be assumed that the majority of donors will be flexible in these challenging times.

GEF still requires individual terminal and midterm evaluations.

7.3.4. Are the new evaluation requirements mandatory for new projects only, or for all ongoing projects?

The new evaluation requirements are mandatory for all new projects starting from January 2019, but we strongly recommend that you also consider ongoing projects with budgets of over \$5,000,000. There is a great opportunity for learning and accountability from these projects and programmes. See [Sections 2 and 3](#) of the Evaluation Guidelines for further details.

In this sense it is important to continuously revisit your evaluation plans and see which type of projects and programmes need to be evaluated, and how. It is also useful to prioritize your evaluations. Instead of several project evaluations you could consider a portfolio, outcome or thematic evaluation, where applicable.

7.3.5. Do we need to include all evaluations in the evaluation plan? Do evaluations which are financed and conducted by donors also need to be included?

Yes, all evaluations where UNDP is involved need to be included in the evaluation plan, even when they are driven by the United Nations country team, resident coordinator or the European

Commission. Only if all evaluations are included can we have a full picture of what is being evaluated. However, we only quality assess those evaluations that are managed by UNDP.

It is mandatory to include evaluations conducted by donors in your evaluation plan, as this provides UNDP and its stakeholders with an understanding of the evaluative coverage of the UNDP programme. Joint evaluations also need to be included, because this helps us to track which joint evaluations have been conducted.

7.3.6. The country office has several planned project evaluations this year, including a number of mandatory GEF evaluations, and we may need to reschedule some. Can we delete those that are below the evaluation guideline thresholds?

We do not recommend deleting evaluations below the threshold. The thresholds were brought in to guarantee that larger projects are evaluated, not to eliminate the need to evaluate smaller ones.

We suggest you review your evaluation plan with your regional monitoring and evaluation (M&E) focal point, looking at the critical importance of each evaluation for management decision-making. From there, you should consider what is mandatory, what could be rescheduled, and what you may be able to combine under a thematic or portfolio evaluation.

7.3.7. Do we have to submit an evaluation plan every year, even when the evaluation portfolio is very dynamic?

The evaluation plan is approved and uploaded at the start of the country programme period. This should be reviewed every year to take account of changes and evaluation needs in the programme unit portfolio of projects. There is no need for a formal submission of a new evaluation plan. Revisions are managed through the ERC.

7.3.8. What do I have to do to change or extend the completion date of an evaluation in the plan?

Make the required date changes in the ERC along with a rationale/ justification for the change. This will be reviewed by the regional evaluation focal point and approved accordingly. There is guidance and a short video on how to upload the evaluation plan to the ERC.

7.3.9. Is it possible to cancel an evaluation, for example due to security reasons?

Yes, evaluations can be cancelled from the evaluation plan through the ERC, but only in exceptional circumstances. Cancellations will need a clear and detailed justification and rationale, which are reviewed and approved by the regional focal points. See section 3.8 of the Evaluation Guidelines for more information on justifications for deleting an evaluation from the plan.

7.4. Evaluation implementation

Evaluation budget

7.4.1. Can we adapt the evaluation budget, for instance to apply a different methodology for which we may need a different consultant or company?

As far as IEO is concerned, the budget can be used however necessary for the evaluation, but there may be some administrative constraints for shifting budgets, for example from travel to consultants' fees or remote data collection.

7.4.2. How can we ensure the availability of funds for an evaluation? What can we do if we do not have sufficient budget for an evaluation?

The budget for an evaluation needs to be defined in the project design phase. It is important to keep the *monitoring* and *evaluation* budgets separate to ensure that the majority of the funds are not spent on monitoring, and enough is left at the end of the project for the evaluation.

If the project budget is high, then it should be able to allocate funds of \$30,000 to \$50,000 for an evaluation.

Often, we do not plan for our evaluations properly. We often plan late and, as a result, we work with the budget allocated for an evaluation, rather than asking early on why we want to do an evaluation (rationale), what we want to cover (scope), and, from there, how many days it will take and how much it will cost.

We have seen some very low evaluation budgets, even below \$5,000. It is very unrealistic to get a decent quality evaluation report with such a low budget. Even if only national consultants are selected, this is not going to provide a decent analysis of the results and project design.

If you do not have enough budget to evaluate a project, you may want to explore pulling several projects together. You may want to think about a portfolio evaluation, thematic or an outcome evaluation.

Another option is to reduce the scope of the evaluation, supporting the evaluator as much as possible in the preparatory phase by providing relevant documentation, so that they can focus. For example, if no theory of change is available, request that the programme team articulate it, ask them to write up some notes on the key elements of the project, so this does not become the task of the evaluator. The new COVID-19 evaluation guidelines also offer some support in this respect.

It is important to mention that all projects with a budget above \$5,000,000 must plan for an evaluation, this is mandatory. See Section 2 for further information.

7.4.3. Many new European Union projects demand an evaluation no matter the project budget. What can we do about that, especially when funds are limited?

You should discuss the evaluation budget with the donor. You can provide an explanation of why it might be beneficial to increase the budget. For example, you could indicate that you do not think that the scope will be covered sufficiently, or that it will influence the quality of the evaluation. If it is not possible for them to increase the budget, then you could enquire in the country office whether it is possible to increase the budget for this evaluation. You may also suggest decreasing the scope of the evaluation and reducing the number of evaluation questions.

Finding an evaluator

7.4.4. Is it possible to use internal staff as evaluators for projects, if they have not worked on the project? Can colleagues from the Resident Coordinator's Office provide evaluative support and staff?

Unfortunately, using a colleague from the Resident Coordinator's Office or country office as an evaluator would not be considered sufficiently independent.

If you do not wish to go through a further recruitment process using the available regional and headquarters rosters, you could continue with a review of the project, rather than an evaluation. Bureau colleagues may be able to give further guidance and support on this.

7.4.5. What are the best ways to identify strong evaluators?

For the credibility of the evaluation, it is very important that you hire a vetted and highly experienced evaluator.

A [guidance note](#) is available on how to use the ERC to identify good evaluators. In the ERC, you can search by country and theme. BPPS has a vetted roster which can be used to identify good evaluators. Regional evaluation focal points can help you access this roster and in some cases regional bureaux also have evaluator rosters. For GEF projects there is a consolidated roster platform, which details evaluator experience in evaluating environmental projects, as well as other topics such as climate change adaptation, climate mitigation, biodiversity, chemicals and waste management, water, and ocean governance. This roster is available to country offices, regional hubs and headquarters.

We have seen cases of evaluators who conducted poor evaluations in the past being repeatedly selected. We encourage you to check the ERC quality assessment scores for evaluators and the level of consistency within the ERC. Also reach out to colleagues who have hired them before. Please, also contact the respective M&E focal point in order to find out what the issues were at the time.

You could also ask your colleagues if they would recommend an evaluator. If you get a name you can then check the ERC, search for the evaluator's name, and see their previous evaluations reports along with the quality assessment scores. A [guidance note](#) is available on how to use the ERC to identify good evaluators. Evaluators should share an example of their previous work, some reports, or anything that they worked on previously, so that you can read some pages and see the quality of the work.

In this respect it is also important to have a separate and adequate budget allocation for evaluations. At the time of budgeting for the evaluation you should broadly define the scope and rationale for the evaluation, which in turn will determine the competencies required of the evaluators.

7.4.6. would it be possible for somebody from the New York office to get involved in the interview process to help us find a good evaluator?

IEO is not able to respond to all requests, given that there are about 350 evaluations a year in UNDP, but we could help for some key and larger evaluations.

Cross-Cutting Issues and Gender

7.4.7. Can we conduct a gender-related evaluation?

Yes, you can decide to have a gender-related evaluation. In fact, gender is one of the most common thematic evaluations, to analyze how gender is mainstreamed across the programme portfolio. Additionally, *all* evaluations should include gender-related evaluation questions, as set out in the Evaluation Guidelines.

7.4.8. How can a project which does not have gender well integrated in its design be evaluated against the gender indicators?

We do not integrate gender as well as we should as an organization. We have not guided evaluators very well on how, or what, to evaluate in relation to gender. However, the revised Guidelines integrate gender for the entire evaluation process and give detailed guidance.

Where you have a project that has not considered gender very well in its design and implementation, the evaluation can look at this very thoroughly. The gender aspects, or lack of them, should be stated in the findings, conclusions and recommendations.

Similarly, you could have a project that has contributed a lot to gender equality and women's empowerment, but with an evaluator who ignores that aspect, resulting in a low quality assurance rating for the gender questions.

The inclusion of gender in evaluations is a consistent issue. Often gender appears in the TOR as a general statement that “gender needs to be considered”, without any concrete gender-related evaluation questions. Sometimes there are specific gender-related evaluation questions, but

without guidance on how they should be addressed. It is surprising that you can have a poverty reduction or livelihood programme, which is about equality and “leaving no one behind”, where gender is not mentioned in the TOR or the final evaluation report.

Here the role of the evaluation manager needs to be emphasized. If the evaluator ignores gender in the inception and/or final draft report, then it is the evaluation manager’s responsibility to get back to the evaluator and request it, assuming that it was specified in the TOR.

7.4.9. What can be done about other cross-cutting issues missing in evaluations, such as human rights, poverty and SDG linkages?

If the analysis of other cross-cutting issues is missing or not sufficiently covered in the final evaluation report, this is an issue of quality assurance. As with the gender issues, the scope of the cross-cutting issues needs to be well defined in the TOR and followed up in the inception and draft final evaluation report processes.

7.4.10. Can we ensure that gender advisors are consulted at some point in the preparation and implementation phases of an evaluation?

Yes, it would be very good practice to engage your gender advisor or focal point, especially at the beginning of the evaluation process.

General

7.4.11. Do we have to send project evaluation TORs to IEO for approval?

No, TORs for project evaluations should be shared with the regional evaluation focal point directly, or through the ERC for technical inputs.

7.4.12. Do we have to use the Development Assistance Committee (DAC) criteria in our evaluations?

We often see errors, or low quality evaluations, when the DAC evaluation criteria are not used. The Evaluation Guidelines strongly recommend using them. It is clear that, when other frameworks are suggested in the TOR or applied by the evaluator, the quality of evaluations declines considerably.

7.4.13. Should the Evaluation Guidelines be shared with the evaluator?

Yes, we strongly recommend that you include a reference and link to the Evaluation Guidelines in the evaluation TOR. We also suggest that you remind the evaluator to look at the guidelines and use the templates provided for the inception and final reports. The UNDP and GEF evaluation guidelines are available in several languages.

It is also essential for the evaluators to make themselves familiar with the quality assessment requirements. Please inform and remind them that their work will be quality assessed against the

criteria stated in the Evaluation Guidelines, including for GEF evaluations. Evaluators should ensure that their evaluations are of the highest standards possible.

Evaluation managers play an important role in ensuring the quality of the evaluation, drafting good TORs, and only approving inception and draft final reports if they meet the quality standards set out in the Evaluation Guidelines. If lower quality reports are approved by the evaluation manager, then the external quality reviewers will score them accordingly.

7.4.14. How can we deal with the fact that end of the year is always a very busy time, when it is often difficult to deal with evaluations on top other work?

We are fully aware that October, November and December are very important months for country offices to focus on delivery. That is why the evaluation planning process is so important, and you may want to consider scheduling your evaluations earlier. Reviewing your evaluation plan in July or August is a good step for getting clarity on this. But if the last quarter of the year is the only option, we recommend that you ensure that everybody involved is aware of the time and agrees to it, and that you have full management support.

Please also note that the evaluation plan considers the date for *completion*, not the start date for the evaluation. If evaluations are all dated October, November and December in your evaluation plan, this does not mean that you have to wait until then to start working on them. Any evaluation needs a six month process, so if the completion date is December, the process and TOR need to start in June at the latest. If your completion date is September, then the process should have started in February or March.

7.4.15. Is the evaluation manager a fixed staff member or will this change project by project?

The evaluation manager cannot be the same person as the manager of the project being evaluated. In the majority of cases, the M&E officer will manage the evaluation. In some cases the project manager of one project may be the evaluation manager of another, for the duration of that particular evaluation.

7.4.16. How can we triangulate data when field visits are not possible?

If field trips are not possible, the only way to triangulate is between desk review, interviews and additional administrative records. Where useful, and possible, contact stakeholders via mobile phone. You may also be able to use alternative data sources, such as big data, geo-spatial data, social media data, etc.

7.5. Evaluation quality assessment

7.5.1. Is there a possibility to link evaluation through ERC with the Project Quality Assurance (PQA) portal?

A feed has been made between the two portals.

7.5.2. How can we support quality assurance work?

The evaluation manager plays an important role in the entire quality assurance process. They will draft the TOR and approve the inception and final reports. If you have a good evaluation process which meets international standards, this will have a positive impact on the quality assurance.

7.5.3. What is the role of M&E focal points to ensure quality, if not acting as the evaluation manager?

From an evaluation perspective it is very important that all colleagues know about basic evaluation requirements. The Evaluation Guidelines are to be used by the entire organization, not only the M&E focal points, and include templates for TORs, inception and final reports, along with relevant links. Try to participate in the evaluation planning process and share these templates, which are the key tools to ensure the quality of the evaluation, with the evaluation managers. You can also share these links with your colleagues, especially for decentralized evaluations.

IEO runs a short training course, “introduction”, which takes about 90 minutes and there is a longer certified training course available, which takes three hours. These courses are not just for the M&E focal points, but for the whole organization.

From experience, we know that a lot of time is spent looking for simple documents such as annual reports, work plans and monitoring data. As M&E focal points you can work with your colleagues to sensitize them on the importance of having a good knowledge management/ filing system.

Another way to improve the quality of evaluations is to share the draft TOR, draft inception report and draft final report with your regional evaluation focal points.

7.5.4. Would it make sense to have a two-step process so that the TORs are quality assured before moving on to the implementation of the evaluation?

It is a great idea to quality assess the TOR prior to an evaluation. At the moment, this is a big weakness in the entire evaluation process. In order to take this further, we need to consider how best to set up the mechanisms for this. This brings us back to the role of the regional evaluation focal points, who can give technical inputs if involved early enough in the process.

7.5.5. What are the major quality concerns in the different phases of an evaluation?

TORs play a big role in the quality of evaluations. If the TOR are not clear and precise, clearly directing the evaluator towards the evaluation rationale, objectives and scope, or do not indicate an initial methodology or cover evaluative questions including gender and other cross-cutting questions, then the evaluation will not be of high quality. In general, good TORs lead to good evaluations, but of course you also need a good evaluator and sufficient budget for the evaluation.

Sometimes, the DAC evaluation criteria are not followed. This results in a loss of structure in presenting the evaluation findings and results. If the DAC criteria are not followed, then the evaluators can lose their way. The DAC criteria provide a very clear framework, and guidance on how to present findings. The Evaluation Guidelines provide very clear instructions, and we have a clear outline on how the final report should be structured.

Some other challenges relate to the evaluation reports. The linkages between findings, conclusions and recommendations are often weak. In particular, the quality of recommendations is very important, since they ultimately link to the usability of the evaluation. We often find recommendations are too generic and superficial, for example “you need to improve your monitoring system,” without providing further details. Recommendations need to help guide the country office and the project / programme on how better implement the programme. Therefore, the recommendations need to be specific, realistic and actionable.

Generally, we must ensure that we take the entire evaluation process seriously. It is not simply about the preparation of an evaluation report. We also have to make sure that the evaluation findings and recommendations are made available to national governments and other donors, to ensure that they are used.

There is also a financial component. Evaluations cost money and, therefore, every evaluation report which is not of decent quality is a financial loss to everybody.

7.6. COVID-19

7.6.1. What are the key points to consider in adapting our evaluations to the COVID-19 context?

In a good year, we all have challenges implementing and maintaining the quality of our evaluations and keeping to our evaluation plan. This is by no means a normal year, and we have issued several guidance documents to help you to implement your evaluations. But there will be additional and new challenges emerging.

Safety first. As a first step, we need to ensure that we are not putting ourselves, our colleagues, consultants, evaluators or partners in any danger.

Our second step is to rethink our evaluation plans.

If international and national colleagues and consultants are not able to travel, we must rethink, and reconfigure our evaluation teams. We are using Skype and Zoom and telephones to interview partners, United Nations colleagues and UNDP staff, alongside surveys and questionnaires and extended desk reviews. Where it is safe, we are also using more national consultants than before, and we are looking at various means to collect data in the field or use secondary data.

There are a few things that we need to do:

- a. Look at our evaluation plans for 2021 (and possibly 2022) and do evaluability assessments.
- b. Enquire how we will implement these evaluations and, equally, think how we will evaluate COVID-19 interventions.
- c. Reconsider combining evaluations where possible and reconfigure evaluation teams.
- d. Use our international evaluators remotely, and clearly identify tasks for national evaluators.

GEF evaluations can be delayed with the approval of your regional technical advisor but different GEF evaluations cannot be combined, and they certainly cannot be cancelled.

There are a few other important questions to be consider: Can people travel to the country? Can we travel within the country? Are government partners and other stakeholders available for meetings (in person or online)?

7.6.2. Are there criteria for re-prioritizing evaluation plans due to COVID-19?

In our view, COVID-19 offers an opportunity to rethink the number of evaluations planned and reflect on how we can make our evaluation plans more of a managerial tool. We need to undertake a criticality assessment of our evaluation plans, asking:

- Is this a **mandatory** evaluation, e.g. for GEF or other donor-funded projects? If yes, then all you can do is reschedule it for a later date, but even that must be decided in collaboration with the funding agency.
- Is the timing **realistic** or **helpful** given the current situation? For example, are the midterm and final evaluations planned close together? Is the evaluation going to give us results that are useful for our portfolios, programmes and projects?
- Is the evaluation **critical** to management and decision-making within the country office at this time?

7.6.3. How can I combine evaluations?

Some planned project evaluations might fall under the same outcome, and could possibly be combined. Interventions with similar themes could also be covered under a thematic evaluation.

You could capture synergies at different levels and only develop one TOR. You could also combine planned midterm and final evaluations in similar areas, which would increase the learning between different but similar projects. The recommendations would need to be specific for each project. This would be a type of joint evaluation.

Where project implementation has slowed due to COVID-19 and may be extended, a midterm evaluation may still be beneficial to support the direction of the project. This should be done with the minimum evaluation standards as set out in the COVID-19 evaluation notes.

7.6.4. Should we be documenting COVID-related changes to the country programme and projects?

It is very important that any possible changes in respect to the country programme due to COVID-19 are captured and documented. Otherwise, 18 months down the line, we will find ourselves in a situation with nothing to base our evaluations on.

7.6.5. We were planning a combined evaluation, but due to COVID-19 it is possible that some projects will get a no-cost extension and the end dates of the projects might be different. How do we then conduct an evaluation?

Each project is different, and COVID-19 has posed different challenges at different stages in their implementation. If they all had an extension for 12 months, then you would likely want to extend that evaluation as well. If not, you may still have sufficient flexibility to combine the evaluations. If not, then you will have to revert to individual evaluations, and you can consult us directly.

7.6.6. Some programming modalities have been adapted due to COVID-19. What does that mean for evaluating projects?

It is a particular challenge for an evaluation if the theory of change is only partially applicable. Changes to their rationale, and relevant decisions, need to be captured for the evaluation, to understand the assumptions of the changed portfolio.

Depending on the situation, different types of evaluations could be useful, for example separate project evaluations, portfolio evaluations, or including a separate section in the D-CPE and regional evaluations. On the global level we could expect COVID-19 evaluations at a UNDP or United Nations system-wide level, or together with other donors.

7.6.7. Will the conferences of national evaluation capacities (NEC) initiative continue?

Yes, we will continue with the NEC conferences and the next ones will be conducted in 2021. This is an important area for IEO. We are also embarking in a partnership with the World Bank Independent Evaluation Group on a broader coalition called the “[Global Evaluation Initiative](#)” to support national evaluation systems with additional and substantial resources, combining the efforts of the World

Bank with a number of other donors through a trust fund. Expect to hear more about that initiative soon.

We will also explore how to continue to work better with national think tanks, academic institutions and consultants.

7.7. Other questions and issues

7.7.1. Why are evaluations for pilot projects mandatory?

If your pilot project is aiming to go to scale, the proposal needs to be based on evidence-based results and findings. The best way to get those results is probably an evaluation.

7.7.2. Do the Evaluation Guidelines cover aspects on how to strengthen national evaluation capacity and SDG evaluations?

The Evaluation Guidelines do not cover how to build national evaluation capacities. But there is reference to a [national diagnostic tool](#), available on the IEO website, which can be used to assess evaluation capacities at the country office level. Evaluators or partners can also assess their own evaluation capacities, and you can build capacity development programming based on this assessment.

7.7.3. What is the best way forward for quality assessment of the reports of donor-led evaluations, when their own evaluation guidelines or process do not necessarily follow our expectations?

Engage with donors early, if possible from the outset of the programme, to discuss what the evaluation should look like and come to an agreement. Where an agreement is not forthcoming, you can outline the evaluation process and approach. This could also be mentioned in the TOR.

Additionally, you can ask the donor to speak to us, because we have relationships with bilateral evaluation offices as part of our engagement with other networks, and we share the same evaluation principles.

7.7.4. Where can we find and share evaluation reports to strengthen information sharing across countries, contexts and evaluation practices, for example on least-developed countries, post-conflict countries or countries with humanitarian crises?

If you want to look at other evaluation reports, please visit the ERC. You can search evaluation reports by themes and countries. You can also join our series of webinars and trainings.

7.7.5. Can we use monitoring data to validate information?

Yes, monitoring data should be used as much as possible. Therefore, it is extremely important that the evaluators get all project-related information and data as early as possible in the evaluation process.

7.7.6. Investments in monitoring are important, but often there is no monitoring data available. Is the evaluation a good time to invest in collecting this kind of data?


Certainly, *monitoring* and *evaluation* are closely related, and good monitoring can lead to a good evaluation. However, it is important to distinguish between monitoring and evaluation responsibilities. The type of data collected in an evaluation will depend a lot on your evaluation questions. If necessary, evaluation recommendations should explain clearly how the respective monitoring system of the project and / or office could be improved.





UNDP EVALUATION GUIDELINES

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

Web: <http://www.undp.org/evaluation>

 /UNDP_Evaluation

 /ieoundp

 /evaluationoffice