UNITED NATIONS
OFFICE OF COUNTER-TERRORISM

EVALUATION HANDBOOK

May 2023

This document was prepared under the guidance and supervision of the Evaluation and Compliance Unit (ECU) in the Office of the Under-Secretary-General (OUSG) for Counter-Terrorism (UNOCT), by Heather Ann Sutherland. It benefitted significantly from the input of colleagues representing the following Sections and Programmes in UNOCT: the Human Rights and Gender Section (HRGS), the Programme Management Unit (PMU), the International Hub on Behavioural Insights to Counter Terrorism, the Global Programme on Preventing and Countering Violent Extremism (PCVE) and the Global Programme on Prosecution, Rehabilitation and Reintegration (PRR).

The Evaluation and Compliance Unit of UNOCT may be contacted at the following:

United Nations Office of Counter-Terrorism
Email: OCT-ECU@un.org

FOREWORD

I am pleased to present the United Nations Office of Counter-Terrorism (UNOCT) Evaluation Handbook which has been developed to operationalize the Evaluation Policy of the Office launched in March 2021. This Handbook contains practical guidance for the conduct, management, reporting, and follow up of evaluations of policies and programmes in UNOCT. It demonstrates the Office's strong commitment to building a results culture and ensuring accountability and represents our dedication to continuous learning and organizational development.

The Handbook takes into consideration existing best practices as well as recommendations from relevant studies and evaluations. It responds directly to the meta-synthesis of evaluation results from UN entities participating in the implementation of the Global Counter-Terrorism Strategy, which called for strengthening evaluation knowledge and capacities among Compact entities and for developing a common M&E framework to identify and measure results and impact of the work conducted by Compact entities under the Strategy. Similarly, the 2020 external evaluation of the UNCCT 5-Year Programme recommended results-based monitoring of programmes consistent with UN regulations, norms, and standards for monitoring and evaluation.

The Handbook therefore serves as guidance material to develop knowledge and capacities for conducting impactful evaluations in UNOCT and when jointly undertaking evaluation processes with other UN entities. While the primary audience for the Handbook is UNOCT staff who manage programme evaluations, it also contains valuable resources of interest to a wider range of users, including evaluators who carry out evaluations commissioned by UNOCT, and other stakeholders involved in evaluation processes.

This Handbook is a culmination of collective efforts across the Office under the overall coordination of the Evaluation and Compliance Unit in the Office of the Under-Secretary-General for Counter-Terrorism. I commend all of those involved in its preparation. I encourage all UNOCT staff to read this Handbook and put the guidelines and tools into practice so that we improve how we plan, manage, and use evaluations to further strengthen our results-based implementation.

Vladimir Voronkov
Under-Secretary-General
UNOCT

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

BI Hub	International Hub on Behavioural Insights to Counter Terrorism
CE	Centralized Evaluation
CEDAW	Convention on the Elimination of All Forms of Discrimination against Women
CT/PCVE	Counter Terrorism/Preventing and Countering Violent Extremism
CTTS	Countering Terrorist Travel Section
DI	Disability Inclusion
ECU	Evaluation and Compliance Unit
ERG	Evaluation Reference Group
EMMT	Evaluation Management Tracking Tool
EvalNet	Network on Development Evaluation (of OECD DAC)
GCTS	Global Counter-Terrorism Strategy
GEEW/G	Gender Equality and the Empowerment of Women and Girls
GM	Gender Marker
HRBA	Human Rights Based Approach
HRGE	Human Rights and Gender Equality
HRGS	Human Rights and Gender Section
IE	Internal Evaluation
IP	Implementing Partner
IPE	Independent Programme/Project Evaluation
KMCS	Knowledge Management and Coordination Section
LNOB	Leaving No One Behind
M&E	Monitoring and Evaluation
MSC	Most Significant Change
OECD DAC	Organization for Economic Co-operation and Development Assistance Committee
OHCHR	Office of the High Commissioner for Human Rights
OPDs	Organizations of Persons with Disabilities
OUSG	Office of the Under-Secretary-General
PD	Programme/Project Document
ProDoc	Programme/Project Documentation
PKMCB	Policy, Knowledge Management and Coordination Branch
PMU	Programme Management Unit
PPBME	Programme Planning, Budgeting, Monitoring and Evaluation Regulations and Rules
PRB	Programme Review Board
PWD	People with Disabilities
RBM	Results-Based Management
RCT	Randomized Controlled Trials
SDGs	Sustainable Development Goals
SOP	Standard Operating Procedure
SPIB	Special Projects and Innovation Branch
SPRF	Strategic Plan and Results Framework
SPPSS	Strategic Planning and Programme Support Section
TOC	Theory of Change
ToR	Terms of Reference
UNCCT	United Nations Counter-Terrorism Centre
UNDIS	United Nations Disability Inclusion Strategy
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNGA	United Nations General Assembly
UNGCTS	United Nations Global Counter-Terrorism Strategy
UNODC	United Nations Office on Drugs and Crime
UNSDCF	United Nations Sustainable Development Cooperation Framework
UNSWAP	United Nations System-wide Action Plan
USG	Under-Secretary-General

INTRODUCTION

This Evaluation Handbook aligns with the United Nations Secretary General's commitment to strengthening the evaluation capacity of the UN Secretariat to better inform programme planning and design and enhance reporting on programme performance. Evaluation is also fundamental to the United Nations Office of Counter-Terrorism (UNOCT) efforts to ensure the Office is accountable for the funds received and transparent about progress in achieving its intended results. Equally important is that evaluation enables decisionmakers to learn from successes and challenges, better enabling the achievement of programmes that are transformative, relevant, timely, gender sensitive, human rights-responsive, inclusive, and impactful.

The UNOCT Evaluation Policy provides the overall framework for the planning and undertaking of evaluation. As the custodian of the evaluation function, the Evaluation Compliance Unit (ECU) has responsibility for implementing this policy with accompanying guidelines and supporting continuous improvement of the quality and credibility of evaluations in line with the United Nations Evaluation Group (UNEG) Norms and Standards and with international good practice.

PURPOSE AND AUDIENCE

The purpose of this Evaluation Handbook is to provide concrete guidance on how the Evaluation Policy is to be carried out across the Office. The intended audience is all UNOCT staff, and in particular programme and project managers who have responsibility for managing evaluations, as well as their programming units and senior management. It is also an important resource for independent evaluators who need to understand not only the guiding principles, standards, and processes for evaluations within the UNOCT context, but also where the intervention under review sits within the structure of the organization and the UN system.

The UNOCT Evaluation Handbook is not a manual for how to design and conduct evaluations. It is expected that contracted evaluators will have this expertise, and there are plenty of resources that provide detailed information on basic methods and on more complex and emerging methodological approaches - several of these are referenced at the end of this document. Rather, this handbook aims to set out the procedures and expectations for how UNOCT evaluations are to be planned, managed, and followed up on. General guidance on good evaluation practice is provided throughout the handbook and the last chapter offers fundamentals for those who are less familiar with evaluation processes but who are required to supervise evaluations and ensure the quality of deliverables.

Throughout the handbook, 'intervention' is used to refer to any evaluand (subject of evaluation) which includes programmes, projects, policies, strategies, activities, and thematic areas. In addition, the term 'evaluation manager' is used to refer to the person responsible for organizing and leading evaluations. Although this role is usually assumed by the programme or project manager or the ECU, in some cases it will be the Evaluation Officer assigned to the respective programme unit or section.

STRUCTURE AND CONTENTS

This handbook contains eight chapters. It is designed so that the chapters can be read in sequence or as stand-alone sections. It is intended to be updated as needed to respond to changes in UNOCT policy and approaches as well as to changing needs for evaluation guidance.

Chapter 1 Evaluation and Why It Matters provides a basic introduction to evaluation and discusses how it differs from other types of assessment and how it links to results-based management. The chapter also explains why 'measuring impact' and 'impact evaluations' do not always mean the same thing, and what UNOCT's International Hub on Behavioural Insights to Counter Terrorism (BI Hub) is doing to increase the rigor of programme design and evaluation processes.

Chapter 2 Evaluation in the UN System gives an overview of the standards, expectations, and ethical principles that all evaluations carried out by UN entities are required to follow, including the requirements for integrating human rights, gender equality, leave no one behind, and disability inclusion. It also situates this work within the Sustainable Development Goals (SDGs) and the 'Delivering as One UN' agenda and provides guidance on conducting joint evaluations. This is an important chapter for all audiences. It is particularly critical for evaluators who have not assessed interventions within the UN system as the entire framework for evaluation processes is guided by criteria that are not used by all national and international public sector organizations. The chapter ends with links to the key resources for conducting such evaluations.

Chapter 3 UNOCT Evaluation Function lays out the guidelines for how the Office's evaluations are structured, planned, and implemented. It introduces the Evaluation Policy, the Evaluation Compliance Unit, and the different categories and types of evaluations within UNOCT, as well as roles and responsibilities for each. This chapter is of importance to Programme and Project Managers as it also explains the preliminary evaluation planning and budgeting process. The final section is useful for all audiences as it discusses special considerations for evaluating UNOCT-supported interventions, particularly counter terrorism/preventing and countering violent extremism (CT/PCVE) programming.

Chapter 4 Planning, Costing and Preparing for Evaluations goes into more detail on the steps to be followed before the evaluation begins. It is primarily directed to evaluation managers as it covers the evaluation proposal, scheduling, further guidance on budgeting, how to develop the terms of reference (ToR) for the evaluation, and recruitment of the evaluation team. The Planning Phase and the Preparation Phase are the initial phases of any evaluation process.

Chapter 5 Implementing and Managing Evaluations covers the steps involved in the third and main phase, and as such, it is important for all stakeholders who participate in evaluation processes, including the evaluation team. The Implementation Phase encompasses the initial briefing of the evaluators and others most closely involved in the evaluation process, the inception/design stage of the evaluation, the data collection and analysis stage, and the development of the draft and final evaluation reports, as well as the evaluation brief. This chapter includes suggested meeting agendas, resources, and report templates to be shared with evaluators, ways of engaging stakeholders in the evaluation, and responsibilities and approval processes.

Chapter 6 Using Evaluation Results discusses the final phase; the Follow-up Phase that needs to take place once the evaluation report has been delivered. The steps include planning the activities and who should be involved, the development of the Management Response (which addresses the recommendations emerging from the evaluation), the formal presentation of results by the evaluators, and tips for effectively communicating results. It also covers different ways of sharing what was learned from the evaluation exercise including through the Lessons Learned Repository and UNOCT Connect & Learn platform. This chapter is relevant to all audiences.

Chapter 7 Quality Assurance and Assessment highlights the features of a credible evaluation process. Both evaluation managers and evaluators are encouraged to review this chapter periodically as it succinctly covers the most important things to look for when conducting quality control of data collection and of each deliverable.

Chapter 8 Practical Guide presents fundamental aspects of good evaluation practice. It will be most useful to those who are less familiar with the field of evaluation, helping to demystify terms such as theory of change, evaluation matrixes, stakeholder mapping, sampling frameworks, and most significant change. It introduces different types of evaluation design and a range of data collection and analysis methods. It also provides a list of useful resources for those wanting further information.

The Annexes contain practical information as well as tools and templates for use in the evaluation planning and implementing phases and are referenced throughout the chapters of the handbook. Particularly useful for those overseeing evaluations are management checklists for each of the three evaluation categories; these cover each step and the roles and responsibilities of those supporting

the process along the way. Evaluators should note that the report templates are required to be used for each deliverable.

CHAPTER ONE: EVALUATION AND WHY IT MATTERS

1.1 WHAT IS EVALUATION?

Evaluation supports informed decision making. It does this by assessing the changes that interventions - such as programmes, policies, and strategies - have contributed to and why those changes occurred.

Evaluation in the UN system is guided by the Programme Planning, Budgeting, Monitoring and Evaluation Regulations and Rules (PPBME),[1] the 2021 Administrative Instruction on Evaluation in the United Nations Secretariat[2], and the United Nations Evaluation Group (UNEG) Norms and Standards. UNOCT uses the following UNEG definition of evaluation:

> *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, 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.*

The UNOCT Evaluation Policy provides the overall framework for planning and undertaking evaluations by the Office. It aims to ensure that all evaluative activity aligns with the Norms and Standards for Evaluation adopted by UNEG. Chapter 3 - UNOCT Evaluation Function provides the guidance for how this policy is to be carried out.

1.2 WHY EVALUATION IS USEFUL

Evaluation is an essential component of the work of UNOCT. It ensures accountability to Member States, rightsholders/beneficiaries, funding partners, and other stakeholders for the investments received. As importantly, evaluation supports organizational learning and knowledge generation by providing valuable feedback to improve performance and mitigate potential harms. It can also create the space for discussions and building partnerships. These purposes are highlighted in figure 1.1 and further discussed below.

Accountability: UNOCT is directly accountable to the General Assembly, Member States, and funding partners to show progress in meeting the intended results of an intervention and in achieving value for money.

> Evaluation is not about judging or fault-finding. Rather, it is an opportunity for listening, reflecting, and learning about what went well, what can be done better, and how to move forward.

[1] UN Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation (ST/SGB/2018/3).
[2] The ST/AI/2021/3 also guides Evaluation in the Secretariat. It prescribes instructions and procedures for the implementation of Article VII on Evaluation of the PPBME. This ST/AI also has accompanying guidelines which support its implementation.

UNOCT also has a responsibility to other partners and rights holders to show how the organization's work and ways of working are making a difference, while observing human rights and furthering gender equality. Accurate, fair, and credible evaluations go beyond monitoring reports to help demonstrate that UNOCT is a trusted partner.

Figure 1.1: The Four Main Purposes of Evaluation
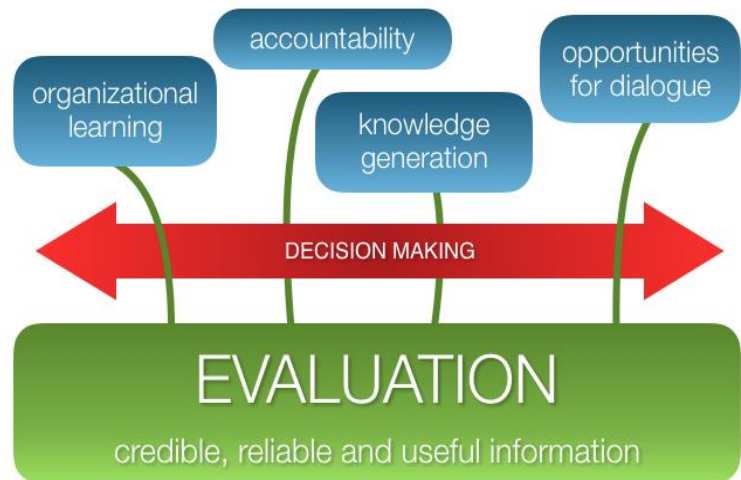
Organizational Learning: Evaluations provide an opportunity to measure progress; as importantly, they help to highlight successes and offer insights into what worked, what did not work so well, and the reasons behind both the accomplishments and challenges. Through the timely incorporation of recommendations and lessons learned into decision-making processes, evaluation aims to make programming and institutions more effective, efficient, sensitive, and sustainable.

Knowledge Generation: Counter-terrorism, and in particular preventing and countering violent extremism (PCVE) are relatively new fields marked by much trial and error in a variety of settings - this places a premium on learning. To keep pace with rapid changes in the terrorist threats and to be effective, UNOCT needs to constantly adapt and learn by basing its technical assistance and engagement in support of Member State responses on empirical evidence and knowledge gained from evaluations of project and programme implementation. Information compiled from multiple evaluations into a meta-synthesis, that is in line with the UN Global Counter-Terrorism Strategy (UNGCTS) and UNEG requirements, helps to build a body of knowledge valuable for UNOCT stakeholders as well as Member States and other organizations engaged in this field.

Opportunities for Dialogue: Inclusive evaluation processes can provide an opportunity for stakeholders to come together to discuss the object of the evaluation and to take time to reflect. This can help to build relationships and ensure a better understanding of the needs and interests for all those involved, as well as opportunities for further collaboration.

1.3 HOW EVALUATION DIFFERS FROM OTHER FORMS OF ASSESSMENT

Although evaluation processes are related to other assessment processes, they serve different purposes.

Monitoring is a continuous part of programme/project management that involves the systematic collection and analysis of data using the intervention's indicators. This data helps determine the progress being made in implementing activities, achieving outputs and milestones, and using allocated resources. Monitoring aims to keep interventions on track and ensure timely decision-making needed to improve programme/project design, ensure alignment and functioning. It typically relies on quantitative information collected on individual indicators on an ongoing basis by programme or partner staff. Such information is an important source of data used in evaluation processes to understand what is happening.

Evaluations often require stakeholders to take time out of their schedules to participate in interviews and group discussions, or to answer survey questions. These processes create the space to pause and reflect, which may prompt useful insights into, for example, significant moments in the intervention's lifecycle and ideas about what to do differently. This may even inspire new initiatives or new partnerships.

In contrast, evaluation is conducted at specific points in time and focuses more on outcome and higher-level results. It usually draws from multiple sources and types of data, enabling a more comprehensive assessment of what happened, why, and how. With the exception of internal/self- and participatory evaluation processes, evaluations are generally undertaken by independent, external consultants.

Audits are an assessment of the adequacy of management controls. They are meant to ensure the economical and efficient use of resources; the safeguarding of assets; the reliability of financial and other information; the compliance with regulations, rules, and policies; the effectiveness of risk management; and the adequacy of organizational structures, systems, and processes. Evaluation is more closely linked to managing for results and learning, while audits mainly focus on compliance.

Inspections are a general examination of an organizational unit, issue, or practice. They look at the extent to which there is adherence to prescribed standards, good practices, or other criteria. Information gained is used to make recommendations for improvement or corrective action. Inspections are often performed when there is a perceived risk of non-compliance.

Reviews are a periodic or ad hoc assessment that typically addresses performance and operational issues of programme/project implementation. Examples include rapid assessments (often conducted as a part of programme/project design) and evaluability assessments[3] (often conducted prior to an evaluation). Reviews are usually undertaken internally and tend to be less rigorous than evaluations.

Research is a systematic examination undertaken to develop or contribute to knowledge of a particular topic. It generally involves a narrower focus as well as a more rigorous scientific methodology and analysis than an evaluation. Research often feeds information into evaluations, other assessments, or decision-making processes. Examples include in-depth baseline studies and impact studies.

1.4 HOW EVALUATION LINKS TO RBM

Evaluation is a fundamental part of Results-Based Management (RBM). RBM is one of the core programming principles for UNOCT and the UN system more broadly. Its definition has evolved since it was first introduced by the Organization for Economic Co-operation and Development's Development Assistance Committee (OECD DAC) in 2007, with **more emphasis now being placed on how results are used**. A current and commonly used definition is:

> *"RBM is a management strategy by which all actors, contributing directly or indirectly to achieving a set of results, ensure that the processes, products, and services contribute to the achievement of desired results (outputs, outcomes and higher-level goals or impact). The actors in turn use information and evidence on actual results to inform decision making on the design, resourcing and delivery of programmes and activities as well as for accountability and reporting"*.[4]

RBM uses a structured, logical approach that identifies expected results and the inputs and activities necessary to achieve them. It aims to promote management effectiveness and accountability through:

- Clearly defining realistic results and targets.

---

[3] An evaluability assessment is carried out to determine the readiness of an intervention to be evaluated. This process typically involves reviewing project documents and monitoring system to make sure the intended results are measurable, and that needed data is being collected so that an evaluation would be feasible and useful.

[4] United Nations Development Group (UNDG), 2011. Results-Based Management Handbook: Harmonizing RBM concepts and approaches for improved development results at country level.

- Linking planned activities to the results to be achieved.
- Monitoring progress towards the achievement of expected results and targets using predetermined performance indicators.
- Assessing whether results were achieved and why, through evaluations.
- Integrating lessons learned into management decisions.
- Reporting on performance.

RBM is first applied during the planning stage of an intervention. It involves the development of a results framework which is ideally based on a theory of change or programme theory. The results framework is important because it shows the links and cause-and-effect relationships between the ultimate goal (also referred to as "objective" or" impact"), the intermediate-term ("outcome"), the short-term results ("output"), and the activities and resources ("inputs") for achieving those results. Evaluations are typically focused on the outcome- and higher-level results as these are the results that the intervention is expected to achieve.

A simple way of showing these relationships is by developing a results chain. Figure 1.2 provides the elements of a results chain and shows which elements are considered results, and how the different types of results build upon each other. More information about results frameworks and programme theories can be found in Chapter 8 Guide to Good Evaluation Practice.

Figure 1.2: Results Chain
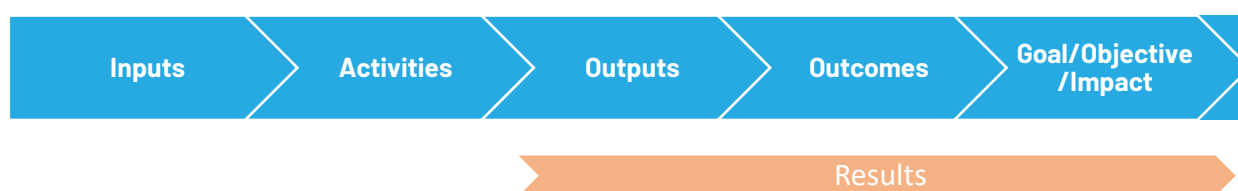
Figure 1.3 below shows how evaluation fits within the entire RBM lifecycle; the main components being (a) planning the intervention, (b) monitoring its implementation, and (c) evaluating to inform further planning and implementation. Even though evaluation happens at a set point in time (most commonly at mid-point and at the end of the intervention), it needs to be considered throughout the lifecycle.

- At the planning phase of the programme/project, the results and performance indicators need to be defined in a way that they can be evaluated (and presented in the results framework), and an evaluation plan and budget should be determined (within UNOCT, this document is the Evaluation Approach).
- At the monitoring phase, it is important that data being collected through the use of pre-determined indicators can be used in the upcoming evaluation(s).

- At the evaluation phase, it is essential that the results, recommendations, and lessons are developed ways that are useful for future programming and organizational learning purposes.

1.5 UNOCT'S RBM SYSTEM

The UNOCT Strategic Plan and Results Framework (SPRF) provides the overarching strategy and direction for the Office. In addition to the Office's Vision and Mission, it sets out a Theory of Change and accompanying Results Framework. The operationalization and monitoring of the content within the Results Framework supports accountability and transparency, forms the basis of planning including for internal workplans, and provides opportunities to demonstrate tangible improvements in results and a 'results culture'.
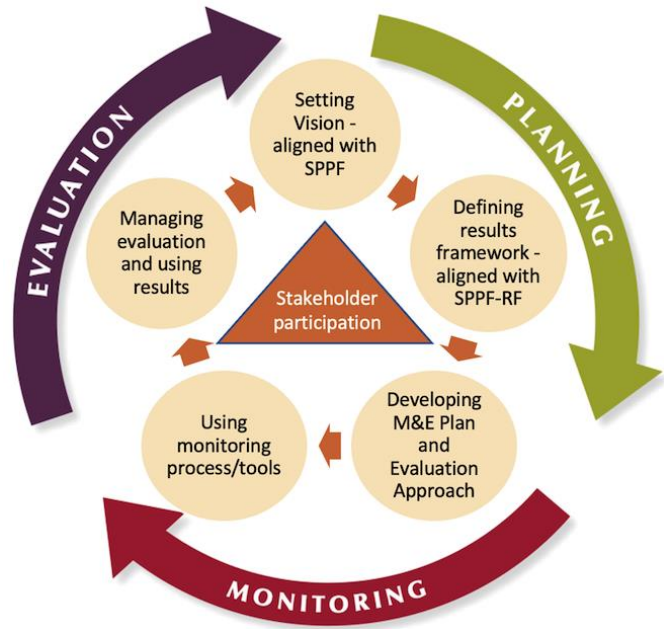
Figure 1.3: Adapted from UNDP, Handbook on Planning, Monitoring and Evaluating for Development Results, 2009

Evaluation then has a critical role in determining and confirming the extent to which results have been attained and the progress that is being made in achieving the UNOCT's Strategic Goals[5]:

Strategic Goal 1: Foster further unity and collaboration within the United Nations against terrorism.

Strategic Goal 2: Create resilience against violent extremism conducive to terrorism.

Strategic Goal 3. Reinforce responses to terrorist threats and attacks.

Strategic Goal 4. Mitigate the risks and impact of terrorism.

Strategic Goal 5. Promote human-rights compliant and gender responsive CT/PCVE efforts.

All evaluations of UNOCT interventions are expected to consider the alignment and contribution of the intervention to UNOCT priority areas. Guidance for doing this is provided in Chapter 3 UNOCT Evaluation Function.

UNOCT is inspired by the vision:

Together, Building a Future Without Terrorism

Mission: UNOCT works together with Member States, the United Nations system, international, regional and subregional organizations, civil society, private sector and other partners in addressing terrorism across its lifecycle, and in supporting them in the implementation of international obligations and instruments to prevent and counter terrorism in accordance with the rule of law, and in a human rights compliant and gender responsive manner.

---

[5] The five strategic goals listed are in the UNOCT Strategic Plan and Results Framework 2022-2025.

1.6 THE CHALLENGE OF MEASURING IMPACT

There is increasing interest in being able to measure higher-level results of an intervention. Understandably, stakeholders want to know what significant difference is being made and if the systematic transformation that is envisioned by an organization's strategic goals is occurring. However, this can be challenging to assess as impact (or objective) generally refers to the longer-term change that an intervention is expected to contribute towards, and it may take years for such changes to become apparent. A further challenge is that it often can be difficult to attribute the observed changes to a particular intervention versus other factors. Since achieving the intended impact is generally beyond the direct control of the intervention, evaluation in the UN system, including UNOCT, primarily focuses on the changes the intervention should be able to control – the outcome-level results.

Depending on the type of intervention being evaluated, there can still be the need to account to stakeholders that progress towards impact is being made. The way in which impact is considered in most UN evaluations is discussed in the next chapter. Examples of questions that aim to capture impact–related information are also shown in Annex 2 of this handbook.

Impact Evaluations

Care must be taken in using the term 'Impact Evaluation' as it is often confused with the broader understanding of impact described above. Impact Evaluation, on the other hand, refers to "specific methodologies for establishing statistically significant causal relationships between the intervention and observed effects" [1]. It is commonly understood as only those evaluations that use a counterfactual (usually a comparison group) to determine what the outcomes would have been in the absence of the intervention. Impact evaluation can also be even more narrowly defined as only those evaluations that use randomized controlled trials (RCTs), whereby both the population receiving the intervention and the control group are chosen at random from the same eligible population. RCTs are not feasible or appropriate in many situations. However, if RCTs are being considered, they need be planned well before the intervention begins.

All types of impact evaluation are resource intensive as they typically require considerably more primary data than for other types of evaluations, usually with a heavy reliance on quantitative survey data. Impact evaluations tend to focus on a tightly defined set of impacts and are usually not designed to answer as many questions as regular evaluation processes. They can focus on results anywhere along the results chain, so do not necessarily look at higher-level change.

Although well-conducted impact evaluations provide important information on the difference that an intervention has made, they should only be built into an evaluation design when resources are likely to be available or can be obtained for this purpose.

For further information see:

OECD/DAC Principles of Impact Evaluation and Better Evaluation (UNICEF) - Randomized Controlled Trials

UNOCT's International Hub on Behavioural Insights to Counter Terrorism (BI Hub)

The BI Hub conducts and advances research in behavioural sciences to better understand the drivers and factors contributing to radicalization to violent extremism and terrorism, and for de-radicalization, desistance, and resilience to radicalization. The BI Hub is a resource for both programme designers and practitioners, including for the planning and implementation of evaluations.

In order to maximize outcomes and cost effectiveness of programme interventions, the behavioural insights approach recommends using well-validated monitoring and evaluation (M&E) data to adequately assess their level of success or failure.

BI Hub's M&E Database is designed to assist practitioners by offering a well-validated selection of measurement and evaluation tools, categorized by variables of interest - including emotional resilience, radicalization, psychological trauma, polarization of opinion, and support for extremism. Instructions and guidance are provided on how to use and score each of these. Over time, the use of this M&E database by different programmes will enable evidence-based comparisons of the efficacy of different types of interventions, including comparing the cost and duration required for a unit of improvement on a particular measure.

Within UNOCT, concerted efforts are being made to measure impact using more rigorous methodological approaches, most notably by the International Hub on Behavioural Insights. This work is often being done as research rather than evaluation but in some cases also bridges the two.

CHAPTER TWO: EVALUATION IN THE UN SYSTEM

This chapter presents the context for undertaking evaluations within UNOCT and within the United Nations more generally. It is primarily aimed at those conducting evaluations. It is important for evaluators to understand how the intervention they are evaluating fits with the structure of the wider UN system and the global mandate of UNOCT, as well as the standards, expectations, and ethical principles to be followed for evaluations within the system. This chapter also provides useful background information for evaluation managers, particularly for showing how the evaluation of individual interventions support the broader work of the UN system, how the cross-cutting themes (of human rights, gender equality, disability, and other aspects of inclusion) are to be integrated, and how to approach evaluations conducted jointly with other agencies.

2.1 SDGS AS THE CONTEXT

The 2030 Agenda for Sustainable Development guides the work of the United Nations system. The 17 Sustainable Development Goals (SDGs) have provided a transformational vision for the world where no one is left behind. They spell out commitments for working together to eradicate poverty, create decent jobs, and promote dignity, equality, and justice for all, while sustaining the natural environment. Each goal has targets (of which there are 169 targets in total), and each target has indicators from which progress can be measured. **Evaluation plays a critical role in assessing progress and the level of commitment towards the attainment of the SDG targets**, and it is incumbent upon those commissioning evaluations and evaluators to undertake such assessments in UNOCT.

Evaluations play an important role in the 2030 Agenda for Sustainable Development by helping to ensure that policies and programming are aligned with the SDGs and corresponding targets. To the extent feasible, evaluations should:

- Include performance indicators that connect with or strive to measure at least one SDG indicator.

- Examine the connection of UNOCT's work to applicable SDGs as part of the assessment of relevance of the intervention.

There is substantial overlap and linkages amongst the SDGs. UNOCT programming supports the achievement of several goals, the most thematically relevant being **SDG 16** – *Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable, and inclusive institutions at all levels.*

The most pertinent targets are *16.3 - Promote the rule of law at the national and international levels and ensure equal access to justice for all* and **16.a** – *Strengthen relevant national institutions, including through international cooperation, for building capacity at all levels, in particular in developing countries, to prevent violence and combat terrorism and crime.*

As gender equality is central to the implementation of the GCTS and **SDG 5 -** *Achieve gender equality and empower all women and girls* – this principle is also highly relevant to UNOCT programming.

Evaluations should, therefore, consider the extent to which UNOCT's programming priorities and results contribute to the SDGs, and how UNOCT's work has assisted countries in advancing achievements of the SDGs, in particular SDG 5 and SDG 16 - target 16.a. Evaluators should also pay attention to other SDGs that are pertinent to the intervention being assessed.

2.2 DELIVERING AS ONE

The 'Delivering as One UN' agenda aims to provide UN development assistance in a more coordinated way at the country level. It increases opportunities to fully access the range of mandates and capacities by leveraging the comparative advantages of each agency and combined synergies to accelerate progress in meeting the SDGs.

In each Member State, the UN country team brings all UN entities working in the country together with government counterparts to plan that country's road map towards its commitments under the 2030 Agenda. The resulting agreement between the Member State and the UN country team is called the United Nations Sustainable Development Cooperation Framework (UNSDCF). **This Framework is used as the basis for aligning the intervention being evaluated with the needs and priorities of the Member State as well as the 'One United Nations' approach** to programme delivery in each country. Such alignment helps to ensure cohesiveness of individual programmes and the overall mission of the UN and is therefore an important part of **assessing coherence** of the intervention.

2.3 UNGCTS AND LINK TO SDGS

The UN Global Counter-Terrorism Strategy (UNGCTS) is a unique global instrument to enhance national, regional, and international efforts to counter terrorism. It was adopted by consensus in 2006, marking the first time that all UN Member States agreed to a common strategic and operational approach to fighting terrorism. The United Nations General Assembly (UNGA) reviews the UNGCTS every two years to ensure its continued alignment with Member States' counter-terrorism (CT) priorities. To facilitate this work, the UN Global Counter-Terrorism Coordination Compact was established in 2018. The Counter-Terrorism Coordination Committee of the Compact is responsible for providing oversight and strategic level guidance to eight inter-agency Working Groups on the implementation of the UNGCTS. **UNOCT provides Secretariat support to the Counter-Terrorism Compact**.

> UNOCT works within the peace, security, and development nexus which recognizes that development cannot happen in the absence of peace. The UNGCTS provides an enabling environment for organizations to work together, creating the conditions for a peaceful society and achievement of the Sustainable Development Goals.

The UNGCTS, in the form of a resolution and an annexed Plan of Action A/RES/60/288 is composed of four pillars, namely:

- Addressing the conditions conducive to the spread of terrorism.
- Measures to prevent and combat terrorism.
- Measures to build states' capacity to prevent and combat terrorism and to strengthen the role of the United Nations system in that regard.
- Measures to ensure respect for human rights for all and the rule of law as the fundamental basis for the fight against terrorism.

UNOCT was created in 2017 through the adoption of UNGA Resolution A/RES/71/291, transferring the Counter-Terrorism Implementation Task Force and the UN Counter-Terrorism Centre, out of the Department of Political Affairs, into the Office of Counter-Terrorism.

UNOCT has five main functions, namely:

1. Provide leadership on the General Assembly counter-terrorism mandates entrusted to the Secretary-General from across the United Nations system.

2. Enhance coordination and coherence across the Global Counter-Terrorism Coordination Compact entities to ensure the balanced implementation of the four pillars of the UN Global Counter-Terrorism Strategy.
3. Strengthen the delivery of United Nations counter-terrorism capacity-building assistance to Member States.
4. Improve visibility, advocacy, and resource mobilization for United Nations counter-terrorism efforts.
5. Ensure that due priority is given to counter-terrorism across the United Nations system and that the important work on preventing violent extremism is firmly rooted in the Strategy.

The Global Counter-Terrorism Coordination Compact brings together 46[6] entities, as members or observers, including 41 United Nations entities, as well as INTERPOL, the World Customs Organization (WCO), the Inter-Parliamentary Union (IPU), the Financial Action Task Force (FATF), and the International Monetary Fund. The launch of the Counter-Terrorism Compact is recognized as a major step forward in building coordination and coherence in the work of the UN system on preventing and countering terrorism and the underlying spread of violent extremism, while respecting the existing mandates of each entity.

When beginning an evaluation, it is important to understand UNOCT's functions and how its work is organized and contributes to the four pillars of the UNGCTS. It is also important to know the other entities with which UNOCT collaborates within the implementation of each of its interventions.

2.4 GUIDING PRINCIPLES, NORMS AND STANDARDS

Two entities have been instrumental in harmonizing and setting standards for how evaluations are to be conducted within the UN system. The United Nations Evaluation Group (UNEG) is an interagency professional network that brings together the evaluation units of the UN system. UNOCT is a member of UNEG and has adopted its norms and standards. The Organization for Economic Co-operation and Development (OECD) provides a forum for governments and a range of partners to collaborate on key global issues. Its mission is to *"promote policies that will improve the economic and social well-being of people around the world".* OECD, through its Development Assistance Committee (OECD DAC), was responsible for developing the widely used evaluation criteria that have provided a normative framework for determining the merit or worth of an intervention (policy, strategy, programme, project, or activity).

This sub-section discusses the key contributions of UNEG and OECD DAC. They provide a baseline for evaluation practice, but it is important to note that each UN entity and their funders may have additional evaluation requirements or areas of emphasis that are specific to their mandates and areas of work.

---

[6] As of May 2023.

UNEG's mission is to promote, strengthen and advocate for a robust, influential, independent, and credible evaluation function throughout the UN system for decision-making, accountability, and learning. UNEG aims to:

- Set the UN evaluation normative framework to reflect emerging or innovative developments as well as good practices in evaluation.
- Provide a forum for professional engagement, support, and exchange globally and locally.
- Advocate for the strategic use of evaluations in all UN entities and beyond to inform decision-making and enhance results.
- Influence evaluation practices in all entities of the UN system.
- Encourage partnerships for joint evaluation.

2.5 UNEG NORMS AND STANDARDS

UNOCT upholds and promotes the evaluation practices, principles, and values to which the UN is committed. Specifically, the following UNEG norms apply to the conduct of evaluations.

NORMS	DESCRIPTION
Utility	There must be clear intention to use the evaluation's analysis, conclusions, and recommendations. This includes relevant and timely contributions to organizational learning, decision-making, and accountability.
Credibility	Credibility requires independence, impartiality, rigorous methodology, and ethical conduct. Key elements include transparent processes, inclusive approaches involving relevant stakeholders, and robust quality assurance systems.
Independence	Evaluators must have the freedom to conduct their work without influence from any party or threat to their careers. In addition, the organization's evaluation function must be positioned separately from other management functions, be responsible for setting the evaluation agenda, and have adequate resources to do its work.
Impartiality	The key elements of impartiality are objectivity, professional integrity, and absence of bias. The requirement for impartiality exists at all stages of the evaluation process, including planning an evaluation, formulating the mandate and scope, selecting the evaluation team, providing access to stakeholders, conducting the evaluation, and formulating findings and recommendations.
Ethics	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. Evaluators must respect the rights of institutions and individuals to provide information in confidence and must ensure that sensitive data is protected.
Transparency	Transparency is an essential element of evaluation that establishes trust and builds confidence, enhances stakeholder ownership, and increases public accountability. Evaluation reports shall be timely and widely disseminated to all stakeholders in line with the disclosure and dissemination parameters established by the Office.

| Human Rights and Gender Equality	Human rights and gender equality will be integrated into all stages of an evaluation. A human rights-based approach and gender equality mainstreaming promote, among others, the commitment to the principle of 'leave no-one behind'.
Professionalism	To ensure credibility, evaluations need to be conducted with professionalism and integrity. These are supported by an enabling environment, institutional structures, and adequate resources.

2.6 OECD-DAC EVALUATION CRITERIA

The Development Assistance Committee of OECD (OECD DAC) and its Network on Development Evaluation (EvalNet) issued a set of widely adopted criteria in the early 2000s to structure the lines of enquiry for evaluation and serve as the basis upon which evaluative judgements are made. In 2021, 'Applying Evaluation Criteria Thoughtfully' was issued; this new guidance aims to help evaluators and others better understand those criteria and improve their use. It also introduced a sixth criteria, Coherence.

Figure 2.1 highlights the full set of OECD-DAC evaluation criteria with each having an embedded link that:

> The new OECD-DAC guidance is a highly recommended and accessible resource.

- Defines and explain the criteria.
- Provides advice on how to use and adapt them.
- Gives tips to avoid common pitfalls.
- Provides examples from real life evaluations.

Figure 2.1: OECD DAC Evaluation Criteria Explained (via hyperlinks)
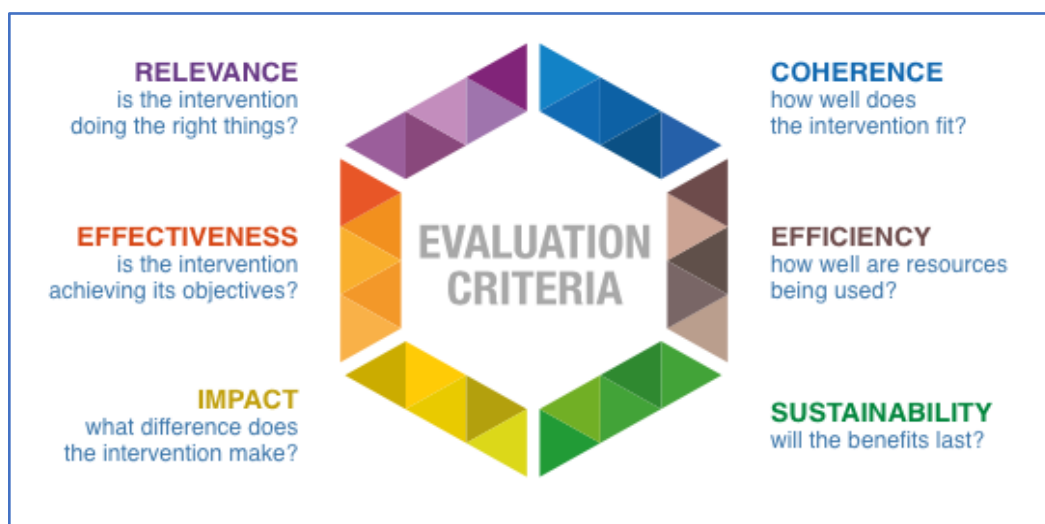
All **UNOCT evaluations are normally expected to address the criteria of Relevance, Coherence, Effectiveness, Efficiency, Impact and Sustainability**. In line with the OECD DAC guidance on the thoughtful application of the standard criteria, and with the increasing focus on producing more concise evaluation reports, it is understood that the scope and subject of the evaluation will be important determining factors in:

- Which of the criteria will have more focus (which usually means having more evaluation questions).
- Whether any of the standard criteria are not needed.

There may be additional areas to consider beyond the standard DAC criteria that may be useful for the context of the evaluation and user needs. One example is "Appropriateness" which can be helpful in understanding the extent to which an intervention, as it has been designed, is the best way to achieve the intended outcomes. There are also criteria specific to evaluations of humanitarian-related interventions, such as "Reach", and "Partnerships", which can also be useful for non-humanitarian contexts, and some funding agencies have further required criteria to consider (as discussed in Chapter 3 in the section on special considerations for evaluating UNOCT-supported interventions). In most cases these additional areas of enquiry can be addressed through evaluation questions under one of the standard criteria. If an in-depth look at such areas is warranted, they can be addressed as a separate criterion.

It is important to note that any changes to the use of the standard OECD-DAC criteria need to be explained in the Terms of Reference for the evaluation and in the evaluation report.

> Evaluations of policy-related interventions highlight the need for flexibility in the application of the OECD DAC criteria. While all of the criteria will typically be used, emphasis will likely by placed on Coherence, specifically policy coherence. The impact of policy work can be particularly challenging to assess without an extended time horizon, but some agencies are now looking at it from an *Orientations to Impact* approach. This is a way to assess the programme's likely direction towards impact, as opposed to actual impact results.[7]

2.7 INTEGRATING HUMAN RIGHTS, GENDER EQUALITY, LEAVE NO ONE BEHIND, DISABILITY INCLUSION and YOUTH

The promotion of human rights, gender equality, disability inclusion, and 'leave no one behind' are guiding principles for all United Nations entities. As such, **these interrelated issues need to be incorporated into all UN evaluations regardless of whether these issues are the focus of the intervention itself.**

The 2014 UNEG Guidance Document on Integrating Human Rights and Gender Equality (HRGE) in Evaluations is the field guide for improving HRGE in the UN system. It also provides the foundation for looking at Leave No One Behind (and, relatedly, intersectionality), and disability inclusion. These topics are commonly referred to as cross-cutting themes and are expected to be mainstreamed into all UN evaluation processes. Where relevant, youth should also be considered as a cross-cutting theme – which will be the case for evaluations of many UNOCT interventions.

This sub-section provides an overview of each of these themes and the underpinning international instruments and provisions on which they are based.

Human Rights

"Human rights are commonly understood as being those rights which are inherent to the human being. The concept of human rights acknowledges that every single human being is entitled to enjoy his or her human rights without distinction as to race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status."

(OHCHR, Human Rights: A Basic Handbook for UN Staff)

---

[7] The Office of the High Commissioner for Human Rights (OHCHR) has developed preliminary indicators for measuring Orientations to Impact. See UNEG Evaluating Policy Influencing Stocktaking Report, 2022 (pp. 5-6).

The 1948 Universal Declaration of Human Rights laid out, for the first time, a set of human rights to be protected. Since then, nine core human rights instruments and their optional protocols have been adopted under the aegis of the UN, together with many more regional instruments. A list of human rights treaties and their optional protocols adopted under the aegis of the UN can be found on the website of the Office of the High Commissioner for Human Rights (UN Human Rights).

> Further information on human rights instruments and processes on the OHCHR website can be found through the following links:
>
> - o Core international Human Rights Instruments
> - o Human rights Treaty Bodies
> - o Jurisprudence of the Treaty Bodies
> - o Human Rights Council
> - o Universal Periodic Review
> - o Special Procedures of the Human Rights Council
> - o Communications by Special Procedures

A human rights-based approach (HRBA) is the strategy for incorporating human rights in United Nations programming. It is a conceptual framework *"that is normatively based on international human rights standards and operationally directed to promoting and protecting human rights."[8]* The six key principles of HRBA are (i) universality and inalienability; (ii) indivisibility; (iii) interdependence and interrelatedness; (iv) non-discrimination and equality; (v) participation and inclusion; and (vi) accountability and the rule of law. Of these key principles, non-discrimination and equality, participation and inclusion, and accountability and the rule of law are particularly relevant to evaluations.

A key aspect of HRBA is the consideration of the obligations of **duty bearers** (principally States) to act in order to respect, protect, and fulfil the human rights of persons under their jurisdiction (**rights holders**). Within the sphere of UN programming, the term 'duty bearers' extends beyond States to also include non-State actors, i.e., any actor with either duties or responsibilities for a development intervention which can additionally be understood to comprise 'moral duty bearers' such as parents and family members as well as corporate entities and UN agencies. (UNEG, 2014).[9] This approach also dispels the notion of rights holders as passive recipients or beneficiaries.

Ultimately, **the rights-based approach requires ensuring that interventions facilitate the claims of rights-holders and the corresponding responsibilities of duty-bearers**. It also emphasizes the importance of addressing the immediate, underlying, and structural causes for such rights not being realized, which in many cases means supporting the development of the capacities of duty bearers to meet their obligations (e.g., knowledge of policies and responsibilities, subject matter expertise, available resources, and tools).

HRBA is simultaneously 1) a goal (i.e., all activities should further the realization of internationally recognized human rights), 2) a process (i.e., international human rights norms and standards should guide all activities at all phases of the programming process), and 3) an outcome (i.e., programming contributes to the capacities of duty bearers to meet their responsibilities/ obligations and/ or that of rights holders to claim their rights).

---

[8] OHCHR, 'Frequently Asked Questions on a Human Rights-Based Approach to Development Cooperation', 2006, https://www.ohchr.org/sites/default/files/Documents/Publications/FAQen.pdf
[9] Please note that, for the purposes of this Handbook, the term 'duty bearer' is used differently from its meaning under international human rights law, also encompassing a variety of actors that would not qualify as duty bearers in the international law meaning of the term.

Human rights-responsive evaluations can be characterized as having the following features. They:

- Are grounded in human rights risk and opportunity analysis conducted prior to the commencement of the activities and assess the ways in which the findings of the analysis have been incorporated throughout all phases. They also consider the appropriateness and effectiveness of any measures taken to address the risks that were identified at the outset or emerged during the project cycle.

- Assess the extent to which an initiative has increased the likelihood that rights holders claim their rights and that duty-bearers fulfil their obligations. Factors in achieving this include improved understanding of rights and duties and why this matters, increased willingness to claim/fulfil rights, and increased capacity to do so by removing obstacles and structural barriers.

- Assess the extent to which results frameworks integrate specific, measurable, attributable, reliable, and targeted indicators that facilitate the evaluation of ways in which activities promote respect for international human rights law (as well as international humanitarian law and refugee law, where applicable).

- Address and seek to balance power relations between and/ or within groups of duty bearers and rights holders. To do this, evaluators must have a full understanding of the context and dynamics concerning power relations as well as their own position in this respect.

- Are participatory by involving stakeholders in the evaluation process as far as feasible, including in decisions about what is to be evaluated and how the evaluation is conducted (i.e., by being part of the Evaluation Reference Group or an evaluation advisory group). Decisions related to participation should be informed by the human rights impact and opportunity assessment and pay due attention to stakeholders who may be most impacted by the intervention.

- Use gender-sensitive and human rights-based language throughout the report (including by referring to 'rights holders' and 'duty bearers').

- Incorporate human rights considerations into lessons learned and recommendations.

A **stakeholder analysis** (also referred to as stakeholder mapping) is important for ensuring evaluations are inclusive of the range of duty bearers and rights holders involved in the programme/project being assessed. In line with UN mandates, ensuring stakeholders' participation, including of vulnerable populations and both men and women, is an obligation of the UN, and it is the right of every beneficiary to have a say on processes and interventions that affect their lives.

Gender Equality

Gender equality is one of the founding principles of the UN. The Charter of the United Nations first article describes one of its main purposes as being *"promoting and encouraging respect for human rights and fundamental freedoms for all 'without distinction as to race, sex, language, or religion"*. The central international legal instrument promoting gender equality is the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW).

Mainstreaming a gender perspective refers *to "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"*[10]. As further defined in the UNEG Guidance on HRGE in evaluations, gender mainstreaming is a strategy for making gender equality-related concerns

---

[10] This is the definition of gender mainstreaming used by UNOCT and is from the United Nations Economic and Social Council (ECOSOC), in its agreed conclusion 1997/2.

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.

UN evaluations should assess the extent to which UN interventions have addressed gender equality considerations in their design, implementation, and outcomes, and whether women and men as well as boys and girls, can equally access the programme's benefits to the degree intended.

Gender-responsive evaluations employ a human right-based approach and also:

- Assess the quality of the context specific human rights and gender analysis undertaken before the start of the intervention, and whether the results of the analysis were integrated into programme design.

- Provide credible and reliable evidence-based information about the extent to which an intervention has resulted in progress (or the lack thereof) towards intended and/or unintended results regarding gender equality and the empowerment of women; as well as assessing gendered harms of the intervention and proposes mitigation measures.

> Evaluation Managers should identify all stakeholders involved in the intervention, their roles, and their interests in the evaluation. This will then provide the basis for which stakeholder groups should be invited to participate as respondents (be included in the sampling framework) in the evaluation. A sample stakeholder analysis matrix is provided in **Annex 7(d)**.

- Assess the extent to which results frameworks integrate specific, measurable, attributable, reliable, and targeted indicators that support the collection and analysis of the intervention's commitments on gender.

- Employ a solid context-specific gender analysis to assess all results, including to determine the extent to which and how the intervention has impacted men and women differently, taking into account intersectional identity factors, such as nationality, age, belief, sexual orientation, race, ethnicity, geographic origin, ability, socio-economic status, and other identity factors, and combinations thereof and whether it has challenged and changed inequalities, including structural causes[11].

- Foster inclusive participation of different stakeholders, particularly women, as well as those individuals or groups who are marginalized or vulnerable due to their status or situation (e.g., members of ethnic, religious minorities, refugees/asylum-seekers, persons with disabilities) and therefore their rights are more likely to had been violated or negatively impacted by the intervention. (Previously conducted context specific human rights and gender analysis can help identifying those groups).

- Minimize barriers to participation – including timing of meetings, language, power dynamics, accessibility of location, etc.

- Ensure respect for cultural differences which includes practices such as acknowledging the complexity of cultural identities, employing context specific gender sensitive

> Evaluation teams should have expertise in HRGE. The UNEG Guidance provides useful tips on resources needed and what this investment would provide. (pp. 66–67)

---

[11] Gender transformative programming and methodologies look closely at the structural barriers to gender equality and how transformative change can be achieved. More information can be found in the joint publication of FAO, IFAD, WFP, EU and CGIAR Gender Platform. 2022.

methods of consultation, and ensuring context specific cultural and gender related competencies of evaluators.

- Ensure that gender is taken into consideration in risk assessments and lessons learned.

Gender-responsive approaches within the context of CT/PCVE

Gender-responsive approaches require:

- "Recognizing the differential impact on the human rights of women and girls of terrorism and violent extremism, including in the context of their health, education, and participation in public life" (S/RES/2242).

- Understanding the "many different roles, including as supporters, facilitators, or perpetrators" that women may assume with regards to terrorist acts as well as the fact that "women are impacted both by violent extremist and terrorist groups as well as the strategies used to counter them, which can result in their being squeezed between terror and anti-terror" (S/2015/716).

- Addressing the ways terrorist and violent extremist groups "use the promise of marriage and access to sex to incentivize recruitment of men and boys, engage in trafficking and other gendered practices that promote and reinforce violent masculinities, perpetuate sexual and gender-based violence and the persecution of individuals on the basis of their sexual orientation or gender identity and [...] continue to demonstrate foundational ideologies and cultures of violent misogyny" (S/2018/900).

- "Integrat(ing) gender analysis on the drivers of radicalization of women and men to terrorism" (A/RES/75/291)

- Considering "the specific impacts of counter-terrorism strategies on women and women's organizations and to seek greater consultations with them when developing strategies to prevent and counter terrorism and violent extremism conducive to terrorism" (A/RES/75/291).

Mainstreaming gender across Counter Terrorism/Preventing and Countering Violent Extremism (CT/PCVE) also acknowledges the enhanced vulnerability of women and girls in conflict-affected contexts, while also understanding and addressing the ways men and boys of diverse sexual orientations, gender identities, and performances of masculinities are impacted by and impact violent extremism, terrorism and CT/PCVE.

Accountability Process for Gender Mainstreaming: The United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP) was adopted in 2012 to make the strategy of gender mainstreaming operational. Since 2013, on a yearly basis, all entities are required to report on their progress in meeting indicators specific to gender equality and the empowerment of women and girls (GEEW/G). This includes the ratings of all evaluation reports for the evaluation performance indicator of the UN-SWAP reporting. In 2018, UN SWAP 1.0 was replaced with UN SWAP 2.0, which refined and expanded the indicators, aligning it with the 2030 Agenda for Sustainable Development. Within UNOCT, the Gender Unit of the Human Rights and Gender Section prepares the annual UN SWAP report based on inputs from all relevant UNOCT sections.

Since 2019, UNOCT has taken important steps to mainstream gender, such as establishing a Gender Unit with a dedicated Gender Advisor and Team, integrating gender analysis into the UNOCT Standard Operating Procedure on Project Initiation, implementing the use of gender markers and gender-responsive results frameworks, setting up a Gender Taskforce with a network of Gender Focal Points, as well as producing guidelines on gender mainstreaming for project managers. These efforts culminated in the development of the UNOCT Gender Mainstreaming Policy and the Action Plan, which was adopted in December 2021.

All evaluators should review the Gender Equality page of the UNOCT website and its video on the gendered aspects of violent extremism on how gender intersects with CT/PCVE and the goals of the GCTS.

The UNOCT Gender Mainstreaming Policy sets the Principles, Priority Areas, and Action Plan to guide UNOCT's direction and strategy for mainstreaming gender equality and the empowerment of women throughout its activities. Through this Policy, the Office aims to ensure that:
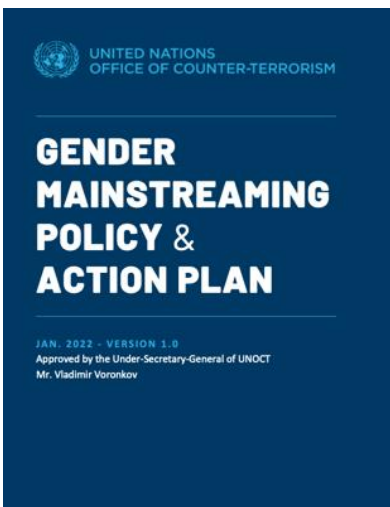
- Its work accounts for the differential impact of both terrorism and CT/PCVE measures on the human rights and needs of women, men, girls, and boys.
- All CT/PCVE initiatives conducted by the Office are well-informed, gender-sensitive, context and human-rights based, and do not reinforce inequalities or jeopardize GEEW/G gains, but rather strengthen them.

Importantly, the Policy also builds on the **knowledge on gender and counter-terrorism and on countering and prevention of violent extremism as conducive to terrorism** produced in related thematic areas by strategic partners. As noted in the Policy, gender mainstreaming is essential for UNOCT to fulfil its mandate of promoting effective implementation of the GCTS[12]; this includes the necessity to:

. . . integrate a gender analysis on the drivers of radicalization of women and men to terrorism, to consider the impacts of counter-terrorism strategies on women's human rights and women's organizations, to enhance cooperation to prevent and combat terrorism in a human rights-based and gender-responsive manner, and to seek greater consultations with women and women's organizations when developing strategies to counter-terrorism and violent extremism conducive to terrorism. [p. 5]

In regard to evaluation, the UNOCT Evaluation Policy aligns with the Office's Gender Mainstreaming Policy, stating that:

"The universally recognized values and principles of human rights and gender equality will be integrated into all stages of an evaluation to promote, among others, the commitment to the principle of 'no-one left behind". "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."

"In addition, all evaluations conducted or commissioned by UNOCT 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, and the United Nations Disability Inclusion Strategy. 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"

---

[12] General Assembly Resolution 75/291.

Disability Inclusion

The United Nations Disability Inclusion Strategy (UNDIS), launched in 2019, provides the foundation for sustainable and transformative progress on disability inclusion (DI) through all the pillars of the UN's work. Evaluation is an important tool for determining the extent to which the strategy is being achieved within the UN system.

All UN entities are required to report annually on their performance in meeting the elements set out in the UNDIS Entity Accountability Framework. **The expectation is for evaluations to assess the extent to which the interventions being reviewed are disability inclusive and for evaluations to be carried out in a disability inclusive manner.** More specifically, the required elements for evaluation are that:

- The ToR of evaluations pay 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.
- Stakeholder mapping and data collection methods involve persons with disabilities (PWD) and Organizations of Persons with Disabilities (OPDs).
- Evaluation findings and analysis provide data and evidence on disability inclusion.
- The conclusions and/or recommendations reflect the evaluators' findings on disability inclusion.

Further information on incorporating DI into evaluations is provided in the UNEG Guidance on Integrating Disability Inclusion in Evaluations and Reporting on the UNDIS Entity Accountability Framework Evaluation Indicator (2022).

UNEG advises programmes and entities that do not specifically address PWD to always take an intersectional approach to assessing disability inclusion, as discussed below.

Leave No One Behind and Intersectional Approaches

United Nations programming must ensure that it incorporates and promotes the principles of equality and non-discrimination. These principles are at the core of international human rights law and require that all persons are recognized as equal before the law and their rights guaranteed "*without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status"*. Duty-bearers must strive for achieving substantive equality (as opposed to formal equality) by eliminating both direct and indirect discrimination both in law and in practice with the aim of ensuring those concerned that these laws and policies do not maintain, but rather alleviate, the disadvantages that particular groups' experience due to inequalities linked to their protected characteristics.

The United Nations works towards promoting the principles of equality and non-discrimination through a human rights-based and gender-responsive approach to programming that fulfils the pledge to 'leave no one behind' and incorporates intersectionality.

Leave No One Behind

Leave no one behind (LNOB) is the central, transformative promise of the 2030 Agenda and its SDGs. LNOB is a guiding principle that is grounded in international law and is a political commitment for states signing on to the SDG Agenda. It is complementary to human rights-based approach and focuses on inequalities, including on ways in which inequalities overlap. The UN Sustainable Development Group (UNSDG) has developed a useful guide that sets out the mandate for LNOB and operational steps for adopting this approach - Operationalizing Leaving No One Behind. As noted in the guide:

LNOB deepens focus on the inequalities, including multiple forms of deprivation, disadvantage, and discrimination, and 'reaching the furthest behind first". HRBA brings to the LNOB a rigorous methodology for identifying who is left behind and why, looking at root causes. [p 10]

To further this approach, **evaluations should address the extent to which UN interventions have advocated for the principles of equality and inclusive development**, including gender and disability issues, and have contributed to empowering and addressing the needs of the most disadvantaged and vulnerable populations in a given society. This can include looking at whether the design of the intervention was based on an analysis of who is being left behind and why, if implementation prioritized actively engaging typically excluded groups and the most marginalised, and if the monitoring and reporting process tracked this information.

Common challenges and issues seen in reaching the 'furthest behind first' are knowing who is being left behind and the often-limited availability of data disaggregated by sex, age, and other characteristics such as class, race, location, disability, educational levels, and migratory status. Interventions that support capacity development of duty bearers could be assessed on whether they looked at addressing data gaps. Evaluators can also look at whether engagement processes used in the intervention hindered inclusion and participation; potential question areas being: was there provision of translation/interpretation into languages spoken by ethnic minorities or migrant groups?; did consultation practices account for power dynamics between and within different ages and groups?; were accessible venues and formats used?; were awareness-raising activities sensitive and tailored to the context and target group (for example, for activities targeting younger people, were the learning materials child- and/or youth-friendly)?

Intersectional Approaches

Intersectionality is grounded in international human rights law and, in particular, in the principles of equality and non-discrimination. **An intersectional approach facilitates recognition of the complex ways in which social identities overlap** (for example, gender, race, class, sexual orientation, physical ability, and others) and, in negative scenarios, can create compounding experiences of discrimination and concurrent forms of oppression[13]. This complexity is depicted in the Intersectionality Wheel in figure 2.2[14].

As noted in UNOCT's Gender Mainstreaming Policy (2022), the Office follows an intersectional approach to gender and CT/PCVE which recognizes how gender is interdependent with other identity factors such as nationality, age, belief, sexual orientation, race, ethnicity, geographic origin, ability, socio-economic status, and other identity factors, and combinations thereof. In doing so, the Policy acknowledges that an intersectional approach contributes to gender-responsive and human rights-compliant CT/PCVE programming by supporting a diverse and comprehensive understanding of the

---

[13] See UN Network Guidance Note on Intersectionality, Racial Discrimination and Protection of Minorities.
[14] UNPRPD and UNWOMEN, Intersectionality Resource Guide and Toolkit: an Intersectional Approach to Leave No One Behind (2021).

underlying dynamics that shape the experiences, challenges and impacts of terrorism, violent extremism conducive to terrorism, and related responses, on different subsets of the population.

Applying a perspective that takes intersectionality into account involves acknowledging and paying specific attention to:[15]

- The fact that the available information and data indicate that people affected by intersectional discrimination generally belong to the groups most at risk of being left behind.
- The socio-structural nature of the discrimination, marginalization and exclusion that perpetuate inequality within a society or specific communities and the role that legal, economic, and political frameworks, institutions and socio-cultural norms play in this context.
- The diversity within groups or communities and the need to recognize the non-homogeneous experiences and needs of individuals affected by intersectional discrimination.

Figure 2.2 Intersectionality Wheel

- The experiences, narratives, and agency of individuals and groups facing intersectional discrimination are key to the development of effective policies and programmes that address, redress, and prevent marginalization, discrimination, and inequality.
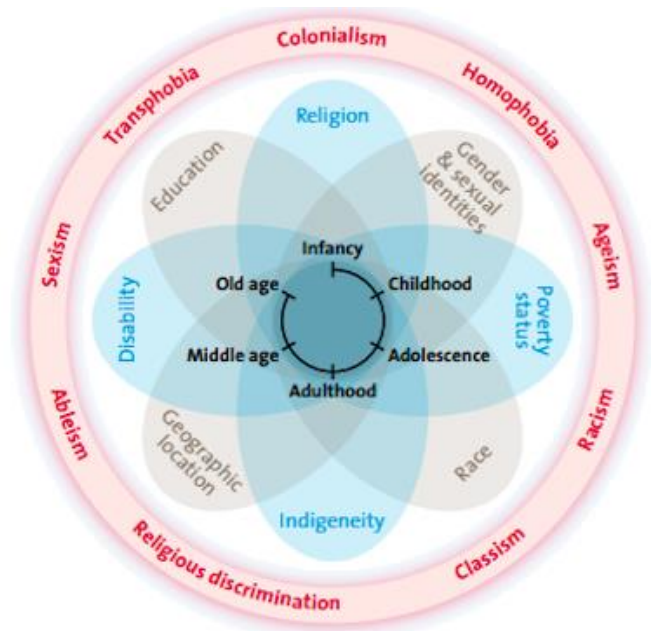
For evaluations, this means paying attention to issues such as:

- What intersectional factors will be considered in the evaluation process?
- What are the problems that may prevent those most marginalised from participating or benefiting from the policy, programme, or action?
- Do programme/project managers and implementors know about the problems of the most marginalised?
- Has the intervention made anything worse?

The following materials provide useful addition information on understanding and operationalizing intersectionality:

- The Intersectionality Resource Guide and Toolkit produced by UNWOMEN and the UN Partnership on the Rights of Persons with Disabilities (UNPRPD)
- United Nations Network on Racial Discrimination and Protection of Minorities, Guidance Note on Intersectionality, Racial Discrimination and Protection of Minorities

---

[15] See Guidance Note on Intersectionality, Racial Discrimination and Protection of Minorities

The meaningful inclusion of young people is also a key commitment in the 2030 Agenda. As part of the United Nations Youth Strategy, Youth 2030, **UN entities are expected to undertake evaluations to guide evidence-based advocacy and action on youth**; this includes providing new evidence of insights on youth left behind. For evaluators, this means ensuring that young people are included as evaluation participants, data is age-disaggregated, youth perspectives are reflected in the analysis and presentation of findings, and that lessons learned include any insights gained on youth-responsive programming. Although not relevant in all cases, it is important for most UNOCT evaluations to be youth-inclusive.

Understanding the Meaningful Participation of Young People

From a rights-based perspective, young women and young men have the right to be informed and consulted and to have their voices taken into account regarding all matters that have a direct or indirect impact on their lives and futures. The realization of their right to participate needs to be guaranteed by duty-bearers and the United Nations.

From an efficiency perspective, ensuring that diverse groups of young people are engaged and have a say in the design, implementation and evaluation of policies and programmes can significantly improve the relevance, legitimacy, sustainability, and impact of projects.

Source: United Nations and Folke Bernadotte Academy, Youth, Peace and Security: A Programming Handbook, 2021.

2.8 UNEG ETHICAL PRINCIPLES FOR EVALUATORS

Evaluation practice often involves ethical dilemmas, with evaluators having to make decisions in complex circumstances. Issues such as power dynamics, different cultural contexts, and security situations can potentially put evaluation stakeholders, including evaluators, at risk or make them feel uncomfortable. To help ensure an ethical lens informs these situations and the day-to-day practice of evaluation, **all those involved in UN evaluations are required to follow** UNEG Ethical Guidelines for Evaluation. This document provides:

- Four guiding ethical principles for those organizing and conducting evaluation activities (see box below).
- Tailored guidelines for entity leaders and governing bodies, evaluation organizers, and evaluation practitioners, including a checklist of ethical issues to consider during each evaluation phase.
- A Pledge of Commitment to Ethical Conduct in Evaluation that all those involved in UNOCT evaluations are required to sign.

The Guidelines discuss a wide range of issues including independence, conflicts of interest, confidentiality, non-discrimination, responsible data management, transparency, evaluator competencies, do no harm, carbon neutrality, and safety.

Good practice calls for the methodology section of both the inception and evaluation reports to include **an explanation of the ethical aspects that have been taken into consideration by the evaluators**. This means it is necessary to go beyond simply stating that UNEG guidance was followed by being explicit about how the relevant issues from the ethical guidelines were addressed. This should include describing how issues such as data security and highly sensitive information, informed consent, confidentiality, and the principle of "do no harm" were addressed. Steps taken to encourage the full participation of all stakeholders should also be specified, which could include information on how differences in power, status, and abilities within group discussions were addressed and what actions were taken to reduce the barriers so that all stakeholder groups could be represented.

Clarity and preciseness about ethical standards are paramount for building trust and encouraging participation in UNOCT evaluations.

Stakeholders, whether they be high-level government officials or community members, may be hesitant to discuss issues that are politically sensitive. Therefore, informed consent documents that include information on the purpose of the evaluation, how data will be stored and used, and the structure and plan for reporting and dissemination are important for ensuring informed participation. Equally critical is for evaluation participants to know that they are free to not answer any question and to end the interview/discussion at any time without repercussions. Depending on the subject of the evaluation, procedures need to be in place for addressing any incidents where violations of human rights are disclosed to the evaluation team.

2.9 JOINT EVALUATIONS

Joint evaluations are defined as "*a joint 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, with the degree of 'jointness', varying from cooperation in the evaluation process, pooling of resources to combined reporting*"[16]. Although it is most common for joint evaluations to be done with other United Nations entities, these can also be undertaken with other development partners, national governments, and/or funding agencies.

Joint evaluations, along with the trend towards joint programming, are increasingly being encouraged to further UN system-wide coherence. There are clear benefits, although also some challenges, in undertaking joint evaluations with key ones being highlighted in table 2.1.

Table 2.1: Advantages and Challenges of Joint Evaluations[17]

Advantages	Potential Challenges
Increased objectivity and legitimacy	Different terminologies and evaluation quality assessment criteria
Fosters collective sense of accountability and ownership of findings and recommendations	Complexity of coordination arrangements
Promotes cost-sharing and cost-efficiencies	Power differentials among partners
Often yields higher quality evaluations	Longer time frames
Can reduce the overall number of evaluations undertaken – thereby reducing transaction costs and administrative demands on host countries	Findings may not address individual agency accountability requirements

A key resource for those involved in these joint processes is the UNEG Resource Pack on Joint Evaluations. It has two components. One is the Joint Evaluations Guidance document, which describes joint evaluation management and governance structures, outlines the steps for implementation, and discusses gender and human rights dimensions in joint evaluations. The Resource Pack also includes a Toolkit that contains reference material, checklists, and examples and good practices of joint evaluation products.

2.10 KEY RESOURCES FOR CONDUCTING EVALUATIONS IN THE UN SYSTEM

The box below provides an easy reference to the main resources discussed in this chapter.

---

[16] UNEG Task Force for Joint Evaluation (TFJE), UNEG Resource Pack on Joint Evaluations, 2014.
[17] Adapted from the UNEG Resource Pack on Joint Evaluations, 2014, p. 21.

- UNEG, 2022, UNEG Guidance on Integrating Disability Inclusion in Evaluations and Reporting on the UNDIS Entity Accountability Framework Evaluation Indicator. This defines how disability inclusion is measured in the United Nations system.

- UNEG, 2020, UNEG Ethical Guidelines for Evaluation. This is a revision of the original document published in 2008.

- UNEG, 2018, Guidance on Evaluating Institutional Gender Mainstreaming. 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. The UNEG Norms and Standards are a foundational document intended for all United Nations evaluation bodies.

- UNEG, 2014, UNEG Guidance Document on Integrating Human Rights and Gender Equality (HRGE) in Evaluations. This is an in- depth handbook designed to serve as a field guide.

- UNEG, 2014, UNEG Resource Pack on Joint Evaluations. This provides practical information for conducting joint evaluations.

- United Nations, 2018 System-Wide Policy on Gender Equality and the Empowerment of Women (SWAP) Evaluation Performance Indicator Technical Guidance. 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. 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.

- UNWOMEN, 2021, Intersectionality Resource Guide and Toolkit. This provides useful guidance to help ensure adequate attention is paid to each of the different stakeholder groups interventions should consider.

- UNSDG, 2022, Operationalizing Leaving No One Behind. This is a useful guide that sets out the mandate for LNOB and operational steps for adopting this approach.

- United Nations Network on Racial Discrimination and Protection of Minorities, Guidance Note on Intersectionality, Racial Discrimination and Protection of Minorities

CHAPTER THREE: UNOCT EVALUATION FUNCTION

This chapter presents the guidelines for how UNOCT evaluations are structured, planned, and implemented. It provides a brief overview of the Office's Evaluation Policy and the Evaluation Compliance Unit (ECU), as well as the different categories of evaluation and the roles and responsibilities for each category. It covers the criteria for determining what interventions require evaluations, as well as evaluation planning and budgeting. The final section discusses some of the challenges in assessing counter-terrorism/preventing and countering violent extremism (CT/PCVE) efforts and suggests a number of evaluation strategies and resources to address these challenges.

Since evaluation is a shared responsibility amongst all of those involved in the planning, implementation, support, and oversight of UNOCT programming, this chapter has relevance for all UNOCT staff and for those who evaluate these interventions.

3.1 ORGANIZATIONAL COMMITMENT TO EVALUATION

UNOCT recognizes that evaluation is a powerful tool for learning, decision-making, and accountability. The Office's commitment to conduct and learn from robust evaluation processes is set out in its Strategic Plan and Results Framework 2022-2025 (SPRF). This document emphasizes that evaluation is considered an essential component of a 'results culture' that ensures the delivery of programmes that are transformative, relevant, timely, and impactful. Evaluation, therefore, contributes to the achievement of the five Strategic Goals of the UNOCT Results Framework 2022-2025.

In respect to Goal 1: *"Foster further unity and collaboration within the United Nations against terrorism"*, evaluation is particularly important for UNOCT's role as Secretariat of the Global Counter-Terrorism Coordination Compact and its associated knowledge management responsibilities. Through the Compact, UNOCT provides the mechanisms and supporting platforms for bringing Member States, UN and Compact entities, and other key stakeholders together to share information, experiences, good practices, expertise, and resources, and to develop innovative partnerships and practical solutions to collectively address a range of counter-terrorism priorities. Evaluations provide important data and lessons learned to inform this role.

"In support of a strengthened results culture in UNOCT, evaluations will help to promote the effective utilization of evidence-based information to inform decision-making at all levels. This will be achieved through the conduct of quality evaluations and timely dissemination of evaluation results and recommendations to inform the design, implementation and review of policies, programmes and projects in UNOCT."

UNOCT Strategic Plan and Results Framework

Evaluation also provides the foundation for UNOCT and UNCCT commitments to promoting human rights and gender responsive CT/PCVE efforts. As highlighted in the SP, *"through the application of a robust monitoring and evaluation framework, UNCCT will make solid contributions to continuous learning and capturing of human rights and gender related results in CT/PCVE which will feed into good practice for developing human rights compliant and gender responsive CT/PCVE interventions"*.

3.2 EVALUATION POLICY

UNOCT's internal approach to evaluation is guided by the UNOCT Evaluation Policy. This policy is based on the UNEG Norms and Standards and guided by the Programme Planning, Budgeting,

Monitoring and Evaluation Directorate (PPBME), ST/SGB/2018/3, as well as ST/AI/2021/3 which specifically addresses evaluation in the Secretariat. The UNODC Evaluation Policy provides the overall framework for the planning and undertaking of evaluations to support accountability and increase transparency, coherence, and efficiency in generating and using evaluative knowledge for organizational learning and effective management for results. More specifically, this policy:

- Sets out the purpose and principles of evaluation.
- Outlines clear roles and responsibilities for evaluation at all levels.
- Presents the UN-wide accepted criteria to be applied to all evaluations.
- Offers direction on the prioritization, planning, and budgeting of evaluations.
- Provides guidance on the different types of evaluations, use of evaluation findings, preparation of management responses, and the establishment of a quality assurance system.
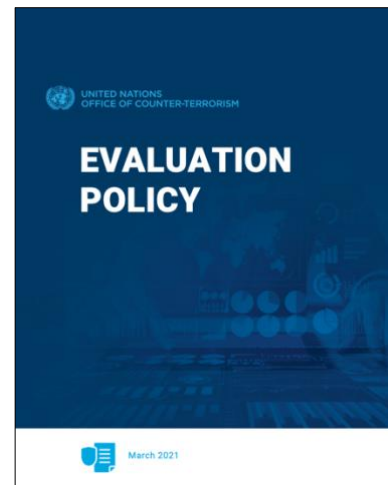
3.3 EVALUATION AND COMPLIANCE UNIT

The Evaluation and Compliance Unit (ECU) in the Office of the Under-Secretary-General for Counter-Terrorism serves as the custodian of evaluation at UNOCT. By being part of the Office of the Under-Secretary-General (OUSG), the ECU is able to maintain independence from the day-to-day operational, management, and decision-making functions of the Office. This helps to ensure credibility of UNOCT evaluations by enabling the ECU to be impartial, objective, and free from undue influence in the conduct and undertaking of evaluations. It also enables the function to coordinate across different sections within the Office, ensuring evaluation is not restricted or otherwise hampered by silos or cross-functional reporting lines.

UNOCT's Evaluation Policy is an important reference document for all staff.

The ECU works to foster a culture of evaluation and to support more strategic and systematic evaluations across the Office. Under the direction of the Chief of OUSG, the ECU helps to elevate organizational capacity for evaluation, provide overall direction and planning for evaluative activity in accordance with the Evaluation Policy, manage and support evaluation processes as appropriate for the different types of evaluation activity, and oversees organizational learning functions from lessons emerging from evaluations. More specific duties are described in the discussion on Roles and Responsibilities below.

3.4 CATEGORIES AND TYPES OF EVALUATIONS

UNOCT has three main categories of evaluations: Centralized Evaluations, Independent Programme/ Project Evaluations, and Internal Evaluations (the latter previously being referred to as Self Evaluations[18]). Although all are expected to be conducted in accordance with the UNOCT Evaluation Policy, and its associated norms, standards, guidance, and templates, including UNEG Norms and Standards, there are different requirements for each category, particularly in respect to evaluation management responsibilities. Table 3.1 provides a brief overview of the main evaluation categories with further explanation provided below.

---

[18] The term 'Internal Evaluation' is used by the Office of Internal Oversight Services (OIOS) and the Business Transformation and Accountability Divisions (BTAD) and will be adopted in the revision of the ST/AI on Evaluation in the Secretariat (planned for 2023).

Table 3.1: Categories of Evaluation within UNOCT

	Centralized Evaluations	In-depth assessments of interventions that are of high strategic interest and relevance across the Office including global and regional programmes, sub-programmes[19], corporate strategies, thematic areas, large-scale or complex initiatives, and joint programmes. Conducted by external evaluators.	Managed by ECU
	Independent Programme / Project Evaluations	Assessments of individual standalone projects or programmes implemented under the framework of global and regional programmes with budgets exceeding USD 1.3 million or with donor requirements for evaluation. Conducted by external evaluators.	Managed by Programme / Project Managers
	Internal Evaluations	Concise assessments of interventions with a modest budget and of short duration. Typically, these are end-of-cycle reviews, but they may also include midterm evaluations of longer-terms projects in which case management arrangements will be similar to IPE above.	Managed by Programme / Project Managers
	Other Evaluative Products	Reviews, meta-synthesis, studies, and other types of assessment processes. They may be scheduled or ad hoc and are typically conducted by external consultants. Responsibility for management will depend on the complexity and scope of the product.	Managed by ECU or Programme / Project Managers

 Centralized Evaluations (CEs) are comprehensive evaluation processes that are managed by the ECU and scheduled on the basis of the annual evaluation plan. They may also include ad hoc assessments requested by the governing body. CEs are conducted by independent external evaluators in consultation with an Evaluation Reference Group (ERG). They require a Management Response and Evaluation Follow-up Plan, and with the exception of strategic evaluations, the final report will be made publicly available on the UNOCT website. There are several types of evaluations that fall into this category.

Global and regional programme evaluations which can be midterm or final evaluations. They address achievements in relation to the programme theory and results chain, and they require a systematic analysis of the programme under review.

Strategic evaluations assess the quality of design, extent of implementation, and results of a corporate or unit-specific strategy or policy.

Thematic evaluations focus on selected topics such as a new approach to PVE, or a cross-cutting theme, or an emerging issue of corporate interest. They entail in-depth analysis of a topic that generally has applicability across organizational units to build organizational knowledge. Thematic evaluations can draw lessons learned from both inside and outside of UNOCT.

---

[19] Sub-programmes will fall within the categories for Centralized Evaluations or Independent Programme/Project Evaluations depending on their level of complexity. This will be determined during the Evaluation Proposal Stage which is discussed in Chapter 4.

Joint evaluations are commissioned and conducted by UNOCT and at least one other organization, typically entities in the United Nations Global Counter-Terrorism Coordination Compact. A joint evaluation may also be initiated by another UN entity. In either case, the degree of 'jointness' may vary depending on the extent to which partners engage in managing the evaluation and pool resources. UNOCT programme/project managers need to set aside funds for the evaluation, even when conducted by multiple partners. Joint evaluations are increasingly encouraged to improve efficiencies and reduce the transaction costs for host countries/implementors who have to accommodate multiple assessments. Joint efforts also help to overcome attribution challenges, provide greater objectivity and legitimacy, gain evaluation knowledge from other agencies, and increase coordination and effectiveness of development assistance. Joint evaluations tend to require greater coordination than other types of evaluations, highlighting the need for advance planning as well as clear management structures and communications systems. The UNEG Guidance for Managing Joint Evaluations provides further information and a useful toolkit.

Ad hoc evaluations may be scheduled at any time to respond to requests that were not included in the annual Evaluation Plan (e.g., those requested by the governing body, the Secretariat[20] or funding partners, or joint evaluations proposed by other organizations).

Independent Project Evaluations (IPEs) are typically lighter-touch assessments than CEs, in part because they are more focused with fewer questions and more concise page lengths for the reports than CEs. IPEs can be midterm or final evaluations used for standalone projects or projects under the framework of global and regional programmes with budgets exceeding USD 1.3 million[21], or with donor requirements for evaluation. They are overseen by the respective Programme/Project Manager and conducted by independent external evaluators in consultation with an Evaluation Reference Group (ERG). IPEs require a Management Response and Evaluation Follow-up Plan, and the final report will be made publicly available on the UNOCT website.

Internal Evaluations[22] (IEs) are typically conducted internally for smaller-scale projects and as mid-term reviews for larger-scale interventions with a timeframe of three or more years. Unless otherwise directed, all projects that do not undergo an IPE require an Internal Evaluation at the end of the project cycle, and those spanning 36 months or more also require an Internal Evaluation at midterm. Internal Evaluations may also be undertaken in combination with requirements for IPEs when more rigorous evaluations are warranted. Programme/Project Managers are responsible for managing Internal Evaluations, although this responsibility is usually delegated to M&E Officers in cases where that position exists. External evaluation consultants can be engaged to provide input at certain phases of the process if needed and if resources are available.

Internal Evaluations will usually have a heavy reliance on existing data that has been collected through project monitoring but should also seek input from the range of other stakeholders involved including rightsholders/beneficiaries. For mid-term evaluations, the focus will be on the quality and relevance of the outputs produced, their usefulness to users, and any initial indications of progress made in achieving the intended results. Final Internal Evaluations will look more at the extent that the project outcomes were achieved. Internal Evaluations can also be conducted at other points during the project lifecycle, including for collecting base-line information during the initial phases of the project

---

[20] As per ST_SGB_2018_3 E which states that "In addition to self-evaluation, the evaluation system shall include the ad hoc in-depth evaluation of selected programme areas or topics conducted internally or externally at the request of intergovernmental bodies or at the initiative of the Secretariat".
[21] As per the UNOCT Evaluation Policy and further guided by internal analysis undertaken to arrive at this threshold.
[22] Note that Internal Evaluations were referred to as 'Self-Evaluations' until 2023 and the latter term may still appear in some documents.

or for rapid assessments that respond to a concisely defined problem. The evaluation report will generally be for internal purposes and therefore a management response will not be required. However, the arising recommendations will need to be incorporated into ongoing and future programming.

3.5 OTHER ASSESSMENT PRODUCTS AND OPPORTUNITIES FOR LEARNING

Meta-Synthesis – An analysis that takes a systematic approach to searching, assessing, extracting, and synthesizing evidence from multiple studies or evaluation reports. Evaluation meta-syntheses are useful for generating knowledge from the processes, results, and lessons of multiple evaluation reports.

Rapid Reviews – An approach designed to quickly and systematically conduct an assessment when time, resources, and/or access to stakeholders is limited. Due to these constraints, it may not be feasible to apply the due process of evaluation. Rapid reviews were commonly conducted during the early stages of the COVID-19 pandemic when information on the suitability of programming and business processes, and how they needed to adapt, was urgently needed.

Studies – Generally refer to discrete investigations to generate answers to specific questions. A study may be part of the monitoring system, the evaluation system, or both. Examples can include situation analyses, literature/desk reviews, mapping exercises, evaluability assessments[23], and case studies. In-depth processes such as baselines studies and impact studies can also be considered as research (depending on the level of methodological rigour) or as evaluations (if the scope of enquiry is broad and encompasses multiple evaluation criteria).

3.6 ROLES AND RESPONSIBILITIES

Under the overall supervision of the Under-Secretary-General (USG), UNOCT has a shared system of roles and responsibilities in performing monitoring and evaluation functions. While monitoring is the responsibility of the Programme Management Unit (PMU) and the respective programme/project managers, evaluations are the responsibility of the ECU, and this Unit will engage other stakeholders based on the type of evaluation.

The following table highlights the main areas of responsibility within the Office for evaluation activities. Specific roles for each type of evaluation are discussed in more detail in Chapter 4- Evaluation Planning and Budgeting, and in the Evaluation Roles and Responsibility Grid in Annex 6.

Table 3.2: Responsibilities for Evaluation within UNOCT

Roles and Responsibilities

---

[23] An evaluability assessment is carried out to determine the readiness of an intervention to be evaluated. This process typically involves reviewing project documents and monitoring system systems to make sure the intended results are measurable, and that the needed data is being collected so that an evaluation would be feasible and useful.

USG	- Provides overall direction, leadership, and management of the evaluation function, including approval of the Evaluation Policy and its revisions, and the annual Evaluation Plan. - Facilitates an enabling environment that values evaluation as part of the organizational culture. - Ensures the evaluation function is accorded sufficient funds and resources, including technical support. - Determines conditions under which the dissemination of particular evaluation findings will be exempted. - Oversees the process of preparing a formal management response to each evaluation (with the exception of Internal Evaluations) including their implementation.
ECU[24](under the direction of Chief/OUSG).	- Develops and updates the Evaluation Policy, harmonizing UNOCT evaluation practices and requirements with UN standards. - Maintains the annual Office Evaluation Plan. - Provides methodological guidance on evaluation activities, including setting standards and maintaining evaluation quality-assurance mechanisms. - Manages the online Evaluation Management and Tracking Tool (EMTT) and coordinates with all programme units to ensure relevant and current data and documents are entered. - Identifies capacity needs and resources to manage and conduct evaluations of programmes/projects and supports the development of relevant expertise and skills through trainings and tools. - Initiates and manages Centralized Evaluations and provides support and technical advice to the Programme/Project Managers for Independent Programme/Project Evaluations and Internal Evaluations. - Synthesizes evaluative knowledge from individual evaluations. - Carries out quality assurance for Centralized Evaluations and Independent Programme / Project Evaluations. - Presents evaluation results and actively disseminates and shares knowledge generated from evaluations; this includes uploading relevant resources to the UNOCT *Connect & Learn* e-learning platform in support of internal knowledge management and information sharing and contributing to the Secretary-General's Data Strategy where appropriate. - Promotes and supports stakeholder buy-in, participation, and support from all levels by engaging Evaluation Reference Groups (ERGs), representative panels, and peer reviews to appraise evaluation plans and reports. - Engages in partnerships with professional evaluation networks, such as UNEG, and supports the harmonization of the evaluation function in the United Nations system.

---

[24] Further details on the roles and responsibilities of the Evaluation Compliance Unit can be found in the UNOCT Evaluation Policy.

PMU/ PKMCB[25]	• Ensures theories of change and logical frameworks appropriate for monitoring and evaluating results of interventions are incorporated into planning and management practices. • Ensures new project documents have adequate budget allocations for evaluation in line with the evaluability criteria established by the Office. • Supports the ECU in implementing Centralized Evaluations; provides support and technical advice to Programme/Project Managers for Independent Programme/Project Evaluations and Internal Evaluations [according to the steps set out in Chapter 4- Evaluation Planning and Budgeting and Annex 6]. • Supports the development of Management Responses for Independent Programme/Project Evaluations. • Oversees the implementation of recommendations of relevant evaluations and coordinates with programme/project managers on the status of implementation. • Promotes the incorporation of evaluation and audit recommendations into ongoing work planning. • Contributes to sharing best practices and lessons learned to improve the quality of UNOCT products and services together with ECU.
Human Rights and Gender Section	• Provides advice, usually as member of the Evaluation Reference Group, to help ensure all evaluations are grounded in international norms and standards related to human rights and gender equality, and reflect issues such as intersectionality and LNOB as appropriate to the subject of the evaluation.
Project/ Programme Managers	• Contributes to CEs; manages IPEs; initiates and manages Internal Evaluations [according to the steps set out in Chapter 4- Evaluation Planning and Budgeting and Annex 6]. • Ensures relevant and current data and documents are entered into the online evaluation tracking tool. • Develops Management Response for each evaluation. • Implements recommendations of relevant evaluations and provides updates on status of implementation. • Incorporates evaluation and audit recommendations into annual work planning. • Shares best practices and lessons learned to improve the quality of UNOCT products and services.
Senior Management (Section Chiefs) and Senior Leadership (Directors)	• Available for consultation and interviews with the Evaluation Team. • Attends presentations of evaluation results. • Reviews the Management Response and ensures follow-up of recommended actions identified in the Management Response. • Ensures the dissemination and utilization of evaluation recommendations to inform workplans and strategies. • Uses evaluation results for evidence-based decision-making.

---

[25] The Policy, Knowledge Management and Coordination Branch (PKMCB) in the case of Strategic Evaluations. The Behavioural Insights Hub (BI Hub) in the case of evaluations conducted under its purview.

M&E Officers	For programmes that have an M&E Officer, their responsibilities for each evaluation should be determined in consultation with the ECU. Potential roles include: • Working with Programme Managers to ensure evaluation planning is embedded into programme/project planning, and to ensure monitoring is established and ongoing. • Collaborating with the ECU in managing the evaluation process. • Being designated as the Evaluation Manager of project evaluations (under the purview of the section/unit). • Conducting quality assurance of evaluation deliverables. • Sharing evaluation results. • Serving as a member of the Evaluation Reference Group (including for evaluations of other UNOCT programmes). • Delivering evaluation capacity-building to Programme Managers and programme staff on an ongoing basis.

3.7 ENGAGEMENT OF STAKEHOLDERS IN EVALUATIONS

In line with UNEG Norms and Standards, additional stakeholders will be engaged to enrich the evaluation processes. Although Evaluation Reference Groups are a requirement for all Centralized Evaluations and IPEs, other groups will be created on a case-by-case basis and their scope of engagement determined by the USG. Examples include Advisory Panels as well as learning groups, peer reviewers, and steering groups. The three most common forms of stakeholder engagement are described here, including ways to involve Member States. It is important for all groups to be inclusive in their membership and in the way they are conducted so that a diversity of voices is heard.

Evaluation Reference Group (ERG) – The ERG is a consultative body representing the most relevant stakeholders in the intervention under review. These individuals and institutions are the natural users of the conclusions and recommendations emerging from the evaluation. Internal members generally include representatives of programme management and other units involved in supporting the intervention, for example, the Human Rights and Gender Section (HRGS) which is a standing member of all ERGs and others who have relevant expertise which will serve to assure the quality of an evaluation. External members generally include other duty bearers (implementing partners, other government partners, and funders) and preferably, also include representation from rights holders (those the intervention is directed towards). Typical tasks include:

• Identifying information needs and contributing to defining the objectives and scope of the evaluation.
• Providing input into the drafting of the ToR.
• Facilitating the evaluation team's access to information relevant to the intervention, as well as to key actors and informants who should participate in data-gathering methods (typically as respondents).
• Monitoring the quality of the process and the reports generated, so as to enrich these with their input and ensure their interests and needs for information about the intervention are addressed.
• Validating preliminary findings and conclusions, as well as providing input into and/or comments on the development of recommendations to ensure they are actionable.
• Sharing the results of the evaluation and supporting the integration of the recommendations and lessons into ongoing and future programming.

Advisory Panels – Advisory groups may be established for the purposes of assessing the quality of work performed and securing the acceptability of the products of the evaluation. These groups or individuals may be people outside of the evaluation process, such as academics or persons from relevant think tanks who have thematic expertise. In-country advisory groups representing a broader set of organizations can be important mechanisms for widening institutional engagement. Typical tasks include:

- Offering advice on the methodological approach.
- Reviewing draft evaluation documents.
- Being on call for referral throughout the evaluation process.

Learning Groups – These are groups formed by Programme Managers to engage their peers (a) for support in the conduct of the evaluation, and (b) in the learning opportunities from both carrying out the process and from the results of the evaluation. Evaluations typically produce recommendations and lessons that are useful to other programmes. The establishment of Learning Groups is not required but is highly encouraged as a way for staff to share insights and knowledge across the Office about conducting evaluations, programme management practices, and thematic expertise.

> Stakeholder participation throughout the evaluation helps to ensure the processes are credible and useful. Advance planning is essential for stakeholder engagement to be meaningful and respectful.

Member States

Member States are key stakeholders in evaluations and can be involved in multiple ways depending on the type of intervention including as counterparts, programme hosts, beneficiaries, and/or funders. Member State representatives are welcome to participate in all phases of the evaluation process and may have different roles and responsibilities including:

- Being part of the Evaluation Reference Group.
- Being available for interviews with the evaluation team.
- Attending the presentation of results of the evaluation.
- Being stakeholders in the follow-up to evaluation recommendations.

3.8 EVALUATION PLANNING, SELECTION, AND BUDGETING

Best practice calls for programmes and projects to undergo a formal evaluation at some point during their lifecycle, and adequate budgets need to be set aside to do this. However, organizational capacity for undertaking multiple evaluations at any one time also has to be considered. It is recognized that all evaluations entail a significant investment of staff time and resources. Therefore, decisions about whether and when an evaluation is necessary and the type of evaluation to be conducted are typically determined by the nature, duration, budget, and complexity of the intervention, as well as the extent to which it is likely to contribute to organizational learning, and commitments to stakeholders, among other factors. This section discusses how evaluations are to be prioritized, planned, and funded.

UNOCT's Evaluation Plan

The ECU in the Office of the Under-Secretary-General for Counter-Terrorism maintains the UNOCT Evaluation Plan[26]. This is a rolling plan whereby new evaluations are added each year and completed evaluations are closed out. The Plan tracks the programmes, projects, topics, and themes for evaluation including the estimated cost and timeframe of each exercise. Interventions are prioritized for an evaluation based on the selection criteria provided below and projected organizational capacity.

---

[26] The ECU also makes the annual Evaluation Plan available to internal audiences as well as the OIOS and BTAD.

Evaluation Selection Criteria

The following assessment criteria are used to determine the type and timing of evaluation needed, including for other interventions (such as strategies, policies, and thematic areas). The more complete description of the evaluation criteria can be found in the Evaluation Policy and in the Selection Criteria Template in Annex 7(a). Programme Managers are encouraged to refer to this information when developing their evaluation proposals during the project/programme design phase.

    a. Significant investment (interventions with a total budget of $1.3 million and above).
    b. Risk associated with the subject including risks related to human rights.
    c. Duration (interventions spanning a period of 36 months and beyond).
    d. Utility and strategic contribution to the United Nations Global Counter-Terrorism Strategy.
    e. Complexity and uncertainty factors of the project.
    f. Level of innovation, and potential to fill knowledge gaps and contribute to organizational learning.
    g. Sustainability of the intervention and/or its intended results.
    h. Formal commitments to stakeholders.
    i. Request from Under-Secretary-General and Senior Management.

Evaluation Proposals

Programme Managers are responsible for developing brief proposals for evaluations they intend to carry out. The proposals set out the case and justification for the evaluation as well as its focus, timeline, budget, and links to any previous evaluations or studies. To be considered for inclusion in the Office Evaluation Plan, proposals for evaluations planned for implementation in Q2 and Q3 need to be submitted to the ECU by the end of March. Proposals for evaluations planned for Q4 and Q1 the next year need to be submitted to the ECU by the end of September. The proposal template is provided in Annex 7(b).

Budgeting for Evaluation

All programme and project budgets need to include direct and indirect evaluation costs. These requirements are specified in table 3.3 and explained here:

- Direct costs of undertaking the required number of evaluations. Guidance on determining the direct costs of conducting evaluations is provided in the next chapter on evaluation planning and budgeting.
- Consultant costs for other M&E-related activities where additional expertise would help ensure adequately robust systems are in place to report on and assess results. Examples include conducting baseline studies and developing monitoring systems.
- Support to centralized, strategic, global, and regional programme evaluations. As part of UNOCT efforts to build its funding for centralized and strategic evaluations, including global and regional programmes, all programmes and projects are required to allocate a percentage of their project evaluation budgets for this purpose to ensure that the ECU can conduct such evaluations effectively. The allotment to the central evaluation support enables the Office to sustain consistent evaluation activity in line with its programmatic growth, recruit necessary ad hoc or permanent evaluation capacity based on need, professionalize the function, and support innovation in evaluation.

Table 3.3: Budgeting for Evaluation by Size of Intervention

Budget components	Project/Programme Budget		
	< $1.3 million	$1.3 - $5 million	Over $5 million

Direct costs of implementing evaluation activities	Minimum US$ 35,000	Minimum US$ 100,000*	Minimum US$ 270,000*
% of overall project/programme budget allocated to support Strategic and Centralized Evaluations of the Office	1%	0.5%	0.3%

* These amounts reflect the cost for one Internal Evaluation and one externally-conducted evaluation.

3.9 SPECIAL CONSIDERATIONS FOR EVALUATING UNOCT-SUPPORTED INTERVENTIONS

Alignment with UNOCT Priority Areas

All evaluations of UNOCT interventions, including those conducted by implementing partners and those conducted jointly with other agencies, are expected to consider the alignment and contribution of the intervention to UNOCT priority areas. These priorities are (a) the strategic goals set out in the UNOCT Strategic Plan and Results Framework (SPRF), and (b) the four pillars of UN Global Counter-Terrorism Strategy which are:

- Addressing the conditions conducive to the spread of terrorism.
- Measures to prevent and combat terrorism.
- Measures to build states' capacity to prevent and combat terrorism and to strengthen the role of the UN system in that regard.
- Measures to ensure respect for human rights for all and the rule of law as the fundamental basis for the fight against terrorism.

The main focus for this aspect of evaluations will be on the outcomes articulated for each strategic goal of the SP. All programme and project documents should include relevant outcomes and indicators in their own results frameworks linked to the SP Strategic Goals. **It is incumbent upon evaluators to consider the extent to which the intervention is appropriately aligned with the overall SP strategic goals and outcomes, and their contribution to at least one outcome.** The associated findings of each programme/project evaluation will be an instrumental data source for the mid-term and final evaluations that are conducted for each UNOCT strategic planning period.

All evaluations of UNOCT and UNOCT-funded interventions are required to:

- Consider the alignment of that intervention with the Office strategic priorities, goals, and outcomes.
- Assess the contribution of that intervention to the achievement of at least one outcome of the SP Results Framework.

UNOCT Evaluation Criteria and Question Requirements

UNOCT evaluations are normally required to address the OECD-DAC criteria[27] of Relevance, Coherence, Effectiveness, Efficiency, Impact and Sustainability. There can be exceptions for very focused evaluation processes, but these should be justified in the evaluation Terms of Reference and the Evaluation Report. Additional criteria to be considered, particularly in the case of evaluations of

---

[27] The evaluation criteria set out by the Organization of Economic Cooperation and Development's Development Assistance Committee (OECD-DAC) are discussed in Chapter Two Evaluation in the UN System.

humanitarian-related interventions[28], include Coverage/Reach, Connectedness, Coordination, Protection and Security, and Partnerships. There are also cases when funding agencies request specific criteria to be considered such as the UN Peacebuilding Fund criteria of Catalytic, Time-Sensitivity, and Risk Tolerance & Innovation[29]. In addition, and without exception, all evaluations are to integrate human rights and gender equality concerns as cross-cutting issues, and **all evaluation reports are to have a specific section on Human Rights, Gender Equality, Disability Inclusion, and Leave No One Behind.** This section should also consider the extent to which youth were adequately engaged and considered in the intervention.

Given the need to consider the intervention's alignment with, and contribution to, UNOCT priorities, **all evaluations are required to include** both of the following (or similarly worded) evaluation questions:

- **To what extent is the intervention relevant to the pillars of the GCTS and the Strategic Goals of the SP?** (usually covered under Relevance).
- **What has been the contribution of the intervention to the pillars of the GCTS and the Strategic Goals of the SP?** (usually covered under Effectiveness).

Assessment of Policy-Related Changes

Policy influencing work is a key driver of Agenda 2030 and the Sustainable Development Goals (SDGs). Much of UNOCT's programming links to policy reform - often developing or strengthening the regulatory environment. Even though the effectiveness of this work can be difficult to establish, UNOCT evaluations are expected to take more than a cursory look at how policy-related work is conducted.

Good practice in policy evaluation calls for a careful analysis and unpacking of the theory of change. The use of theory-based approaches includes examining the assumptions on which the intervention is based as well as the pathways (arrows) in the theory of change to see if there is a clear logic between the activities and the intended results within the context of the intervention. UNEG's Evaluating Policy Influencing Stocktaking Report (2022) provides useful guidance; some of the key takeaways being the importance of:

- **Conducting a thorough contextual analysis** to help understand how change happens in that particular context, and how it impacts people of different genders and other identity factors differently. Determining if the intervention design was well informed in regard to the political, social, and economic contexts helps to answer the question of whether the intervention was technically sound and politically viable, and if it has been able to adapt to changing circumstances.
- **Considering process factors** by looking not just for end results but also at the strategies and approaches employed. Undervaluing the importance of processes can lead to an underestimation of achievements. In this sense, process milestones such as changes in ways of operating, in ways of actors coming together, and in building consensus, are accomplishments that should be acknowledged and learned from.
- **Considering how the intervention is positioned** within the different stages of the policy cycle (agenda setting, policy formulation, policy approval, policy adoption, policy implementation, policy monitoring, and evaluation) and if additional stages need to also be targeted.

---

[28] The Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) has developed guidance on the use of humanitarian-specific criteria of Coverage, Connectedness, Coordination and Protection.
[29] The UN Peacebuilding Fund (PBF) typically requires one or more of the topics of Catalytic, Time-Sensitivity, and Risk Tolerance & Innovation to be considered separately or integrated into the questions of other evaluation criteria.

Evaluators should also consider whether the type of intervention is the best mode of delivery.

Policy influencing work in UNOCT is often focused on capacity-building. Capacity-building initiatives related to CT/PVE are guided by assessments made by the Counter-Terrorism Committee Executive Directorate (CTED) on behalf of the Counter-Terrorism Committee and Security Council resolutions. It is prudent for capacity needs to be clearly identified and then supported with clearly targeted programming in a balanced manner across all four pillars. Given the many players in CT/PVE interventions and the fluidity of the contexts in which they operate, evaluations should be structured to account for UNOCT's specific contribution to outcomes achieved. Contribution analysis, described at the end of this chapter, can be useful for assessing the links between the programme activities and results, and ultimately to help determine if capacity-building is the best mode of delivery in a given context.

Consideration of International Conventions and Instruments

UNOCT evaluators should also be familiar with the key international conventions and instruments relevant to the respective thematic areas addressed by the intervention under review.

TERRORISM

- UN Global Counter-Terrorism Strategy and its successive reviews.
- The 19 Legal Instruments related to the Prevention and Suppression of International Terrorist Acts.
- Model Legislative Provisions to Support the Needs and Protect the Rights of Victims of Terrorism.
- Security Council resolution S/RES/2242 (2015) establishing the links between the Women, Peace, and Security Agenda and CT/PCVE.

CRIMINAL JUSTICE

- UN Strategic Guidance Framework for International Policing (SGF).
- Compendium of United Nations Standards and Norms in Crime Prevention and Criminal Justice.
- International Covenant on Civil and Political Rights.
- The 2021 Kyoto Declaration on Advancing Crime Prevention, Criminal Justice and the Rule of Law: Towards the Achievement of the 2030 Agenda for Sustainable Development.

A discussion of conventions and other resources related to the cross-cutting issues of human rights, gender, youth, and disability can be found in Section 2.7 of this handbook.

Unique aspects of evaluating CT/PCVE

Most evaluation processes involve challenges that need to be mitigated, to the extent feasible – such as missing baseline data, scant theories of change, travel restrictions, and language barriers. However, evaluations of interventions aimed at counter-terrorism/preventing and countering violent extremism (CT/PCVE) frequently have to also deal with higher levels of complexity.

CT/PCVE faces challenges inherent in any new field of investigation, including lack of agreement on terminology, difficulty in finding experienced evaluators with sufficient subject-matter expertise, and differences in how CT/PCVE interventions are packaged to beneficiaries and Member States. Given the dynamic and complex environments in which UNOCT interventions operate, there can also be uncertainty about what outcomes should be expected, how to measure change, and how to deal with issues of attribution. CT/PCVE also involves multi-dimensional issues that can be difficult to tackle in a linear fashion. As such, theories of change will have multiple pathways that can be much more

complicated to determine and to assess than programme theories, for example, for agriculture and education-related interventions.

Furthermore, the thematic areas UNOCT engages in are often sensitive, involving aspects of national security and criminal justice systems. This can mean that data is not readily available to evaluators, that some stakeholders may be reluctant or unable to share their experiences, and that project sites may be difficult to access. These factors and other risks need to be carefully thought through when planning the evaluation, and especially for field missions conducted by the evaluation team.

> Evaluators need to be aware that stakeholders may have different understandings of how CT/PCVE terms are defined. It is important to get agreement on the meaning of key terminology and to provide those agreed definitions in the evaluation report.

CT/PCVE interventions may also involve marginalised groups and people in vulnerable situations where issues of human rights, including freedom of expression, and protection of minors, non-discrimination and the rights of persons belonging to racial, ethnic, religious, and other minorities, etc. require special attention in evaluations. Such issues **elevate the need for human rights-based, gender responsive and disability inclusive evaluation processes with well-articulated ethical considerations**.

Strategies for conducting evaluations to help to address such complexities are discussed in Chapter 8 - Practical Guide to Good Evaluation Practice.

UNOCT International Hub on Behavioural Insights to Counter Terrorism

The UNOCT International Hub on Behavioural Insights to Counter Terrorism (BI Hub) can provide assistance to programme designers, practitioners, and evaluators. Its three core functions are:

- Conducting and advancing research in behavioural sciences to better understand the drivers and factors contributing to radicalization leading to violent extremism and terrorism.
- Providing capacity-building assistance to Member States, regional organizations, and civil society partners to develop and implement programmes, projects, and initiatives that integrate behavioural insights to counter terrorism.
- Promoting communication, outreach, and partnerships to share knowledge, expertise, experiences, and lessons learned on behaviourally-informed counter-terrorism interventions.

The BI Hub is developing an M&E Database which will have a synopsis of tools, methods, and indicators from peer-reviewed research that are specific to researching and evaluating CT/PVE interventions.

PVE Practitioners' Perspectives on Evaluation

A report by UNESCO[30] captures the experiences of 57 professionals involved in PVE efforts across the globe including their experiences of evaluation of these programmes. Their perspectives provide interesting insights into practitioners' ambivalence about evaluations as well as how evaluations can be more useful for improving PVE practice. Among the report's recommendations are the importance of the following factors:

- Building evaluation into the initial programme design and budget.
- Balancing the summative aspects of evaluation (what worked?) with the formative aspects (how could this work better?).

---

[30] Clément, P., P. Madriaza and D. Morin, "Constraints and opportunities in evaluating programs for prevention of violent extremism: how the practitioners see it", UNESCO Chair in Prevention of Radicalization and Violent Extremism (UNESCO-PREV Chair): Sherbrooke University, 2021.

- Developing qualitative indicators with the programme staff to help ensure these reflect local conditions and contexts.

- Including an assessment of the organization's resources (training, talent retention, obstacles encountered).

- Increasing the focus of impact evaluations on the quality of the relationships built with the community, and on the intervention context.

- Ensuring evaluation processes are culturally sensitive.

- Conducting cross-evaluations – defined as *"self-evaluations by practitioners to capture data from the field, internal evaluations within the organization to measure the program's effectiveness, and external evaluations to ensure that all organizations are treated equally"*.

3.10 ADDITIONAL USEFUL RESOURCES

More in-depth information specific to evaluating CT/PCVE interventions can be found in the following publications:

- United Nations Development Programme (UNDP) and International Alert, Improving the Impact of Preventing Violent Extremism Programming: a toolkit for design, monitoring and evaluation, (Oslo, 2018)
- United Nations Office on Drugs and Crime, UNODC Toolkit for Evaluating Interventions on Preventing and Countering Crime and Terrorism, (Vienna, 2021)
- Global Counterterrorism Forum, Gender and Preventing and Countering Violent Extremism Policy Toolkit (2022), Ch. 4 Monitoring and Evaluation

Lessons from evaluating counter-terrorism related programming can be found in the following meta-synthesis report. It was commissioned by the Resource Mobilization, Monitoring and Evaluation Working Group of the United Nations Global Counter-Terrorism Coordination Compact through its Sub-Group on Evaluation and provides collective results of 118 evaluations and oversight reports from 18 Compact entities on their implementation of the GCTS.

- United Nations, Learn Better, Together: Independent Meta-Synthesis Under the Global Counter-Terrorism Strategy, (Vienna, 2021)
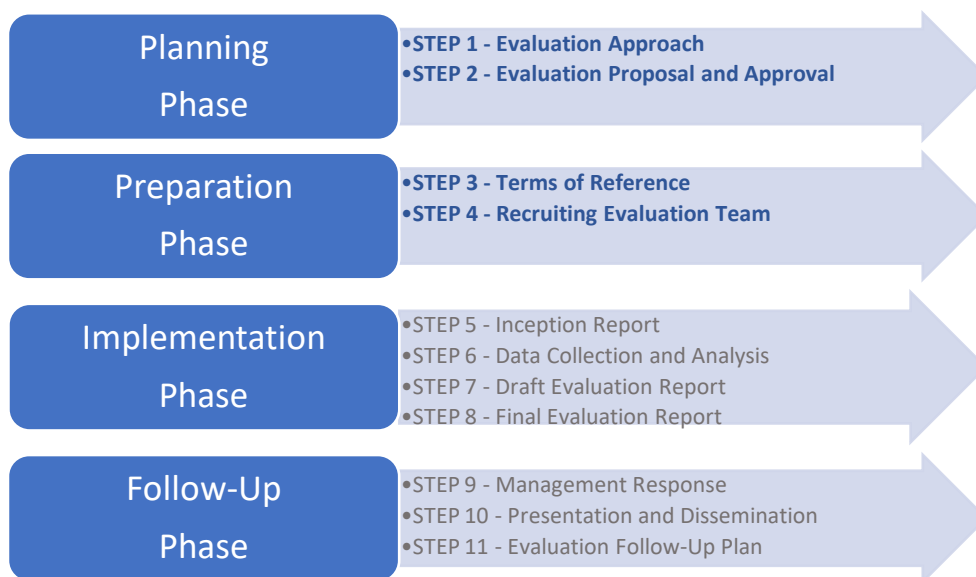
Further information on counter-terrorism, along with guidance and tools for different programme areas, is available on the Publications page of the UNOCT website.

CHAPTER FOUR: PLANNING, COSTING AND PREPARING FOR EVALUATIONS

This chapter is applicable to all programmes and projects. It covers the planning and preparatory work that needs to be undertaken before an evaluation is carried out. These are the initial phases of the evaluation process as illustrated in **figure 4.1**. The Implementation and Follow-Up Phases are covered in Chapters 5 and 6 respectively.

The main audience for this chapter is those who plan and manage evaluation processes. Depending on the category of evaluation, this may be the responsibility of the Programme Manager, Project Manager, or the Evaluation Compliance Unit (ECU) in the Office of the Under-Secretary-General for Counter-Terrorism. For simplicity, the term 'Evaluation Manager' is used in the remainder of this handbook to refer to the person who has designated primary responsibility for overseeing the preparatory and implementation phases of the evaluation process.

Figure 4.1: Phases and Steps of the Evaluation Process

Planning Phase	•STEP 1 - Evaluation Approach •STEP 2 - Evaluation Proposal and Approval
Preparation Phase	•STEP 3 - Terms of Reference •STEP 4 - Recruiting Evaluation Team
Implementation Phase	•STEP 5 - Inception Report •STEP 6 - Data Collection and Analysis •STEP 7 - Draft Evaluation Report •STEP 8 - Final Evaluation Report
Follow-Up Phase	•STEP 9 - Management Response •STEP 10 - Presentation and Dissemination •STEP 11 - Evaluation Follow-Up Plan

PLANNING PHASE

STEP 1 – DETERMINING THE EVALUATION APPROACH

Evaluation planning starts at the programme/project formulation stage and is reflected in the approved Programme/Project Document (PD)[31]. This conceptualization of the evaluation requirements considers the categories and types of evaluation needed, their timing, budget, and how they will be carried out. Most programmes/projects will have two evaluations during their lifecycle (a mid-term evaluation and a final evaluation), with the exception of those with smaller budgets and timeframes. In some cases, small projects may be exempt from having a formal evaluation. Programme/Project Managers should refer to the Evaluation Selection Criteria in Annex 7(a) to confirm whether an evaluation is required.

> It is very important to allocate adequate funding in the programme/project budget to cover all required evaluation activities throughout the intervention's lifecycle. Guidance for preparing evaluation budgets is provided later in this chapter.

Box 4.1 provides an outline for what the **Evaluation Plan** is to include, both in the PD and in any subsequent PD Revision Documents. Note that the Evaluation Plan is to be discussed with and cleared by the Evaluation Compliance Unit (ECU) in the process of preparing the PD. This helps to ensure that all evaluation information is clearly identified during the formulation phase of all interventions. The PD will then need to be approved by Section Chiefs prior to its submission to the Programme Review Board (PRB) for review and recommendation for USG's approval of the programme/project.

Box 4.1: Outline for the Evaluation Plan

Required information for the Evaluation Plan in the Programme/Project Document (PD) and PD Revision

---

[31] In the formulation stage of the programme/project management cycle, Programme Managers develop the full programme/project documentation (ProDoc), which includes the PD and its annexes, as per SOP on Programme Management.

1. **Category(s) of evaluation anticipated** – Centralized Evaluation (CE), Independent Programme/Project Evaluation (IPE), Internal Evaluation (IE).

2. **Type of evaluation(s)** – Formative, Mid-term, Final, Ex-post, Other such as Ad hoc.

3. **Purpose of evaluation(s)** – Explain the reasons for the evaluation and how results will be used.

4. **Timing of evaluation(s)** – Proposed months and year for undertaking each evaluation; whether rescheduling of evaluation is needed due to project extension.

5. **Relationship to relevant past and planned evaluations** – And to relevant evaluations at other levels (global, regional and national programmes/projects, centralized evaluations, other).

6. **Evaluation budget** – Specify reserved amount and if further funds are needed.

7. **Stakeholder involvement** – Expected involvement of other UNOCT sections/units, partners, or other groups in the conduct of each evaluation, such as being part of an Evaluation Reference Group or having an advisory role.

8. **Evaluation consultants** – Expected number of consultants and type of expertise required (if applicable).

9. **Methodological approach** (optional) – If known, identify the expected approach such as mixed-methods, quasi-experimental design, etc. in consultation with the Evaluation and Compliance Unit.

STEP 2 – DEVELOPMENT OF THE EVALUATION PROPOSAL

Approximately eight months to one year before the evaluation is scheduled to begin, the Programme/Project Manager should prepare the **Evaluation Proposal**. At this point, the original evaluation plan from the PD should be revisited to take into account any issues related to changes in the context, environment, organizational priorities, or other conditions from when the initial planning was done. This is also the time when an evaluability assessment[32] may be conducted if needed (this may only be applicable to some interventions that are undergoing a Centralized Evaluation).

The Evaluation Proposal is an updated version of the evaluation plan that was submitted in the PD. However, it also includes more specific information on the definitive timeline for the evaluation process, the budget, and the plan for stakeholder engagement – these issues are discussed below. The Evaluation Proposal should also suggest the methodological approach to be followed that would be suitable for the evaluation purpose. General guidance and resources on different evaluation designs, methods, and tools can be found in Chapter 8 - Guide to Good Evaluation Practice. Although the Evaluation Proposal is usually developed internally, with support being available from the ECU, there may be cases where it is appropriate to engage external expertise.

The template for the Evaluation Proposal is provided in Annex 7(b). The completed Evaluation Proposal document will be reviewed by the ECU prior to the initiation of the evaluation. Evaluation Proposals for CEs and IPEs also require final approval by the USG.

---

[32] An evaluability assessment is carried out to determine the readiness of an intervention to be evaluated. This process typically involves reviewing project documents and monitoring systems to make sure the intended results are measurable, and that needed data is being collected so that an evaluation would be feasible and useful. Evaluation assessments are also an opportunity to collaboratively plan and design the evaluation process, including the methods to use and stakeholders to involve.

Scheduling Evaluations

The type of evaluation will determine the general timeframe.

- **Formative evaluations** are typically scheduled in the early phases of the intervention so that the learning can be incorporated into ongoing planning and any required re-design processes. These are also referred to as ex ante evaluations which are done before or at the early stages of an intervention, an example being evaluations that collect baseline information. Formative evaluations are only required for more complex programmes and their need should be discussed with the ECU.
- **Mid-term evaluations** should be scheduled as close to the middle of the intervention life cycle as feasible. These have both formative and summative purposes. As such, they serve an accountability function and the insights gained from assessing performance to date can be used to guide the remainder of the intervention and to adjust direction if needed.
- **Final (or summative) evaluations** are scheduled near the end or after the completion of the intervention lifecycle. These are undertaken for accountability purposes as well as to inform the development of future programming.
- **Ex-post (or post hoc) evaluations** are done some time after the intervention has closed with the aim of assessing the sustainability of results and impacts. Ex-post evaluations are only occasionally conducted at UNOCT.
- **Ad hoc evaluations** are those designed for a specific purpose and can be commissioned at any time.

> Most UNOCT programmes and projects will undergo mid-term and final evaluations. The size and complexity of the intervention will determine whether these are conducted as Centralized Evaluations, Independent Programme/Project Evaluations or Internal Evaluations. See Chapter 3 – UNOCT Evaluation Function for further guidance.

Table 4.2: Illustrative Examples of Evaluation Requirements

Illustrative Examples of Evaluation Requirements		
Evaluation requirements depend on the type of intervention (whether global, regional or national programmes or projects, strategy, etc.), its duration and budget, as well as contribution agreements. Larger projects and programmes will, at a minimum, undergo a mid-term and final evaluation, with each potentially being conducted under a different category. For example, the mid-term assessment might be conducted as an Internal Evaluation and the final assessment might be conducted as an Independent Programme/Project Evaluation. Some interventions will not require an evaluation. The final determination will be made at the development of the Evaluation Plan during the formulation stage (when the PD is being developed) in consultation with the ECU. The Evaluation Selection Criteria is used to determine any exemptions from evaluation processes. The table below provides examples of the kinds of evaluation that may be required for each type of intervention.		
	Evaluation Category[33]	

---

[33] Further information on the different categories of UNOCT evaluations can be found in Chapter 3.

Types and Examples of Interventions	Internal Evaluation Self-assessments managed by PM	Independent Programme/Project Evaluation External assessments managed by PM	Centralized Evaluation In-depth, external assessments managed by ECU
Global Programmes e.g., Countering Terrorist Travel Programme			Mid-term evaluation and Final evaluation
Corporate Strategies e.g., Strategic Plan	Formative evaluation		Final evaluation
Large Programmes /Projects e.g., Global Programme on Countering Terrorist Use of Weapons	Mid-term evaluation	Final evaluation	
Small Programmes /Projects e.g., Building capacity in the field of eliminating the supply of weapons to terrorists in Africa, Phase I	Final evaluation		
Large Programmes/Projects are those with a timespan of 3+ years and those with budgets of USD 1.3 million or more. **Small Programmes/Projects** are those with a timespan of less than 3 years and with budgets of less than USD 1.3 million.			

Global Programmes that are ongoing (without specified end dates) are required to have periodic evaluations at least every three years. The category and type of evaluation is determined in consultation with the ECU.

Within these parameters, there are other factors to be considered in determining the amount of time needed to complete an evaluation. One of the most important is for the data collection phase to be held when key stakeholders are more likely to be available. This means avoiding periods when people are more likely to be on leave or holidays (including major cultural and religious celebrations in countries where data collection is planned to be carried out), during mid-year or end-of-year reporting periods when UN and partner staff may have very limited time availability, and when seasonal weather events may affect travel, particularly in cases where field missions are planned for regions outside the main cities.

Table 4.3 provides an approximate timeline for the different categories of evaluations processes to be completed. It shows the duration of Centralized Evaluations (CE) and Independent Programme/Project Evaluation (IPE) ranging from 8–15 months, with Internal Evaluations taking from 6–11 months. Of course, there are situations where each stage can be compressed or extended.

Table 4.3: Indicative Timeframe by Type of Evaluation

Task	Centralized and Independent Programme/Project Evaluations	Internal Evaluations
Evaluation Proposal (confirming budget, schedule, stakeholder participation, final approval)	1-4 weeks (depending on size and scope of intervention)	1-3 weeks
Development of ToR	3-6 weeks (+ 2 weeks for review by ERG)	1-2 weeks (+2 weeks for clearance by ECU)
Recruitment of evaluation consultants	8-12 weeks if hiring individual team members and 10-16 weeks if issuing tender for evaluation firm	8-12 weeks for hiring individual team members
Onboarding evaluation team (orientation and inception meeting, sharing key documents)	1 week	1 week
Development and approval of Inception Report	3-6 weeks (once contract is signed)	3-6 weeks (once contract is signed, if using external evaluators)
Data collection, including field missions	6-16 weeks (depending on size and scope of intervention)	4-8 weeks
Analysis and drafting (data analysis, validation, development of draft reports, feedback)	6-10 weeks (depending on number of drafts and extent/timing of feedback from ERG and other stakeholders)	4-8 weeks
Final report and presentation	2-3 weeks	1-2 weeks
Follow-up activities (management response, dissemination)	1-4 weeks	1-3 weeks
Approximate total timeframe	30 – 60 weeks	24 – 45 weeks

Additional time should be factored in for evaluations that are undertaken jointly with others, for example with governments and other UN entities. Joint evaluations can bring a number of benefits, including shared ownership of the results. However, the planning and approval times can be longer than for evaluations commissioned by a single organization as there has to be agreement on issues such as evaluation management, cost-sharing, the methodological approach, reporting formats, quality standards, etc. More information can be found in Chapter 2 – Evaluation in the UN System.

It is critical to plan ahead – evaluations should be initiated 6 to 12 months before data collection is envisioned to give time, in particular, for hiring and contracting evaluation consultants. A longer lead time will be necessary when engaging an evaluation firm through a bidding process.

Budgeting for Direct Costs of Evaluations

All interventions undergoing evaluation need to include the direct costs of any evaluation process[34]. The total amount for evaluations within UNOCT can range from $35,000 for small projects (that involve two evaluators) to over $270,000 for large, complex programmes (which typically are conducted by evaluation firms), with Internal Evaluations being considerably less expensive, but requiring some internal evaluation expertise. The costs will depend on the scope and methodology of the evaluation (such as whether large scale data collection is envisioned), the duration of the intervention, geographic and thematic coverage, etc. Table 4.4 shows the main considerations for developing a budget for a single evaluation.

Table 4.4: Factors to consider in developing an evaluation budget

	Potential budget items
Evaluator Fees UNOCT requires a minimum of two evaluators (one evaluation expert and one subject matter expert), with the exception of internal evaluations. The engagement of national evaluators and junior evaluators as team members is encouraged. More complex evaluations may need to include additional consultant fees for external quality assurance[35].	• Number of consultants and grade level • Daily rates[36] • Number of working days (for orientation; inception phase; travel time, enumerator training, data collection and analysis, quality assurance and validation processes; draft report and expected number of revisions; final report, evaluation brief, and presentation of results) • Local data collectors (if applicable)
Travel Consider the need for in-person versus remote data collection processes. Consider minimizing international travel by increased use of local consultants for in-country data collection.	• International and in-country flights and visas • Other local travel • Number of working days when daily subsistence allowance (per diem) is needed • Participation of UNOCT staff in field missions • Security costs (if applicable)
Data collection and analysis Consider the costs of supporting effective data collection and analysis. Data collection could include stakeholder participation, such as interviews and workshops, which can require logistical costs.	• Facility rental fees and food/refreshments • Per diem for participating stakeholders • Interpreters for meetings, interviews, focus groups in other languages • Translation of materials and tools into other languages (e.g., information forms, surveys) • Equipment and supplies • Software subscriptions (e.g., survey platforms) • Recording and transcription of qualitative data collection (e.g., interviews, focus groups) for analysis

---

[34] In addition to the direct costs of evaluations, each programme/project must also allocate a percent of the total budget to support the undertaking of centralized strategic, global, and regional evaluations to which the project/programme is aligned. Further explanation is provided in Chapter 3 - UNOCT Evaluation Function.
[35] Further information can be found in Chapter 7 - Quality Assurance and Assessment.
[36] See Secretary-General's bulletin ST/AI/2013/4, Annex 3. The ECU can also provide advice in determining daily rates.

Evaluation Report If the final report needs to be professionally presented, consider whether this would be done internally or externally. Also consider whether there may be costs involved in sharing the report and lessons learned with stakeholders.	• Number of revisions to draft report envisioned. • Graphic design, editing and printing • Translation of reporting into other languages • Workshop or webinar to share results • Production of knowledge products in addition to main Evaluation Report (e.g., pamphlets, policy briefs etc.)
Contingency Since there will usually be a time lapse between initial planning and the start of the evaluation, it is suggested that approximately 5% be budgeted for unknown expenses.	• Contingency

Note that when a firm is engaged to undertake an evaluation, most of the costs (evaluator fees, air fare, and per diems) are managed by the firm. When UNOCT directly hires individuals or groups of individuals to carry out the evaluation, the costs are typically managed by UNOCT. Guidance can be found in the UN Procurement Manual.

In cases where evaluation costs were not included in the PD and the programme/project budget or where the allocated amount is insufficient, it will be necessary for the Programme Manager to get Programme Review Board (PRB) recommendation and USG approval for budget reallocation.

Stakeholder Engagement

It is good practice to engage in the evaluation process the range of stakeholders involved in the intervention. This is particularly important for national counterparts where effective engagement and communication is critical at any time and even more so when evaluating a topic that may be politically sensitive and involve some risks.

Engagement can take two forms. The most common is when stakeholders provide data to the evaluators through participation in interviews, group discussions, and surveys. The other is the engagement of stakeholders in the design and implementation of the evaluation. More 'participatory evaluation' practices entail the involvement of stakeholders in a range of processes from developing the ToRs to selecting evaluators, defining questions, developing methods, collecting data, validating findings, drawing conclusions, contributing to the development of recommendations, and sharing the final results.

Participatory evaluation practices can be considered on a spectrum. Highly participatory approaches[37] can engage rights holders (beneficiaries) and duty-bearers (such as implementing partners and programme staff) in all stages of the evaluation. At the other end of the spectrum is the engagement of a fewer number of select stakeholders (usually duty-bearers) in an Evaluation

---

[37] This can be exemplified in human rights-based approaches (HRBA) to evaluation that not only involve highly participatory processes but also require examining the capacity of duty-bearers to meet their human rights obligations and encouraging right holders to claim their rights. Further information on HRBA can be found in Chapter 2 Evaluation and the UN System.

Reference Group (ERG) that provides feedback on the evaluation products such as the draft evaluation report and its recommendations.

Stakeholder engagement can have many benefits. It typically contributes to a stronger sense of ownership of the intellectual outcome of the process, increasing the likelihood that stakeholders will be more invested in the intervention and act on the recommendations. Higher levels of engagement can also garner a richer set of data when participants can see how their input feeds into programme improvements and become more trusting of the process. Broader participation even helps to build technical knowledge of the subject matter and knowledge about evaluation practice, thereby serving a capacity building function. There are also some drawbacks to participatory processes, most notably the increased time investments and the potential to be drawn off course by divergent viewpoints or agendas. Evaluation Managers will need to seek a balance between different attributes but are encouraged to make the process as participatory as feasible also bearing in mind that a participatory approach is part and parcel of human rights and gender-responsive evaluations. A useful resource for more information on stakeholder engagement, particularly for end users of evaluations, is Michael Quinn Patton's *Utilization-Focused Evaluation: 4th edition* (Sage Publications, 2008).

> Ideally, programmes will have a high level of engagement of stakeholders throughout their lifecycle. If that is not the case, evaluations can provide a good opportunity to establish stronger connections.

PREPARATION PHASE

Once the evaluation proposal has been approved by the USG, it is time to develop the terms of reference and engage the evaluation team.

STEP 3 – DEVELOPING THE TERMS OF REFERENCE

The Terms of Reference (ToR) sets out the specifications of the assignment for all stakeholders and for the evaluation team. The ToR describes **what** is being examined (evaluation object and scope), **why** the evaluation is being undertaken (purpose and objectives), and **how** it will be accomplished (design and methods). It also outlines **when** the various phases of the evaluation will take place (timeframe and deliverables), **who** will be involved (as stakeholders and users) and their roles, and the **use** of the evaluation once it is completed. Finally, the document should refer to the **quality standards** UNOCT evaluations are expected to uphold.

In cases where more than one evaluation consultant is being engaged on an individual basis it will be necessary to have multiple versions of the ToR. This will require an overall ToR for the whole process and supplementary ToRs for each member of the evaluation team that clearly specify their respective roles, responsibilities, and the duration of their contract. One ToR is normally sufficient when an evaluation firm is being contracted.

The ToR needs to be very clear and agreed upon by key stakeholders. An evaluation usually cannot address all components of the intervention, therefore developing the ToR involves strategic choices about the specific focus, parameters, questions, and outputs, given available resources.

The ToR will also suggest the methodologies to be used. However, it should provide enough flexibility for the evaluation team to determine what approaches and tools will be most appropriate and to fine-tune the evaluation questions in consultation with the evaluation manager and the ERG. It is important that any such changes from the ToR be agreed and reflected in the inception and evaluation reports.

The initial draft ToR should be developed by the Evaluation Manager with input from the ECU, those with expertise on cross-cutting issues (human rights,

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

gender, and disability inclusion), funding partner(s), the ERG, and others whose input is appropriate. Sharing the preliminary ToR is an important way to make use of collective knowledge, to establish a shared understanding and ownership of the evaluation, and to clarify expectations about the roles and responsibilities of each party. The final draft should also be shared with the same stakeholders before being cleared and uploaded to the Evaluation Management and Tracking Tool (EMTT).

> The ToRs for Internal Evaluations should be tailored to the needs of that situation. For example, if the report is for internal use only then the context and background section may be briefer, and there may be fewer evaluation questions. If consultants are only engaged for pieces of the evaluation, then their scope of work should be well defined so that is clear where their contribution will fit into the full evaluation process.

A general ToR template that can be adapted and used for all categories of UNOCT evaluations is provided in Annex 7(c). It has instructions for each section of the ToR as well as required and suggested text where relevant. It also includes more specific ToRs for evaluation specialists and for subject matter specialists.

STEP 4 – RECRUITING AND BRIEFING THE EVALUATION TEAM

The selection of the evaluators is a crucial part of ensuring a quality and credible evaluation. To meet this standard, UNOCT relies on independent external evaluators who have extensive evaluation and subject matter expertise, amongst other qualifications and skill sets.

Team Composition: Centralized and Independent Programme/Project Evaluations require a minimum of two evaluators. Typically, the team is led by an evaluation expert and supported by one or more subject matter experts although, ideally, all members should have experience in both areas. Depending on the size and complexity of the intervention being evaluated, additional inter-disciplinary national and international team members with evaluation expertise may be added. Key considerations for recruiting teams include the following:

- Proven expertise in conducting evaluations (preferably with UN entities), managing evaluation teams, and report writing.

- Technical knowledge and experience in UNOCT thematic areas, and particularly in the area of focus of the evaluation, including behavioural sciences.
- Technical knowledge and experience in human rights and gender-responsive evaluation processes, and in other cross-cutting areas (disability inclusion, youth inclusion, leave no one behind, and intersectionality).
- Knowledge of national/regional contexts and situations.
- Strong English language writing skills as well as fluency in other languages relevant to areas where the evaluation is being conducted.
- Demonstrated experience in delivering evaluation reports that meet quality standards, including for good presentation (incorporating useful visual aids and other design elements, succinct and accessible writing style, etc.).
- Having a gender-balanced team with members from the countries/regions involved.

Internal Evaluations (IEs) do not always require the engagement of evaluation consultants. If the internal capacity exists, IEs can be carried out by staff and other stakeholders. In other cases, IEs may require some or full support from one or more external evaluators.

Selection and Contracting Processes: There are two ways in which potential evaluators can be identified. One is through a competitive process in accordance with UN rules and regulations for procurement. This is the required process for contracting firms which can provide an individual evaluator or a team. Such opportunities are typically posted in Inspira or on the United Nations Global Marketplace (UNGM) Tender Alert Service (TAS). Individual evaluators can also be identified through the resources list in the box below.

Resources for identifying potential evaluators:

- UNOCT Evaluation and Compliance Unit (ECU) for a list of potential consultants.
- International Hub on Behavioural Insights for subject matter experts such as behavioural scientists.
- Department of Management Strategy, Policy, and Compliance (DMSPC) Evaluation Consultant Resource (ECR)[38].

The selection process needs to be transparent and will usually be done with the assistance of the ECU. Selection also needs to follow the official UN recruitment process. This generally includes submitting proposals that show how they will meet the requirements of the ToR (in the case of tender processes), interviews, reference checks to verify candidates' competencies as evaluators, and a review of relevant samples of work. The identified candidates will need to be approved by the Strategic Planning and Programme Support Section (SPPSS) prior to their engagement.

It is advisable to begin the selection process early as well qualified evaluators are often in high demand and may have limited availability.

Conflict of Interest: Evaluators need to be independent and free from any potential conflict of interest so that they are able to conduct an impartial assessment. Conflict of interest can be due to:

- Having a vested interest in, or past involvement with, the intervention (for example, having had a role in its design, implementation, or decision-making).

---

[38] https://unitednations.sharepoint.com/sites/DMSPC-BTA_COMMS/SitePages/ECR2.aspx

- Potential future involvement in an intervention (for example, being engaged to re-design or implement the next phase).

As per the United Nations Evaluation Group (UNEG) UNEG Ethical Guidelines for Evaluation, 2020 *(p. 20)*, evaluators are required to disclose any potential conflicts of interest. In cases where there is uncertainty about a potential evaluator's suitability for any assignment, the Evaluation Manager should seek advice from the ECU. Should a conflict of interest arise during the course of the evaluation, the ECU will assist in providing an assessment and guidance, including determining if the evaluator should be dismissed and/or the evaluation terminated. Evaluation reports should also address any potential or actual conflicts of interests declared by evaluators and measures taken to mitigate any negative consequences.

> Independence means that the evaluator should be able to work freely without any outsider interference, while impartiality means that the evaluator should not be biased with respect to what is being evaluated. (UNEG Ethical Guidelines for Evaluations)

It is equally important that evaluators have the freedom to conduct their assignment impartially and to make informed evaluative judgement without undue pressure, including potential negative effects on their career development (*see UNEG Norms and Standards for Evaluations 2017, Norm 4 Independence, p.11*).
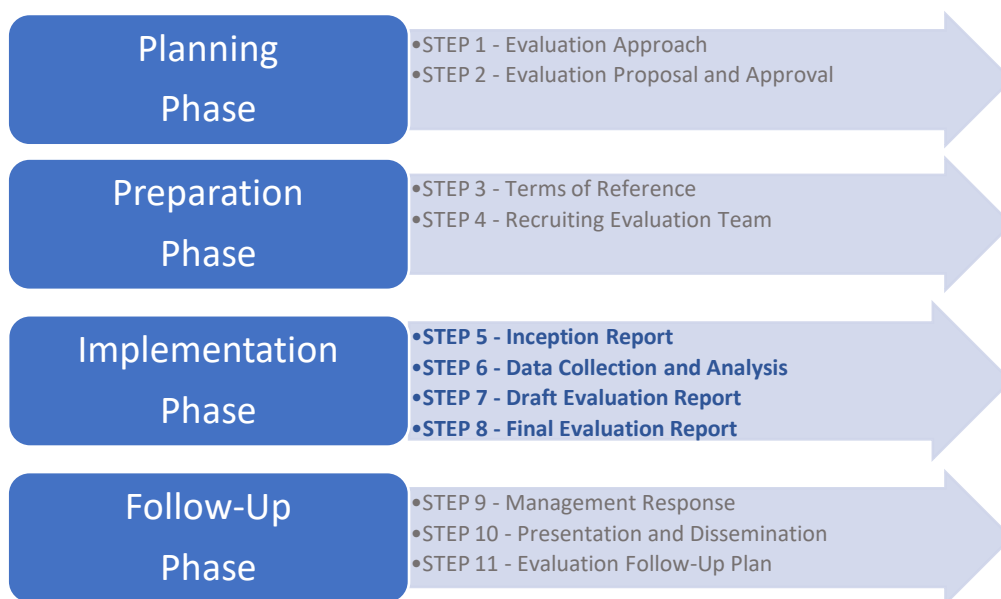
Further information on ethical considerations for evaluation processes can be found in Chapter 2 Evaluation in the UN System.

CHAPTER FIVE: IMPLEMENTING AND MANAGING EVALUATIONS

Once the terms of reference have been agreed upon and the evaluation team is in place, the actual evaluation can finally begin. This chapter covers the steps involved in the Implementation Phase of the evaluation process, from the initial briefing of the evaluation team through to the approval of the final evaluation report.

The main audience for chapter five is those who are involved in the management of evaluation processes and evaluation team members, although it will also be of interest to other stakeholders including senior managers and those who participate in advisory capacities such as Evaluation Reference Group members. As was the case in the previous chapter, the term 'Evaluation Manager' is used to refer to the person who has designated primary responsibility for overseeing the preparatory and implementation phases of the evaluation process.

Figure 5.1: Phases and Steps of the Evaluation Process

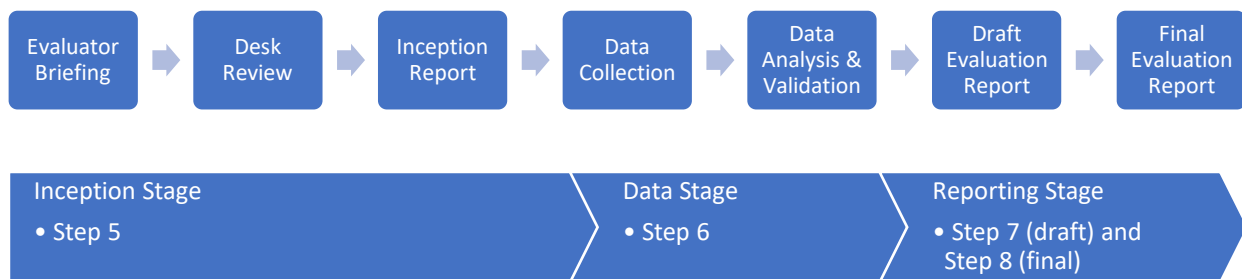| Planning Phase	•STEP 1 - Evaluation Approach •STEP 2 - Evaluation Proposal and Approval
Preparation Phase	•STEP 3 - Terms of Reference •STEP 4 - Recruiting Evaluation Team
Implementation Phase	**•STEP 5 - Inception Report** **•STEP 6 - Data Collection and Analysis** **•STEP 7 - Draft Evaluation Report** **•STEP 8 - Final Evaluation Report**
Follow-Up Phase	•STEP 9 - Management Response •STEP 10 - Presentation and Dissemination •STEP 11 - Evaluation Follow-Up Plan

IMPLEMENTATION PHASE

Figure 5.1 shows the various activities that take place during the Implementation Phase. They are presented in four steps following on those in the previous phases. **Step 5** covers the inception stage activities that lead up to the approval of the Inception Report, which is the document that sets out the road map for the evaluation. **Step 6** focuses on collecting, verifying, and analysing data from multiple methods and sources. **Step 7** covers the preparation of the draft evaluation report. **Step 8** involves the finalization of the evaluation report, which takes into account stakeholder feedback on the draft report.

Ideally, data collection only begins when the Inception Report is agreed and approved. However, sometimes there is some overlap between stages, for example, data collection begins during the desk review process and can extend into the development of the draft report if a data gap is uncovered.

Figure 5.1 Implementation Phase Steps and Activities

```
Evaluator      Desk        Inception     Data        Data         Draft        Final
Briefing   →   Review  →   Report    →   Collection  Analysis &   Evaluation   Evaluation
                                                     Validation   Report       Report
```

Inception Stage	Data Stage	Reporting Stage
• Step 5	• Step 6	• Step 7 (draft) and Step 8 (final)

STEP 5 – INCEPTION STAGE AND REPORT

Introductory Meetings – This stage should begin with a formal briefing session, which can be held virtually or in-person. It is scheduled by the Evaluation Manager for the evaluation team and key staff who will be involved in supporting the evaluation process. The Evaluator Briefing provides an opportunity to welcome the evaluation team on board, review the purpose and expectations of the exercise, have a substantive discussion on how the evaluation will be carried out, and discuss logistical and administrative issues. This is also a time to revisit the desk review materials that have already been provided to the evaluators and to identify other documents, templates and resources that could provide useful information to the team as they prepare the Inception Report. A sample agenda for the briefing meeting and documents to share with evaluators are provided in the boxes below.

For evaluations that involve Evaluation Reference Groups (ERGs), a second meeting should be held to introduce the team to the ERG members and other key partners and stakeholders. This Kick-Off meeting is useful for facilitating initial contact and for helping evaluators to understand stakeholders' interest and needs.

In the case of Internal Evaluations (IEs) conducted without an external evaluator, kick-off meetings are still important for signaling the launch of the evaluation and making sure there is a shared understanding by the project teams of the purpose, scope, and expectations for everyone's participation in the process.

> The briefing meeting is a key opportunity to discuss quality standards for UNOCT evaluations and to emphasize the need for the process to be human rights and gender responsive, and disability inclusive.

Box 5.1: Briefing Meeting Agenda

Suggested Agenda Items for Briefing Meeting

- Welcome and introductions
- Overview of programme/project being evaluated
- Evaluation purpose and scope
- Evaluation deliverables and timeline
- Stakeholders to be involved in the evaluation
- Evaluation quality standards, including the need for the process to be human rights and gender responsive and disability inclusive
- Expectations regarding evaluation reporting (length, style, structure, required templates, etc.)
- Administrative and logical issues (lines of communication, roles of those supporting evaluation process, scheduling and travel arrangements, invoicing, approval procedures, etc.)
- Evaluation project risks and mitigation/management strategies
- Desk review material

Desk Review – Following the briefing and introductory meetings, the evaluation team will carry out a thorough review of strategy and planning documents, programme reports, monitoring data and other relevant information. It is incumbent upon the Evaluation Manager to provide this material and the report templates to be used in a timely manner. At this time, it may also be useful for the evaluation team to have preliminary conversations with key informants such as UNOCT programme unit or section heads who can provide additional background or context for the evaluation.

Based on insights gained from the desk review[39], and before moving on to the next step, it is recommended that the evaluators discuss with the Evaluation Manager any proposed revisions to the theory of change/results framework[40], as well as to the set of criteria and questions. These elements provide the foundation for the evaluation. It is particularly important that the number of questions is manageable and that it is feasible to answer them within the scope and timeframe of the evaluation process.

Box 5.2: Desk Review Documents

Key Documents for Evaluation Team

In addition to relevant programme/project documents, the following should be shared with the evaluation team to help ensure the evaluation process and products consider UNOCT's strategic direction and adhere to UNOCT quality standards. These should be provided right after the kick-off/introductory meeting so as not to delay the process.

- UNOCT Strategic Plan and Results Framework 2022-2025
- UN Global Counter-Terrorism Strategy and its Biennial Reviews
- UNOCT Evaluation Policy (2021)
- UNOCT Gender Mainstreaming Policy and Action Plan (2022)
- UNOCT Disability Inclusion Policy (2023)
- UNOCT Evaluation Handbook and annexes (2023)
- UNEG Norms and Standards for Evaluation (2016)
- UNEG Guidance on Integrating Human Rights and Gender Equality in Evaluation (2014)
- UNEG Guidance on Integrating Disability Inclusion in Evaluations (2022)
- UNEG Ethical Guidelines for Evaluation (2020) and Pledge of Ethical Conduct (2020)

Inception Report – This document is produced by the evaluation team and provides a conceptual framework or roadmap for the evaluation. It ensures a shared understanding between UNOCT, other key stakeholders and the evaluators of how the evaluation will be carried out. Box 5.3 provides an outline of the type of information to be included in the Inception Report, with the required template being provided in Annex 7(f).

---

[39] The desk review should allow the evaluators to provide an initial analysis based on the programme-related documents received, e.g., timeliness of donor reports submitted, relevance of the intervention, effectiveness of the results framework and ToC.

[40] At this stage, suggested changes to the results framework would focus on any needed adjustments to the indicators to ensure that they are measurable for the purposes of the evaluation, i.e., developing proxy indicators. Shortcomings found in the results framework, the ToC, and/or in monitoring processes should all be addressed in the recommendations emerging from the evaluation.

Inception Report

- **Background and context** to demonstrate understanding of the intervention being evaluated.
- **Evaluation objective, purpose, and scope** to demonstrate understanding of the evaluation's focus. Any changes from the ToR should be noted.
- **Evaluation criteria and questions** that have been refined and agreed with the Evaluation Manager.
- **Summary of the desk review** material which would include preliminary findings for each evaluation criterion and inform the course of action for the evaluation.
- **Analysis of the results framework and theory of change** (ToC), if available, and preliminary reformulations if needed.
- **Stakeholder map** that identifies all stakeholder groups (duty-bearers and rights holders) involved in the intervention being evaluated and their roles. A template for a stakeholder analysis is provided in Annex 7(d).
- **Further elaboration of the methodology** proposed in the ToR which should include the overall approach (including a specific section on addressing cross-cutting themes of human rights, gender, leave no one behind, and disability inclusion), the criteria and evaluation questions, a sampling strategy that shows how different stakeholder groups will be included, data collection and analysis methods and tools, quality assurance processes, ethical considerations, identification of risks and proposed mitigation/management strategies, and limitations to the approach. Any changes from the ToR, discussed in the kick of meeting or otherwise, should also be stated.
- **Comprehensive evaluation matrix** that includes the evaluation criteria, questions, assumptions, data collection methods and sources, and indicators for measuring success. A template is provided in Annex 7(e). The full matrix is usually included as an annex but should also be referenced in the main body of the report.
- **Workplan** with associated activities, timeline, deliverables, roles, responsibilities, and well as the travel schedule for any field missions.
- **Resource requirements** associated with the evaluation activities and deliverables in the workplan. Assistance required from UNOCT for each activity (e.g., scheduling interviews, organizing focus groups, sending out survey invitations), should be discussed with the Evaluation Manager and specified in the Inception Report.
- **Annexes** to include:
  - Evaluation matrix
  - List of documents reviewed
  - Draft data collection tools
  - List of stakeholders to be involved as participants in the evaluation (via interviews, focus group discussions, etc.)

Review and Approval Process

The draft Inception Report should be first shared with the Evaluation Manager for feedback on its overall quality and suitability for meeting the expectations set out in the ToR. The Inception Report will subsequently be shared with the Evaluation Compliance Unit (ECU), as well as the ERG and programme funders (as relevant to the type of evaluation) for their review and comments. The review process may require several rounds of reiteration before approval. Once approved, the formal data collection process can begin.

See Chapter 7 – Quality Assurance and Assessment for guidance on reviewing Inception Reports.

Inception Report Approval	Approval by
Centralized Evaluations	Input from ERG, PM, and PMU. Approval by ECU.
Independent Programme/Project Evaluations	Input from ERG and ECU. Approval by ECU.
Internal Evaluations	Input from Programme Team. Approval by ECU.

STEP 6 – DATA COLLECTION, ANALYSIS AND VALIDATION

Data Collection – Key Considerations

Scheduling missions, meetings, and other data collection activities – Once the inception report has been agreed, with the list of stakeholders (both rights holders and duty bearers) to be interviewed or met, the Evaluation Manager will need to work with the evaluation team lead in preparing a detailed schedule of events. In order to allow for broad consultation with the variety of stakeholders, any field missions should be carefully planned in advance, devoting a sufficient number of days for this exercise so that the range of stakeholders can be met, recognizing that some of them can be hard to reach (often the case with rights holders). Consultations should be organized in a gender-sensitive manner, taking into consideration context and gender-specific needs and concerns for the meaningful participation of all stakeholders. Data collection activities can take a considerable amount of time to organize regardless of whether remote methods are used or if travel is involved. Both the Evaluation Manager and the evaluation team need to take this into account in their work planning.

Mixed methods – Even though scheduling can be challenging, it is important that a variety of methods be used as part of data collection in order to elicit the range of stakeholders' views, opinions, and assessments. The most common methods are individual and group interviews, surveys, and questionnaires as well as visits to project sites and observation. As every method has the potential to bring new information and insights, additional and innovative approaches are encouraged. It is this variety of methods and stakeholders that helps to confirm (triangulate) the evidence obtained.

> Evaluators are encouraged to use additional approaches beyond those that are most common in order to obtain the most comprehensive set of qualitative and quantitative data feasible within the scope of the evaluation.

Logistical support – The Evaluation Manager is responsible for ensuring that the evaluators are provided with adequate support to carry out their work. This includes providing introduction to and/or setting up meetings with key stakeholders (and any additional interlocutors deemed necessary by the evaluation team), developing a comprehensive agenda for the field missions, travel, and transport, as well as translation and interpretation services if needed. The support also includes ensuring the safety of the evaluation team. The evaluation team members must undertake relevant United Nations security training and be aware of, and compliant with, related security protocols (see UN Department of Safety and Security). The Evaluation Manager will be assisted by project teams and administrative staff as needed. In the case of country missions where UNOCT has a field presence, that office should be delegated the responsibility for all in-country logistical arrangements and support for the evaluation team.

> Interviews and meetings are confidential and evaluation reports should not attribute any statements or findings to specific individuals.

Ensuring independence – There is a delicate balance between providing adequate support for the evaluation and maintaining the independence of the exercise. To ensure independence and

confidentiality, UNOCT staff should not directly participate in any data collection activity other than when their input is needed as a data source. Independent interpreters may be present for language support during qualitative data collection, such as interviews and focus groups.

Data Analysis and Validation – Key Considerations

Data analysis – Although analysis of information being obtained occurs throughout the implementation stage, once all data has been collected, a different analytical process is undertaken. This involves systematically organizing, comparing, and synthesizing information that was obtained through each method. Increasingly, evaluators rely on software applications particularly for analysing large data sets. More details about undertaking data analysis can be found in Chapter 8 - Guide to Good Evaluation Practice.

Quality assurance – The evaluation team is responsible for conducting internal quality assurance to make sure that data collection is being carried out as agreed in the Inception Report and in an ethical manner, and that it is accurately recorded, analysed, and interpreted. Data validity, reliability, and triangulation are also essential components of rigorous and credible evaluation methods that quality assurance should address.

- **Validity** refers to the accuracy and relevance of the data, for example whether or not the data collection tools are measuring what they are intended to measure.

- **Reliability** refers to having consistency in results using the same method, for example whether similar findings would come from using the same tools multiple times.

- **Triangulation** refers to using multiple sources of data and multiple methods of data collection and analysis in order to verify and substantiate information.

For large and more complex evaluations, internal processes may be supplemented by evaluation team member(s) whose only role is quality assurance. Alternatively, an external firm can be contracted specifically to conduct quality assurance.

Presenting Preliminary Findings – It is good practice for evaluators to share their initial findings with stakeholders immediately after the data collection phase and before drafting the evaluation report. Referred to as 'validation meetings' or 'debriefing meetings', these events provide the opportunity to get feedback on initial findings and conclusions, address any factual errors or misunderstandings, and generally heighten engagement with the evaluation process. Preliminary recommendations may also be discussed. Depending on the scope and complexity of the evaluation, such meetings may initially be held with the Evaluation Manager and UNOCT programme staff and managers, and then later with the ERG and other key stakeholders. In cases where data is collected in multiple countries, validation meetings can also be held at the conclusion of each country mission. At a minimum, all evaluations should include at least one validation meeting. Programme/Project Managers should take into account the time required to prepare for these sharing sessions in calculating the level of effort required of the evaluation team, particularly when formal presentations are required.

STEP 7 – DRAFT EVALUATION REPORT

The evaluation report is the principal product of the evaluation. It is first delivered as a well-written and carefully constructed document in full draft form. The draft evaluation report must be developed according to the UNEG norms and standards as well as other specific requirements of UNOCT. As such, it should be formatted as per the requirements set out in reporting templates.

Box 5.2 provides an outline of the type of information to be included in all evaluation reports. Annexes 7(g), 7(h) and 7(i) includes the required template for each category of report – Centralized Evaluations (CEs), Independent Programme/Project Evaluations (IPEs), and Internal Evaluations (IEs).
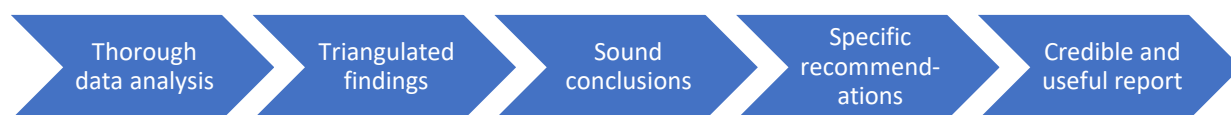
Evaluation Report

- **Executive Summary** that can serve as a stand-alone document, concisely presenting the object of the evaluation, the methodology used, and the main results including recommendation and lessons learned.

- **Introduction** which includes an introduction to the intervention, background and context information, stakeholders involved in the intervention, and the intended chain of results.

- **Evaluation purpose and scope** to confirm the reason for the evaluation, its main focus, what it includes and doesn't include, evaluation questions, and stakeholder participation in the assessment. Any changes from the ToR should be noted.

- **Methodology** should provide an updated account of the approach and methods that were proposed in the Inception Report. This includes the methodological approach, data collection and analysis processes, sampling strategy, approach to cross-cutting themes (human rights, gender equality, leave no one behind, disability inclusion), ethical considerations, and methodological limitations. Supplementary information can be provided in the annexes. Any changes from the ToR should also be stated.

- **Findings** systematically present the evidence obtained for each evaluation question and reflect the perspectives of the range of stakeholders. Preferably, this section is organized by evaluation criteria.

- **Conclusions** convey the overall strengths and shortcomings of the intervention found, highlight accomplishments and areas for improvement, and discuss any significant unintended results/impact. It is good practice to also present this section by evaluation criteria.

- **Lessons Learned** should be the most important ones that emerged from the evaluation process about what worked well and/or did not work in the process of planning and implementing the intervention. The lessons should be of interest to a wide audience and applicable to other situations.

- **Recommendations** should be the main ones emerging from the evaluation, align with the evaluation purpose, and be actionable.

- **Annexes** to include:
  - Terms of Reference
  - Evaluation Matrix
  - Theory of Change and/or Results Framework (indicating if it was changed and required PRB recommendation and USG approval of the project revision)
  - Evaluation Tools
  - List of Documents Reviewed
  - List of Stakeholders Consulted

Key Considerations

The evaluation report describes the intervention being evaluated and the evaluation methodology, but its main purpose is to present evidence-based findings, conclusions, recommendations, and lessons learned in an accessible format. In addition to following the reporting template and guidelines, top-of-mind issues for evaluation reports are that (a) there is a logical and systematic flow to the case being presented, and (b) it is crafted in a way that is easily read and accepted.

Systematic approach - As figure 5.2 helps to illustrate, useful and quality reports are those that systematically link evidence from the findings through to the conclusions and recommendations. The questions are connected to findings and the findings are based on solid data collection as well as analysis. Conclusions are drawn from those findings and referenced accordingly, providing the basis for the recommendations given.

Figure 5.2 Systematic flow of an evaluation

Thorough data analysis → Triangulated findings → Sound conclusions → Specific recommend-ations → Credible and useful report

Presentation and usability – High priority needs to be given to the report's presentation and its usability by the intended audiences. As such, both the draft and final evaluation reports, should:

- Use clear and concise language, so that they are accessible to all users.
- Set a positive tone that avoids blame and presents the results in a balanced and constructive manner.
- Ensure anonymity at all times, whereby the source of any finding cannot be attributable or identifiable.
- Employ visual aids[41] (beyond charts and graphs) and design techniques that communicate key information in an eye-catching manner.
- Have clear, useful, time-bound, and actionable recommendations aimed at enhancing the project performance and improving the sustainability of results.
- Be proofread for spelling and grammatical errors.

This topic is discussed further in Chapter 6 - Using Evaluation Results.

Language Requirements

UNOCT evaluation reports are typically submitted in English. There are exceptions for the official United Nations languages, but this needs prior discussion and agreement with the ECU. Evaluation reports written in a language other than English need to be accompanied by an English translation of the executive summary prepared by a competent translator who is cleared in advance by the ECU.

Review and Clearance Process

The review process may take several rounds of revisions[42] before the Draft Evaluation Report is cleared for distribution. The initial draft evaluation report should be first shared with the Evaluation Manager for feedback. The Manager's primary responsibility is to make sure that it corresponds with the inception report, particularly the agreed upon evaluation questions and methodologies, and does not contain factual errors. Where errors of fact or misunderstanding of approaches are found, documentation should be provided to the evaluators to support comments and requests.

See Chapter 7 – Quality Assurance and Assessment for guidance on reviewing Evaluation Reports.

The Evaluation Manager also considers what improvements may be needed for the document to meet UNOCT quality criteria for evaluation reports. Once this feedback has been addressed by the evaluation team, the next version of the draft report (the provisional draft) will be shared with other internal stakeholders. Depending on the type of evaluation, this would include the Programme Manager (PM), the Human Rights and Gender Section (HRGS), and the Evaluation Compliance Unit (ECU).

---

[41] There are many resources for use of info graphics and data visualization in presenting evaluation results. See, for example, https://www.betterevaluation.org/en/resources/data-visualization-design-process-step-step-guide-beginners

[42] Good practice is to have no more than three rounds of revisions unless there are extraneous circumstances based on the quality of the report, the complexity of the intervention, etc. It is important that the number of anticipated revisions is reflected in the evaluation costing.

One issue that all reviewers should pay particular attention to is the recommendations - to ensure they appear substantiated by the findings and are realistic and feasible to implement.

After all comments have been taken into account by the evaluation team (which may involve further revisions), the Evaluation Manager will clear the 'Provisional Draft Evaluation Report'. The approval processes are shown in the table below and in figure 5.3.

	Provisional Draft(s) Evaluation Reports	Cleared by
	Centralized Evaluations	EM -> Section Chiefs -> ECU
	Independent Programme/Project Evaluations	EM -> Section Chiefs -> ECU
	Internal Evaluations	ECU

The provisional report is then shared with the ERG and any other key stakeholders such as programme funders (as relevant to the type of evaluation) for their review and comments.

Figure 5.3 Review and Clearance Processes for Draft Report(s)
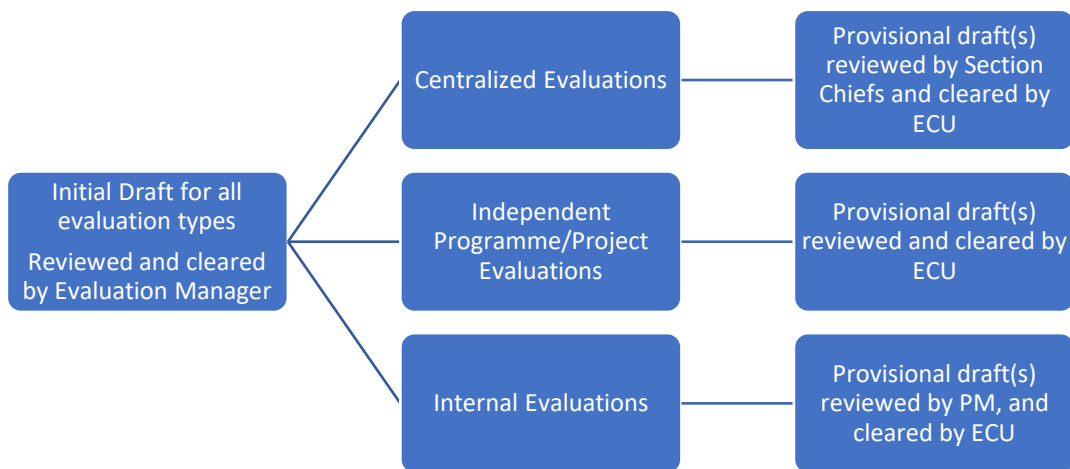
The Evaluation Manager should compile the collection of comments, questions, and requests for clarification received into one document, an Evaluation Report Comments Matrix (see figure 5.4), which is then shared with the evaluation team.

Figure 5.4 Template for Evaluation Report Comments Matrix

Chapter, section, and page number	Paragraph number / line number	Comments	Evaluation team response and/or action taken

STEP 8 – FINAL EVALUATION REPORT

Addressing feedback on draft report

Evaluators need to respond to all comments and questions provided on the draft report. However, they are not compelled to make changes other than those related to factual errors, to the language used (as this needs to be appropriate for an official United Nations document), and to issues raised about gaps in meeting evaluation report quality standards. The evaluators are independent and as long as findings and recommendations of the evaluation report are reliably based on sound methods for data collection and analysis and evidence, they are not required to adjust their report substantively in response to general criticism.

Disagreements – These should be discussed and documented, and any efforts should be made to resolve them. In circumstances where the commissioning unit and/or stakeholders cannot come to agreement on a finding with the evaluators, a note about the disagreement should be made in the report. If the disagreement is on a substantial finding or if there is continued disagreement with a conclusion or recommendation, this should be clearly stated in the Management Response, with the supporting rationale. Evaluators should use the Evaluation Report Comments Matrix to provide further evidence to substantiate their findings. In cases where the comments matrix is not used, evaluators can provide formal documentation in another format and send this to the Evaluation Manager and the Evaluation Compliance Unit.

Final formatting and editing

The final report should be carefully proofread by the evaluation team for errors and accessible language. Any final photos or other visual aids should be added. There should also be a final review of formatting and design to make sure the report aligns with UNOCT Visual Identity requirements[43], including for use and placement of the UNOCT logo.

Review and Approval Process

The Evaluation Manager is then responsible for reviewing the final evaluation report to ensure that relevant input, comments, and corrections have been adequately considered. If not, a revised version should be requested from the evaluation team that addresses the outstanding issues. The final version will then need to be approved by the parties identified below.

	Final Evaluation Report Approval	Cleared by
	Centralized Evaluations	Section Chiefs -> ECU -> C/OUSG -> USG
	Independent Programme/Project Evaluations	Section Chiefs -> ECU -> C/OUSG -> USG
	Internal Evaluations	Section Chiefs -> ECU

---

[43] Evaluation Managers should share the UNOCT Visual Identity document (2020) with evaluation team undertaking CEs and IPEs. It outlines the branding requirements including logo, color schemes and typography needed for all UNOCT information products made publicly available. Internal Evaluation reports do not need to adhere to the branding requirements if they are for internal use only.
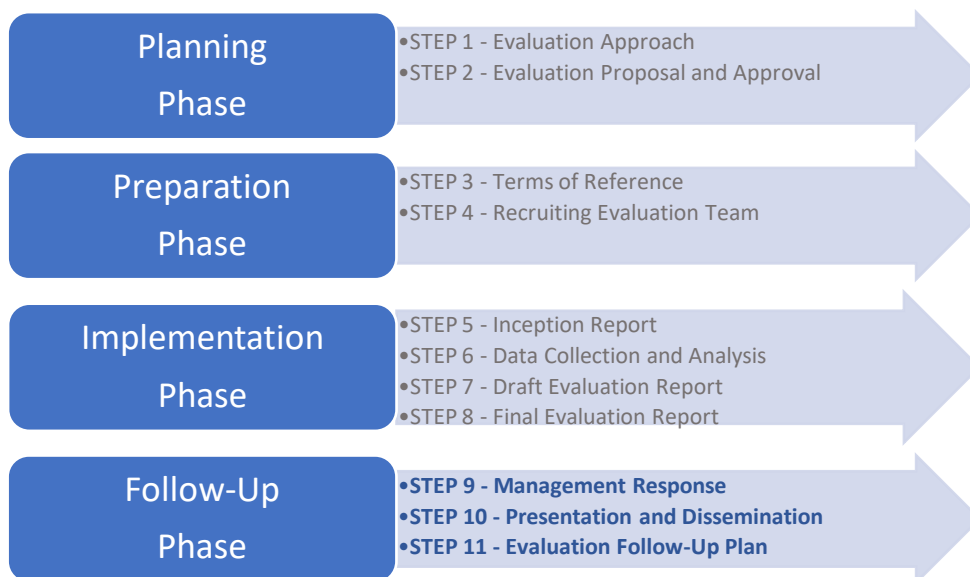
CHAPTER SIX - USING EVALUATION RESULTS

Evaluations are undertaken for the purposes of accountability, organizational learning, and knowledge generation. However, **evaluations can only be effective in advancing UNOCT's work if resulting analyses are used.** This chapter addresses the final phase of the evaluation process. It focuses on the utilization of evaluation results and covers follow-up planning, the Management Response process, and the dissemination of the evaluation report and lessons learned.

This chapter is primarily designed for an internal audience, in particular evaluation managers, programme/project managers, senior management, and others involved in the implementation and tracking of the lessons and recommendations emerging from evaluation processes. However, it also includes tips on using visual aids and other presentation techniques that may be useful to evaluators as they develop the design and format of the evaluation report and follow up communications materials.

Figure 6.1: Phases and Steps of the Evaluation Process

Planning Phase	•STEP 1 - Evaluation Approach •STEP 2 - Evaluation Proposal and Approval
Preparation Phase	•STEP 3 - Terms of Reference •STEP 4 - Recruiting Evaluation Team
Implementation Phase	•STEP 5 - Inception Report •STEP 6 - Data Collection and Analysis •STEP 7 - Draft Evaluation Report •STEP 8 - Final Evaluation Report
Follow-Up Phase	•**STEP 9 - Management Response** •**STEP 10 - Presentation and Dissemination** •**STEP 11 - Evaluation Follow-Up Plan**

FOLLOW-UP PHASE

STEP 9 – EVALUATION FOLLOW-UP PLANNING

The completion of the evaluation report is not the end of the evaluation process. Further work is still needed to help ensure the sharing and uptake of results, learning from the evaluation, and the incorporation of findings into the overall programming cycle. This focus on utilization should be part of the initial evaluation planning and continue to be considered as the evaluation proceeds. The Evaluation Compliance Unit (ECU) in the Office of the Under-Secretary-General for Counter-Terrorism is available as a resource for this task.

Ideally, the stakeholder mapping process undertaken during the preliminary stages of the evaluation will have identified groups that need to be informed, consulted, or actively involved in the follow-up process.

> The main value of an evaluation is ultimately what is done with the results, specifically how results are used in decision-making processes to improve and accelerate CT/PCVE results

The first priority in follow-up planning is determining who needs to be involved in the development of the Management Response. This will primarily be UNOCT staff but should also include members of the Evaluation Reference Group (ERG) and Implementing Partners (IPs) who are not represented in the ERG. If a fully participatory evaluation approach is being used, then representatives of those most affected by the intervention (rights holders/beneficiaries) should also be involved in this process.

The next priority is determining how the evaluation results and lessons should be shared internally and with the broader stakeholder community, including those involved in consultations, Member States and the wider public, as is relevant to the subject of the evaluation. Table 6.1 is an example of a Communications Plan that sets out issues to consider for different target audiences with further guidance on dissemination provided at the end of this chapter. The target audience columns should be adjusted for the type of intervention being evaluated. Note that the list of evaluation products includes an 'evaluation brief'. This is a two-page document that provides a succinct and visually appealing overview of the evaluation and its main results. It is a required final deliverable, along with a PowerPoint presentation, for evaluators conducting Centralized Evaluations (CEs) and Independent Programme/Project Evaluations (IPEs). A template for evaluation briefs can be found in Annex 7(k).

Table 6.1: Sample Communication Plan for Sharing Evaluation Results

Target Audience	UNOCT	Implementing Partners	Policymakers	Wider Public
Knowledge needs / interests	Targeted evaluative evidence to inform decision-making and organizational learning.	Targeted evaluative evidence to inform decision-making and organizational learning; to support further engagement with UNOCT.	Targeted evaluation results to inform legislative actions, decision-making and policy development.	Targeted evaluation results to support awareness raising.
Evaluation products to meet needs	Evaluation report; executive summary; presentation; evaluation brief.	Evaluation report; executive summary; presentation; evaluation brief.	Evaluation report; executive summary; presentation; Policy briefs; infographics.	Infographics; videos; blogs; photo stories.
Dissemination channels and platforms to use	Final presentation by evaluation team; workshop; webinar; programme/project manager forums.	Final presentation by evaluation team; workshop; webinar.	Conference; workshop; webinar; face-to-face engagement; policy briefs, newsletters.	Website; social media.
Timing of dissemination*				
Costs involved*				
Responsible unit / person	ECU; Programme Managers.	ECU; Programme Managers.	ECU; Communications Unit.	Communications Unit.

*The timing of dissemination and potential costs incurred will be dependent on the type and subject of the evaluation.

STEP 10 – MANAGEMENT RESPONSE

The Management Response is a formal mechanism to ensure the results of the evaluation, and in particular its recommendations, will be used. It provides management views of the evaluation recommendations, including whether and why management agrees, partially agrees, or disagrees

with each one. For those agreed to, the Management Response identifies who is responsible for their implementation, what the action points will be, and the associated deadlines. The form for the response also enables the tracking of the status of implementation of the agreed actions.

All UNOCT evaluations are required to have a Management Response. The template for this is provided in Annex 7(j). Responsible programme managers are required to prepare this document for evaluations undertaken under their portfolio. It is important to have stakeholder participation in the decisions on how to respond to the evaluation and agreement on clear roles and responsibilities for the implementation of the action steps. It is advisable for the consideration of responses to begin once it becomes clear what the emerging recommendations are likely to be, which would be at the draft evaluation report stage, if not earlier. In the case of CEs and IPEs, the Management Response is intended to be included in the published evaluation report, and therefore should be finalized within one week of receipt of the final evaluation report from the evaluators. If it is not produced within that timeframe, then the Management Response will stand as a separate report that will be published alongside the evaluation report.

> The Management Response should be included in the published version of the evaluation report as this practice increases utilization of the evaluation results.

The process for the review and clearance for all management responses is shown below.

	Management Response Approval	Cleared by
	Centralized Evaluations	Section Chiefs -> ECU -> C/OUSG -> DUSG ->USG
	Independent Programme/Project Evaluations	Section Chiefs -> ECU -> C/OUSG -> DUSG -> USG
	Internal Evaluations	PM -> Section Chiefs -> ECU

C/OUSG = Chief, Office of the Under-Secretary-General   DUSG = Deputy Director to the Under-Secretary-General

It is important for the evaluators to be aware of the management response process prior to developing recommendations. **Management response actions are more feasible to implement when**:

- There is a reasonable number of overall recommendations - a range of between 7 to 10 is typical and there can be sub-recommendations within each.
- Recommendations are designed with their follow-up in mind - that means they are clearly written, implementable, specific about what needs to be done, and are targeted in that they identify the party(ies) responsible for follow-up.
- There is brief guidance provided on their implementation (for example, by discussing their operational, technical, and resource implications).

> Evaluation managers should remind evaluators of the need to structure recommendations so that a Management Response can be prepared.

Evaluators also need to be mindful of the need to target recommendations to UNOCT and not Member States. Recommendations can be directed to other UN entities or other organizations in cases where they are taking part in a joint evaluation process.

Joint Programme Management Responses

In the case of joint evaluations, evaluation managers should coordinate with programme/project partners in the development of management responses and key actions. UNOCT units are only responsible for those recommendations targeted at them and should develop management responses and key actions only for these.

Implementation of the Management Response

The Chief of the Substantive Unit/Section/Branch, in collaboration with PMU and section/branch evaluation focal points, will oversee the implementation of the actions provided in management responses, such as follow-up reports or tracking systems. As noted in the UNOCT Evaluation Policy, ensuring follow-up is the responsibility of management. Therefore, follow-up will be overseen by the Programme Review Board as delegated by the Under-Secretary General.

The ECU in the Office of the Under-Secretary-General for Counter-Terrorism will provide oversight to ensure that management responses to evaluations are provided, that the actions contained in management responses are adequate to substantially address agreed recommendations, and that the recommendations are appropriately implemented.

Once the Management Response has been cleared, the Evaluation Compliance Unit will enter the actions into the Evaluation Management Tracking Table (EMTT) and track the status of their implementation. It will then be up to the Programme/Project Manager, or their designate, to monitor implementation of key actions and use the Management Response form to share this information with the ECU. **Updates should be provided on a quarterly basis.**

STEP 11 – DISSEMINATION OF RESULTS AND LESSONS LEARNED

Stakeholder Meeting or Workshop

Once the final evaluation report has been cleared, it is usual practice for a meeting or workshop to be held for UNOCT staff, the Evaluation Reference Group, implementing partners, and other stakeholders as are relevant to the subject of the evaluation including funders, government officials, and rights holders as feasible. This in-person and/or online event is hosted by UNOCT with the evaluation team leader presenting (in a slide presentation) an overview of the methodology used along with the main findings, conclusions, lessons, and recommendations. Ample time should be set aside for questions and discussion.

A workshop format provides more opportunities than a regular meeting for in-depth conversations about programming implications and can therefore inform the Management Response and facilitate and build support for the adoption of the recommendations. In some cases, it may be appropriate to hold separate sharing events for different stakeholder groups or to facilitate attendance for people from different time zones.

Publication of the Evaluation Report

The UNOCT Evaluation Policy calls for all final evaluation reports, including the terms of reference and complete set of annexes, to be published alongside the Management Response on the UNOCT website. This is a key element of organizational transparency and aligns with the United Nations Evaluation Group (UNEG) norms and standards. The ECU will facilitate the posting of reports.

Discloser and Dissemination – In fulfilling its accountability principle, UNOCT will ensure that stakeholders have easy access to evaluation reports. Depending on the nature of the evaluated work, some cases may require an exception to the disclosure rule. The Under-Secretary-General shall specify the conditions under which an exception to the general disclosure and dissemination provision can be granted. Notwithstanding, UNOCT will uphold stakeholders' confidentiality and anonymity to protect informants and to avoid unintended consequences related to the evaluation findings.

> Lessons learned are usually focused on sharing information about what went well and successful practices to emulate. However, equally important are lessons about what failed and why.

Using Professional Networks and Social Media

Sharing of results can be facilitated through a variety of professional networks. Participating in thematic or evaluation-related conferences, communities of practice, working groups, and

other sharing-type events provides beneficial networking opportunities for learning about and disseminating evaluation results. It should be kept in mind that the direct users of the evaluation, including implementing partners, are likely to also be able to use their own networks and channels to contribute to dissemination.

Sharing of Lessons Learned

The Evaluation Compliance Unit is responsible for ensuring that lessons are recorded and shared. Lessons will be uploaded into the Evaluation Management and Tracking Tool in the **Lessons Learned Repository.** This repository will have functionalities for push (automatically delivering to a user) and pull (where a user manually searches for lessons).

In addition, a collaborative knowledge platform will be created on the **UNOCT Connect & Learn platform** which will be also used as a repository and interactive platform for sharing lessons learned. The platform will support joint reflection on good practices in CT evaluations, including case studies on good evaluation approaches and tools on CT/PCVE, behavioural insights, gender equality, and human rights.

Formatting and Presentation Techniques

Also key to the utilization of evaluation results is making sure the information is available in accessible and suitable formats for the intended audience. To this end, increasing emphasis is being placed on:

- Having shorter and more concisely written evaluation reports and executive summaries. Annexes can be used to present important information that may not fit well into the main report such as additional information on the methodological approach and more detail on survey results.
- Incorporating design elements such as having variation in font type and colour, ample white space to avoid text-heavy pages, use of bullet points and text boxes, and formatting the annexes in a consistent style as the main report.

- Increasing the use of well-designed visual aids to convey key information such as graphs, charts, pull-out quotes from evaluation informants, photos[44] and infographics.
- Making evaluation reports easy to navigate including by hyperlinking the table of contents to the respective chapters and annexes, and by using distinctive typeface to clearly delineate headings and sub-headings.
- Providing barrier-free access to the evaluation products which includes making sure the language and format of the report is accessible to all potential users, avoiding excessive use of acronyms, and defining key terminology.

The evaluations products should also look professionally presentable. Although the evaluation team is not expected to have the reports expertly designed and edited, it is expected that they will be able to produce a well-formatted and designed report that has also been fully proof-read and reasonably well edited.

An infographic (information graphic) is a graphic representation of information that is designed to make data or other key points from the evaluation easily understandable at a glance. It requires the evaluator to condense results into very concise messages. In addition to quickly engaging readers, infographics have the benefit of being easily shared in multiple types of communication products and online. As can be seen in the example provided in figure 6.2, infographics generally go beyond regular graphic or charts.

More information on producing infographics as well as tools can be found at BetterEvaluation, USAID Learning Lab and Okanagan College Library.

---

[44] Identifiable images of evaluation participants and programme beneficiaries is typically not permitted and can only be included if cleared by the Evaluation Compliance Unit.

CHAPTER SEVEN: QUALITY ASSURANCE AND ASSESSMENT

As pointed out in the United Nations Evaluation Group (UNEG) norms and standards, evaluations must be credible, and a key element of credibility is having robust systems in place to ensure that evaluations meet quality standards. This chapter outlines UNOCT's quality assurance and assessment processes, the roles and responsibilities for this work, the various stages at which they take place, and what those expectations are.

7.1 ECU ROLE IN ENSURING ROBUST AND QUALITY EVALUATION PROCESSES

The Evaluation Compliance Unit (ECU) in the Office of the Under-Secretary-General for Counter-Terrorism has primary responsibility for promoting and supporting good evaluation practice across UNOCT. The Unit provides the standards, guidelines, and tools to help ensure evaluation quality, and in particular, adherence to UNEG Norms and Standards. As discussed in more detail below, the ECU carries out quality assurance and assessment for most UNOCT evaluations that are undertaken. Although their formal level of engagement is dependent on the category of evaluation, ECU staff are available to answer questions and provide advice on all evaluation processes.

7.2 DEFINING TERMS

Quality Assurance is a shared responsibility that happens throughout the different phases of the evaluation. It focuses on ensuring the activities undertaken during the evaluation align with UNEG and UNOCT standards and with the guidance and procedures set out in the UNOCT Evaluation Policy and this Evaluation Handbook. This is achieved by providing feedback, primarily to the evaluation team lead, on whether processes are meeting standards or if there are deficiencies or additional elements that should be considered, as well as advice on how to address these. Such feedback should not reflect negatively on the evaluation team. UNOCT interventions are complex and external evaluators need to quickly process a considerable amount of information about the organization, the context, the varied needs, and interests of multiple stakeholders, as well as the intervention. A successful evaluation requires a collaborative effort to make sure the evaluators have the necessary guidance, materials, insights, and access to the most important data sources in order to meet quality standards.

Quality assurance begins at the time of developing the Terms of Reference (ToR) and follows through to the design of the evaluation, data collection and analysis processes, and report writing. **Early investments in quality assurance tend to pay off** – issues uncovered in the initial stages of the evaluation can be addressed much more easily than after the report has been presented.

> The Evaluation Manager has to safeguard the integrity of an external, independent evaluation process. Quality assurance is meant to ensure adherence with quality standards for evaluations within the UN system and that evaluations are useful for UNOCT accountability and learning purposes. It is not meant to steer the evaluation team to a specific set of results or recommendations.

Responsibility for quality assurance rests with multiple actors. The Evaluation Manager plays a key role, as does the evaluation team itself. Depending on the category of evaluation, the Evaluation Compliance Unit (ECU), the Programme Managers, members of the Evaluation Reference Group (ERG), and audit and oversight bodies (e.g., OIOS, Joint Inspection Unit) will also be involved in providing feedback on the quality of the evaluation process. In most cases, the evaluation team lead is responsible for internal quality assurance/control, although in situations where there is a larger team, it is good practice for one

senior team member to have quality assurance as their sole responsibility. For high-level and complex evaluations, the ECU may determine that an external evaluation expert or firm that is independent from those conducting the evaluation be engaged specifically for quality assurance at all stages of the evaluation. In respect to the ERG, it is advantageous to have at least one member with some technical expertise and experience in evaluation design, implementation, and quality assurance so that they can provide relevant input and feedback on the quality of the process and deliverables.

Quality assurance roles and responsibilities for the different phases of the evaluation process are highlighted throughout this chapter.

It is highly recommended that the evaluation team be provided with the UNOCT EQA template and be reminded to refer to it before drafting the inception report. The evaluation manager and others with responsibility for quality assurance should also be familiar with this tool.

Quality Assessment happens at the end of the evaluation process. The final evaluation report is rated against a set of criteria to determine the extent to which the report meets UNEG and UNOCT standards. The expectations for UNOCT evaluation reports are described later in this chapter and the UNOCT evaluation report quality checklist is provided in Annex 4. The ECU is responsible for conducting quality assessment for Centralized Evaluations (CEs), Independent Programme/Project Evaluations (IPEs), and Internal Evaluations (IEs), and for providing the results to the Evaluation Manager. Evaluation Managers are then encouraged to share the EQA with the evaluation team. The ECU will also provide a brief report to Under-Secretary General (USG) on an annual basis that summarizes the overall strengths, weaknesses, and areas for improvement for the overall portfolio of evaluations conducted over the past year.

7.3 QUALITY ASSURANCE OF THE TOR

ToRs play a significant role in the quality of evaluations. If the TOR clearly directs the evaluator towards what the evaluation is expected to cover and accomplish, it is more likely that the evaluation will be of high quality.

Table 7.1: Quality Assurance of ToR

	Centralized Evaluations	Independent Programme/Project Evaluations	Internal Evaluations
Responsibility Party	ECU	ECU	PM and ECU[45]

What to look for:

- Does the document adhere to the requirements and guidance provided in the ToR template in Annex 7(c)?
- Does the stated purpose emphasize the need for the evaluation to be human-rights and gender responsive, as well as disability inclusive?
- Are the evaluation criteria and the proposed evaluation questions and methodology appropriate for the purpose of the evaluation?
- Are the proposed questions the most important ones to ask?
- Are the timeframe and budget realistic given the stated objectives, scope, and proposed methodology?
- Is the ToR clearly and precisely written, giving sufficient direction on the expectations for the exercise to the evaluators and stakeholders?

---

[45] In the case of Internal Evaluations, quality assurance will be undertaken by the ECU in consultation with the PM until the PMU has increased capacity to assume full responsibility for this role.

7.4 QUALITY ASSURANCE OF INCEPTION REPORTS

Quality assurance at the design stage of the evaluation focuses primarily on the intervention logic / theory of change and the methodological approach and tools. The evaluation team is expected to ensure that this work meets the required standards and that the design and methodology are appropriately expressed in the Inception Report. The Inception Report is then quality assured by responsible parties, and these are identified in table 7.2 for the different evaluation categories. Since all UN evaluations need to be designed to be human rights and gender-responsive and disability inclusive, this subsection includes guidance on these topics.

> It is critical to invest time in reviewing the Inception Report. If it is well done, it is more likely that the rest of the evaluation will proceed smoothly.

Table 7.2: Quality Assurance of Evaluation Design Phase and Inception Reports

	Centralized Evaluations	Independent Programme/Project Evaluations	Internal Evaluations
Responsibility Party	Evaluation Manager, ECU, and ERG	Evaluation Manager, PM, and ECU	Evaluation Manager, PM, and ECU

What to look for:

- Does the document adhere to the requirements and guidance provided in the Inception Report template in Annex 7(f)?
- Is there sufficient clarity in the programme/project logic or theory to assess the links between the intervention's activities and intended results?
- Do the proposed methodology and tools adequately address the ToR, particularly the evaluation questions, and UNEG norms and standards for evaluation?
- Are the proposed methodology and tools able to reliably and validly provide data to explore the different aspects of the evaluation criteria?
- Are the proposed methodology and tools able to capture issues of human rights, gender equality, disability inclusion, youth, leave no one behind and intersectionality, as well as the different experiences, views, and assessments of the variety of stakeholders? Do they integrate human rights and gender concerns throughout?
- Are the most relevant stakeholders identified for data collection?
- Are the potential risks accounted for, including risks to participation in the evaluation?
- Are ethical considerations adequately covered?
- Are adequate quality assurance processes described?

What does a gender-responsive evaluation look like?

Gender-responsiveness relates to both what the evaluation is examining and how the evaluation is carried out. This means:

- Assessing the extent to which gender and power relationships, including structural and other causes of inequities, have changed as a result of the intervention. The evaluation should consider how the intervention may have affected women and men and those with diverse sexual orientations and gender identities differently, and how it has contributed to the achievement of gender equality and women's empowerment (GEEW) and to human rights. As such, the aim to address GEEW should be reflected in the evaluation scope, criteria, and evaluation questions.

- Conducting the evaluation in a way that is inclusive, participatory, and respectful of all stakeholders, both rights holders and duty-bearers. As also discussed in Chapter 2, a gender responsive methodology is defined in the UN system as having the following parameters[46]:
  - 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 to enable triangulation and validation, and to ensure 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[47].
  - Evaluation conclusion and recommendations integrate lessons learned, good practices, and gaps related to gender mainstreaming throughout the intervention cycles.

> **Human rights and gender-responsive approaches to counter-terrorism and preventing and countering violent extremism (CT/PCVE) are discussed in Chapter 2 and in UNOCT's Gender Mainstreaming Policy.**

Other gender- and youth-responsive evaluation practices
  - Determining if it is more appropriate to hold separate meetings for women and men, and for girls and boys, as some participants in specific contexts might not feel comfortable speaking out in mixed settings.
  - Using active facilitation and participatory exercises (such as games, drama, ranking exercises, and drawings), having small working groups, asking specifically for youth/women's inputs, etc.
  - Identifying the priority issues for women and for young people, ensuring these are included in the consultation agenda, and giving them equal weight.

Additional considerations for human rights-responsive evaluations
A human rights-based approach (HRBA) follows the same practices as for gender-responsive evaluations but also takes into account issues such as:
  - Power dynamics amongst all stakeholders, not just between women and men.
  - The extent to which an initiative has increased the likelihood that rights holders claim their rights and that duty-bearers fulfil their obligations.
  - Use of gender sensitive and human rights-based language throughout the report (including by referring to 'rights holders' and 'duty bearers').

---

[46] This definition comes from the **UN System Wide Action Plan (UN-SWAP) on Gender Equality and the Empowerment of Women (GEEW)** which is part of the accountability framework for coordination of gender mainstreaming in the UN system.
[47] Ethical standards and UNEG Ethical Guidelines for Evaluation are also discussed in Chapter 2.

What does a disability inclusive evaluation process look like?

Disability inclusion (DI) also relates to both what the evaluation is examining and how the evaluation is carried out. DI is a requirement for all UN evaluations and guidance on its application is provided in the United Nations Disability Inclusion Strategy (UNDIS). Although the degree to which DI is taken up will vary based on the nature of the intervention, at a minimum, all evaluations should:

- Assess the extent to which the intervention paid attention to persons with disabilities (PWD) and addressed their needs and interests.
- Conduct the evaluation in a way that is inclusive and respectful of PWD. A DI evaluation process would be exemplified by including the following practices:
  - The evaluation team has 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 PWD and their representative organizations.
  - Accommodations are made to reduce the barriers for PWD to participate.
  - Evaluation findings and analysis provide data and evidence on disability inclusion.
  - The perspectives and voices of PWD are reflected in the findings.
  - The conclusions and recommendations of evaluations reflect findings on disability inclusion.

Note that these same practices are applicable to ensuring that evaluations are inclusive of youth.

7.5 QUALITY ASSURANCE OF DATA COLLECTION AND ANALYSIS

Quality assurance during the field phase of the evaluation is an ongoing process, and the responsibility is exercised by both the team leader and the evaluation manager. Throughout this phase, the team leader needs to ensure that all members of the team correctly understand which types of information must be collected, and how this information should be recorded and archived. Specific tasks in this regard are provided in table 7.3. This subsection also explains three processes that underpin quality evaluations – data reliability, validity, and triangulation.

Table 7.3: Quality Assurance of Data Collection and Analysis

	Centralized Evaluations	Independent Programme/Project Evaluations	Internal Evaluations
Responsible Party	Evaluation Team Lead (or team member/independent firm responsible for quality assurance)		

What to look for:
• Is the approach to data collection and analysis following the processes set out in the Inception Report? Are adjustments needed? (Any challenges to following the Inception Report must be brought to the attention of the Evaluation Manager).
• Did the testing of data collection tools confirm both validity and reliability? (these terms are explained below).
• Is the selection of data sources (both documents and key informants) following the stated sampling processes and the evaluation questions? Is the final sample balanced, and is it adequately representative (taking into account differing viewpoints and perspectives of the range of stakeholders)?
• Are all stated human rights and ethical practices for data collection being followed by all team members, including the principle of do not harm and the safe storage of data?

• Is data being carefully recorded with sufficient detail and in a uniform and systematic way so that is it useful for data analysis and report writing purposes?
• Is the collected data reviewed and cleaned so that it is usable for analysis processes?
• Are data analysis procedures double-checked for errors?
• Are there sufficient amounts of reliable data sources to show triangulation of key findings?
Responsible Party	Evaluation Manager
What to look for:
• Are the evaluators' selection of interviewees and other data sources appropriate? (The evaluation manager is responsible for supporting the evaluators in arranging interviews and accessing other data sources. The evaluation manager is expected to use their knowledge of the context as well as of the stakeholders to ensure that all of the main interest groups and stakeholders are taken into account during data collection.) • Do the preliminary findings presented at the end of the field phase appear valid? (The evaluation manager is tasked with assessing the validity of the preliminary findings and answers to the evaluation questions as presented by the evaluators during the debriefing session at the end of the field phase. They should point out weaknesses in the reasoning of the evaluators and point out those findings, conclusions or preliminary recommendations that do not appear to be sufficiently backed by evidence.)

Data Validity, Reliability, and Triangulation

Evaluations are considered more credible, and therefore are more useful, when the methodology is sound and rigorous. A sound approach includes the use of valid and reliable methods of data collection and analysis, and the use of triangulation for developing the findings.

- Validity refers to the accuracy and relevance of the data, i.e., whether or not the data collection tools are measuring what they are intended to measure.
- Reliability refers to having consistency in results using the same method, i.e., whether similar findings would come from using the same tool multiple times.

There are multiple methods for ensuring that data collection tools exhibit high validity and reliability. For example, to ensure validity, information obtained from using a tool multiple times should be reviewed to ensure it corresponds to the evaluation question. To ensure reliability, the tool could be tested more than once on the same individual; the tool could be administered by several people; or the tool could contain multiple questions that are aimed at yielding the same response. It is a good and necessary practice to test all data collection tools to ensure high validity and reliability.

Another important concept for ensuring the credibility of data is triangulation. Evaluation findings are strengthened when multiple pieces of evidence point in the same direction. Triangulation is accomplished when different data or types of information are used that come to broadly the same conclusion.

- Triangulation refers to using multiple approaches, methods, and sources for data collection and analysis to verify and substantiate information. It enables evaluators to overcome the bias that comes from single informants, methods, observations, or points of view. The purpose of triangulation is to increase the reliability of evaluation findings.

7.6 QUALITY ASSURANCE OF EVALUATION REPORTS

Most of the investment in quality assuring the evaluation report happens at the draft report stage. As noted in table 7.4, multiple parties have responsibility for providing feedback on whether the report meets UNOCT standards. The main tool to assess this is the Evaluation Report Quality Checklist

provided in Annex 4, with the table below providing overarching questions for reviewers to keep in mind.

Table 7.4: Quality Assurance of Evaluation Reports

	Centralized Evaluations	Independent Programme/Project Evaluations	Internal Evaluations
Responsibility Party	Evaluation Manager, ECU and ERG	Evaluation Manager and ECU	Evaluation Manager, PM and ECU

What to look for:

- Does the document adhere to the requirements and guidance provided in the relevant Evaluation Report Template (Annexes 7g, 7h, 7i)?
- Are key findings adequately supported by clearly referenced and triangulated evidence?
- Is there a clear and logical flow between the main questions, the respective findings, conclusions, and recommendations?
- Does the report meet the purposes of the evaluation and provide useful information for decision-making?
- Is the presentation clear and professional?
- Does the report meet the criteria set out in the Evaluation Report Quality Checklist?

7.7 UNOCT QUALITY EXPECTATIONS FOR EVALUATION REPORTS

The following provides an overview of what each section of the evaluation report is expected to cover with more details being provided in the Evaluation Report Quality Checklist (Annex 4).

	Report Structure and Presentation: The evaluation report is expected to be well structured, clearly written, and complete. It should be accessible in terms of language and include the use of visual aids to succinctly convey key information.
	Executive Summary: The Executive Summary needs to be able stand on its own, very concisely presenting the main results of the evaluation as well as a brief description of the intervention, the parameters of the evaluation and how it was carried out. As this is often the only part of the report that it is read in full, it is critical that it be well written. The recommendations are particularly important for decision-making; therefore, the Executive Summary must be complete enough to convey the credibility of the evaluation and to show the main findings that lead to and substantiate those recommendations.
	Introduction: This section presents the background and context of the intervention, including its links to the pillars of the Global Counter-Terrorism Strategy, the SDGs, and other strategic priorities of UNOCT and government partners. The intervention needs to be succinctly presented, along with the stakeholders involved, its theory of change (ToC), and main intended results.
	Evaluation Purpose and Scope: It is important that a clear description be provided of what the evaluation was intended to do, including why it was commissioned, what it covers, and who the intended audience is. This section should also explain how stakeholders, such as an Evaluation Reference Group, were involved in the conduct of the evaluation.

	Methodology: This section should provide an explanation and justification for the methodological approach utilized by the evaluation. It needs to describe the data collection methods, the data sources and how these were selected, how the data was analysed and the limitations of the study. The methodology should also address how rigour, credibility and ethical conduct were ensured, and how the process was human rights and gender responsive, and disability inclusive.
	Findings: The findings are typically considered to be the most important part of the evaluation report. They need to be clearly formulated, and address each of the evaluation criteria and questions. Findings should be based on rigorous analysis of multiple sources of data. The cause-and-effect links between the intervention and its intended and unintended results should be explained, and the reasons for accomplishments and/or lack of progress identified.
	Conclusions: These are expected to emerge from, and also add value to, the findings. They should present a comprehensive picture of the intervention's overall strengths and areas for improvement.
	Lessons Learned: This section is important for organizational learning. It should present the most significant lessons that can be drawn from the evaluation that are also useful for other programmes or contexts. They should capture not only good practice about why things worked but also learnings from what did not work.
	Recommendations: Recommendations need to align with the purpose of the evaluation and be clearly derived from the findings and conclusions. They also need to be useful. Usefulness is promoted when recommendations are clearly stated, realistic, and actionable.
	Cross-Cutting Themes: All UNOCT evaluation reports are expected to align with the UN-SWAP requirements for Gender Equality and the Empowerment of Women (GEEW) and Disability Inclusion (DI). They are also expected to address human rights, youth, and leave no one behind (LNOB). Guidance can be found in Chapter 2.

# CHAPTER EIGHT: GUIDE TO GOOD EVALUATION PRACTICE

The purpose of this chapter is to provide an overview of fundamental components of evaluation practice. It is not intended to be a comprehensive methodological resource but rather a general guide particularly for those with minimal evaluation experience who may be required to oversee such processes. It should also serve as a useful reminder to evaluation practitioners of the core evaluation concepts that need to be considered in any evaluation.

This chapter begins with an overview of planning frameworks and theories of change and the central role these have in evaluation processes. The next section discusses and provides commonly used examples of different types of evaluation design, methodological approaches, methods, and tools. There is then a brief discussion on the importance of the evaluation matrix as a tool for organizing and presenting the methodological information. This is followed by explanations of sampling and data analysis processes; areas which are often not adequately covered in evaluation reports. The chapter goes on to discuss evaluation complexity and suggests strategies that can help to address complex situations as well as methods and tools that are particularly relevant for assessing counter-terrorism and preventing violent extremism (CT/PVE) interventions. The final section provides selected resources that have more in-depth information on good evaluation practice generally, and resources more specific to the evaluation of CT/PVE.

8.1 THEORY OF CHANGE AND PLANNING FRAMEWORKS

Theories of Change, Results Frameworks, and Logical Frameworks all have the purpose of showing the cause-and-effect relationships between an intervention's activities and its intended results. There are different levels of results and a simple depiction of the chain between the inputs through to the ultimate result is shown in figure 8.1. To be useful for evaluation purposes, though, the relationships between all components of the results chain should be logical and the results stated in a way that can be easily measured.

Figure 8.1: Results Chain
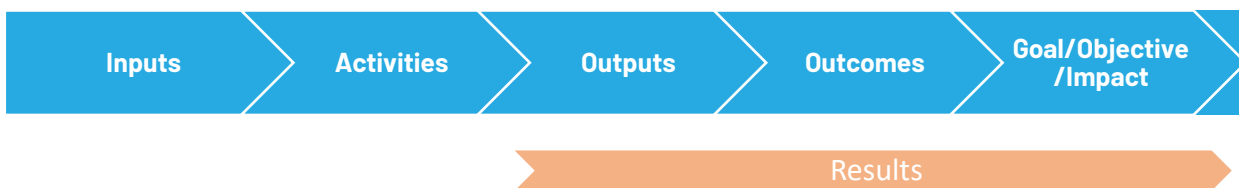
A theory of change (ToC) (also referred to as a programme theory) focuses on the change pathways, explaining how an intervention, or set of interventions, is expected to lead to the ultimate desired goal. It incorporates different levels of change, different actors, and different causal pathways. The ToC should also present contextual factors that help or hinder the envisioned change, such as economic conditions, gender dimensions, the human rights context, and the range of perceptions people hold about the topic. For evaluation purposes, it is important to make clear the assumptions on which change pathways are built (the conditions necessary for the change to happen but which are not under the control of the intervention). There are various ways to depict the ToC. Usually these are shown in diagram form (with different formats shown in figure 8.2) along with a narrative description. The level of detail also varies but the complexity of the presentation should align with the intended audience – for example, a diagram with numerous cascading boxes and interconnecting arrows may be appropriate for potential funders and the programme team, but it may be overwhelming to other audiences. On the other hand, a diagram that simply restates the intended results without depicting

contextual factors is less useful for communicating how the intervention is appropriate for that particular situation.

Figure 8.2: Examples of Different Styles of Graphic Depictions of Theories of Change
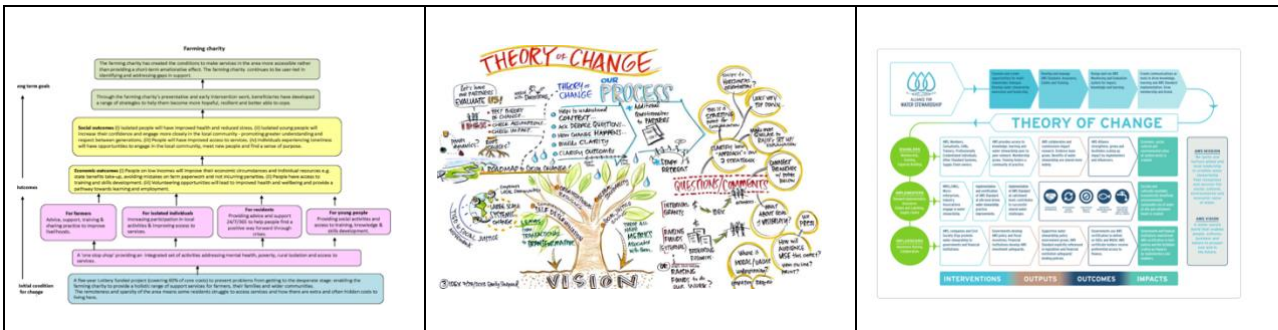
Although not commonly used in UNOCT interventions, **results frameworks** are another way of showing expected levels of change and are usually presented in matrix form. While the term is often used interchangeably with logical frameworks, results frameworks serve more of a strategic planning purpose. Their main intent is to make the causal logic of the intervention apparent by focusing on the different levels of results from short-, medium- to long-term, and the activities and outputs supporting them. Results frameworks typically have less contextual information than the ToC and less information on the measurement processes than what is included in logical frameworks.

A logical framework (or logframe) is a planning, management and monitoring tool that was originally created for military and space programmes before being adapted by USAID in the 1970s for use in the international cooperation sector. Earlier versions of the logframe were based on a simple grid but the tool has since evolved with different terminologies and features being applied in different organizations. Still, the main purpose continues to be to show how an intervention is to achieve its goal by drilling down into the hierarchy of results, activities, and inputs as well as how these will be measured or verified. The logical frameworks used by UNOCT (figure 8.3) are detailed matrices that state the intended impact and objectives and then provide the logic for each preceding level of result along with the indicator, baseline information, target, source and means of verification, and assumptions.

Figure 8.3 UNOCT Logical Framework Template

	Results Chain	Indicator	Baseline	Target	Source and Means of Verification	Assumptions
Impact Statement:						
Objective(s):						
Outcome						
Output						
Activities						

Use in evaluation - Both ToCs and logframes are central components of an evaluation process at UNOCT. The ToC generally provides the framework for determining the scope of the assessment, the evaluation criteria to be applied, the questions to ask, and the design and methods to be used. For example, if the ToC includes as a key assumption that civic organizations not directly involved in the project boost an advocacy campaign, then the scope of the evaluation should encompass a review of it and how such groups actually did provide support; this process is also referred to as examining the extent to which the assumptions held true.

Equally important to the evaluation is the logframe as it provides the basis for measuring success and is key for assessing the criterion of Effectiveness. However, evaluators need to take into account how the context may have changed since the logframe was developed and the assumptions on which it was based. It is important for evaluations to capture how and why an intervention may have diverged from the original plan (or in some circumstances, how it should have diverged but did not) and to consider unintended effects, both positive and negative. It is common for interventions to have impacts beyond what was initially envisioned, and these serve as important lessons for future programming.

Assessing and revising these tools - The evaluation team is normally expected to look at both the ToC and the logical framework during the Inception stage of the evaluation to ensure these are adequate for assessing the design and level of achievement of the intervention. For interventions without a ToC, the evaluators should develop one. Where shortcomings are found in existing ToCs and logframes, the evaluators should reconstruct them for the purposes of the evaluation. Ideally, this process should be done in a participatory process with the range of stakeholder groups. At a minimum, any proposed changes should be made in discussions with the evaluation manager and then be reviewed as part of the approval process of the Inception Report. Importantly, if there was not broad consultation on the development or reconstruction of the ToC, the new version can be used for the evaluation but should be more thoroughly discussed by the range of stakeholders before being formally adopted.

Key issues to look for in the ToC are whether there is a plausible progression between the various levels of results (are the pathways of change well-articulated?) and whether the assumptions and risks are made clear. It is also important that the theory addresses the human rights context, power dynamics and gender dimensions, and how these and other cross-cutting themes may affect the counter-terrorism interventions in specific cultural and institutional contexts.

Common shortcomings found in logical frameworks are (a) results statements that include terms such as 'strengthen', 'enhance', and 'capacity building'[48] which do not clearly state the change that is expected, (b) the use of indicators that are not precise enough to capture the respective result, (c) data not being available to support indicator achievement, and (d) straying from the generally accepted results terminology of:

> Inputs – what UNOCT provides (human and financial resources)
>
> Activities – what UNOCT does (technical assistance, training, provision of equipment, etc.)
>
> Outputs – what UNOCT delivers (products and services)
>
> Outcomes – what governments and other counterparts do differently (in part or in full due to UNOCT efforts)

---

[48] These types of terms are not easily measurable unless they are accompanied by specific metrics. For example, if the intent is to build capacity to enact a policy, a clear and measurable results statement would be "***develop*** *a documentation system that is used by officers on a daily basis to . . .* ". Another example is *"**increase** the rate of social contacts with people outside of community X by 10% each week"*. Other words that can be helpful in framing outcomes include 'expanded', 'eliminated', 'improved', 'lessened', 'prevented', 'reduced', 'shrunk', and 'raised'.

Goal/Impact/Objective – what governments and other counterparts achieve (in terms of bringing about changes)

Since the logical frameworks of programme and projects being evaluated will already have been recommended for USG approval by the UNOCT Programme Review Board (PRB), any revisions to logframes will require a change in the PD and re-submission to the PRB for review. The need for major changes in logframes should be reflected in the evaluation recommendations.

8.2 CHOOSING THE DESIGN, METHODOLOGICAL APPROACH, METHODS, AND TOOLS

There are a number of decisions to make about how an evaluation should be carried out so that it provides credible and useful information. This includes selecting the evaluation 'design', 'methodological approach', 'methods,' and 'tools'. These terms are often used interchangeably, which can cause some confusion. They each refer to how an evaluation process is implemented and can be considered as cascading categories with the first step being the determination of the evaluation design. The choice of which design, approach, and methods to use should primarily depend on the evaluation purpose, scope, and questions. However, the choice is also often constrained by the time and resources available to carry out the evaluation.

This section provides a brief explanation of these terms as well as examples of standard evaluation approaches and methods that are commonly used. There are many more options available and evaluators are encouraged to use more innovative evaluation processes to supplement the standard ones where that is feasible.

Evaluation Design

The evaluation design generally refers to the overall strategy for carrying out the assessment, the broad ones being:

Experimental design – involves the random assignment of subjects to treatment and non-treatment conditions and the pre- and post-measurement of each group. It is more commonly used in research than in evaluation processes although some impact evaluations will employ this design.

Quasi-experimental design – this design also aims to identify cause and effect relationships but takes into account the challenges of doing a full experiment in most real-life conditions. Instead of randomly assigning participants to different conditions, it involves the use of comparison groups and then post-measurement of each group. Most impact evaluations employ a quasi-experimental design.

Non-experimental design – considers the extent of change only for those affected by the intervention and does not involve a comparison group. It is the most commonly used evaluation design.

The evaluation design should be determined at the planning phase of the intervention to ensure adequate budgets and processes are in place for the evaluation to proceed. Programme Managers interested in using experimental or quasi-experimental designs are encouraged to contact the UNOCT BI Hub for advice on tested tools and approaches that may be applicable.

Evaluation Approach

The overall evaluation approach can be broadly categorized as qualitative, quantitative or mixed-methods. The terms qualitative ("QUAL") and quantitative ("QUANT") can also be used to describe the method and type of data. Both approaches bring important evidence to an evaluation process with

each having its advantages and limitations. For that reason, most evaluations apply a mixed-methods approach, and this is a requirement for all UNOCT evaluations.

Quantitative approaches use structured means to gauge and explain what is being studied. QUANT data is numerical or measurable, and comparatively easier to analyse. Although this type of data is sometimes perceived to be more objective than QUAL data, there are still subjective elements in designing QUANT data collection processes such as determining what variables to measure, how the questions are asked and answered, and the ways in which results are analysed and reported.

Qualitative approaches generally use more semi-structured techniques to gather descriptive data that helps to answer the 'how' and 'why' questions. QUAL data focuses on the constructed nature of social programmes and considers context, perspectives, and values as part of determining the results of an intervention. As such, it can provide more meaningful insights into an intervention's applicability and performance, and the lesson to be drawn. Qualitative processes can be designed in ways that provide quantifiable data, i.e., through the application of codes and the inclusion of ranking questions in interview protocols. The rigor of QUAL approaches is achieved through use of structured and systematic data analysis processes (such as coding and content analysis) and incorporating evaluator practices such as inter-rater reliability and self-reflection.

Mixed-method approaches, as the term implies, allow evaluators to utilize the advantages of both qualitative and quantitative types of approaches and data. Notably, quantitative data, even though often perceived as being more rigorous, usually needs to be explained by qualitative data.

Evaluation Methodology

The methodological approach (or methodology) also describes the conceptual approach and principles to be used, identifying how the data collection and analysis will be structured and undertaken. An evaluation can employ multiple approaches, the more common ones being:

Theory-based – this approach is typically explicitly based on the theory of change, using it to guide the evaluation. In addition to looking at the extent to which the intended results were achieved, it also looks closely at the pathways of change and whether the overall theory and the stated assumptions held true.

Participatory – this approach covers a range of processes but has at its core the principles of purposely engaging the range of intervention stakeholders, and in particular rights holders / beneficiaries, in conducting and making decisions about the evaluation. This term is often used incorrectly to refer to an evaluation process that simply engages stakeholders as respondents but with no opportunity for further input.

Case Study – an approach that has an in-depth focus on a unit (an individual, a geographic location, situation, or a specific activity/project) in its context and extracts learning from those cases that can be applicable to whole intervention. It is common for evaluations to use case studies in combination with other data collection methods in order to do a deeper dive into a specific aspect of the intervention under review.

Developmental Evaluation – an approach designed to support ongoing learning and adaptation, whereby the evaluator(s) are embedded for a period of time in the institutional structure of the intervention being evaluated.

Randomized Controlled Trials – an impact evaluation approach employing an experimental design that compares results between a randomly assigned control group and a group receiving the intervention.

Outcome Mapping – an approach that identifies (maps out) the steps that link the activities of the intervention to the outcomes that are defined in the theory of change.

Other approaches that are useful for evaluating complex interventions are highlighted in Section 8.6 below. These include Contribution Analysis, Outcome Harvesting, Most Significant Change, and Process Tracing.

Human rights-based and gender responsive approach – It is important for the chosen methodology to align with the overall aims of the evaluation. However, in addition to whatever other approach(es) are used, all UNOCT evaluations need to be human rights-based and gender responsive. This approach deliberately integrates human rights and gender equity (HRGE) dimensions, including the principles of equality, inclusion, respect for cultural sensitivities, and non-discrimination into both what is assessed and how the evaluation is carried out. These principles should be explicit in the design, methodology and analysis.

Good practice calls for relevant HRGE issues to be highlighted in the context section of the evaluation, for the assessment of HRGE to be part of the evaluation objectives and questions and therefore well discussed in findings, and for the key results of the analysis to be reflected in the conclusions and recommendations. The methodology section of the report should discuss what was done to ensure that the above principles were followed, for example:

- Ways in which the evaluation incorporates the perspectives and voices of a diversity of stakeholders.
- How disaggregated data was collected and analysed.
- What accommodations were made to facilitate everyone's participation such as the provision of interpreters, childcare and transportation to focus group discussions, and ensuring the location is accessible to those with mobility issues.
- What ethical considerations were attended to by the evaluators, including the risks to participation, and how the UNEG Ethical Guidelines for evaluations were followed.

Further details about human rights, gender equality, disability inclusion and leave-no-one behind in the context of evaluations are provided in Chapter 2 Evaluation in the UN System. UNEG Ethical Guidelines are also addressed in that chapter.

Evaluation Methods

The methods are the particular techniques used to collect and/or analyse data. Each evaluation will involve the use of multiple methods for triangulation purposes, the most common ones being:

Document Review – a basic method used in all evaluations, this encompasses an examination of materials produced by the institution, programme/project, or partners such as strategies, policy documents, programme documents, theory of change and results frameworks, annual workplans, technical and financial monitoring and progress reports, previous evaluation reports, etc.

Data obtained from document and literature review, including from online sites, is considered as being acquired from *secondary sources*.

Literature Review – this method looks beyond the documents that pertain directly to the intervention and were produced internally or by implementing partners to include relevant research papers, best practices and lessons, meta-synthesis reports, national development plans, UN Sustainable Development Cooperation Framework, etc. This can also include big data which can be obtained from a variety of sources such as social media platforms, national statistics, and third-party monitoring systems.

Surveys and questionnaires – these provide a standardized approach to acquiring information from a large number or diversity of stakeholders. They can be administered in-person or remotely and have the advantage of gathering QUAL and QUANT data on a range of topics relatively quickly. The respondents are able to be anonymous and the information obtained is generally easy to analyse. On the other hand, the descriptive responses may be quite brief, and this type of self-reporting can be subject to bias. Online surveys have the added disadvantage of typically low response rates while larger-scale surveys conducted in-person are resource intensive.

Interviews – semi-structured or structured discussions in which questions are asked by an evaluator to individuals or a group of interviewees about their impressions or experiences. As this format enables the use of probing and clarification questions, it is a way to obtain more information than what is generally provided in a survey. Interviews are also used as follow up to, or in conjunction with, other data collection methods.

Key Informant Interviews (KIIs) – these are typically in-depth interviews with stakeholders who have first-hand knowledge of the initiative's operations and context. They are the most common means of collecting data from managers and senior officials. Often KIIs are used amongst a smaller set of key informants at the beginning of the evaluation process to help identify how to frame subsequent data collection.

Focus Group Discussions (FGDs) – these are sessions facilitated by a moderator with a relatively small group of people (ideally five to eight people) on a limited set of topics. This term is often loosely applied to interviews with multiple people but in its truer form involves engagement and conversation that goes beyond direct questions and responses. The facilitation process can involve use of participatory techniques (such as those described below) to help participants feel comfortable and have freer flowing conversations about their experiences related to the intervention being evaluated. FGDs are most commonly used with rights holders/beneficiaries but can also be an effective way to engage decision-makers.

Observations – this involves careful observation and recording of information about how a programme operates and its observable results including its activities, processes, context, sites, and physical construction. To be considered as an evaluation method, the observation should be structured and recorded using a pre-agreed checklist.

Benchmarking – this method involves identifying outstanding practices from within the same organization or external sources as a way to rate or gauge internal performance.

Social media analysis – this refers to the gathering and analysis of unstructured real-time data from social media platforms. Such analysis can include practices such as web scaping whereby a web scraping software programme or bot crawls the internet and looks for data that fits predefined parameters. When it finds the data, the bot downloads, organizes, and displays it for the user. Social media platforms are also increasingly being used to collect raw data, in particular from youth populations.

Participatory methods – this term encompasses a wide variety of techniques to actively engage participants and generate open discussions that often bring out a wider range of ideas and perspectives than more conventional forums. Such methods are also used for starting conversations about challenging topics and building relationships between the evaluators and

respondents, as well as between respondents. Examples include ranking exercises and having participants draw diagrams, maps, timelines, and other visual displays to examine the study topics. Such techniques can be particularly effective in situations where there are differences in languages and literacy levels, where not all participants would otherwise feel comfortable contributing ideas, where the desired information is not easily expressed in words or numbers, or where a group would benefit from a change of pace from a structured question and answer format.

Evaluation Tools

The tools are the actual instruments used in data collection – for example, the questionnaires, survey forms, FGD protocols, and observation checklists. In most cases these will need to be tailored to the specific stakeholder group from which data is being collected. Tools can be bespoke, meaning they are prepared for that specific evaluation. Tools can also be standardized and validated instruments that have a scientific basis for measuring results.

The Office of Internal Oversight Services (OIOS) has developed useful guidance for preparing and conducting interviews and surveys that can be used in most types of evaluations; the links to these are provided in the Resources section at the end of this chapter. Suggested tools and methods that have been found useful in evaluating interventions relevant to UNOCT programming are discussed in the 'Dealing with Complexity' section below. As well, the UNOCT International Hub on Behavioural Insights to Counter Terrorism (BI Hub) is a source for tools that can be used to assess attitudinal and behavioural changes.

8.3 EVALUATION MATRIX

An evaluation matrix is the form used to capture the methods and tools that will provide the data for each evaluation question and sub-question. The matrix should also show the assumptions that are being assessed, the indicators or other metrics by which progress will be measured, and the sources of information for each method. Other elements can include the data analysis methods and the key findings for each question.

All UNOCT evaluation processes are to be guided by an evaluation matrix which is to be developed during the inception phase. The evaluation matrix is to be included in the Inception Report and in the Evaluation Report. A sample template is provided in Annex 7(e).

8.4 EVALUATION INDICATORS

Indicators are used for assessing the extent to which the intended results have been achieved. Indicators are expressed as benchmarks or variables usually in quantitative or qualitative form and should be specific to what is being measured, meaningful, observable, and measurable. Common types of indicators are those that look at incidence, prevalence, participation rates, attitudes, and behaviours.

An example of a clearly-stated and measurable objective-level indicator that would be found in a logical framework is, "*number of beneficiary countries that have Ministries assuming responsibility for API (Advance Passenger Information) and PNR (Passenger Name Record) data*". This type of information should be tracked by the intervention's monitoring system and the evidence should be readily available to the evaluation team. However, evaluators will also need to develop indicators specific to the evaluation questions beyond what is provided in the logframe. In most cases multiple indicators will be needed for each question to ensure triangulation. As with indicators found in the logical framework, those included in the evaluation matrix also need to be specific, meaningful, observable, and measurable.

> Triangulation involves the use of multiple methods and data sources to address an evaluation question. This cross-verification of evidence is an important way help ensure validity and credibility of the findings.

8.5 SAMPLING PROCESSES

It is usually not feasible for an evaluation to collect data from all stakeholders that are affected by an intervention or in all the locations from which it is implemented. Therefore, sampling is used to enable the evaluator to generalize findings from a representative selection to the broader population. Sampling can be used for a number of purposes including the selection of evaluation respondents, sites to be visited, and documents to be reviewed.

> Sampling is the process of selecting a representative subset of a population. It is an important part of all evaluations and needs to be clearly described.

Stakeholder Mapping

The basis for the sampling of respondents should be the stakeholder analysis or mapping process which identifies all of the stakeholders involved in an intervention along with their roles and relationships in the intervention. It should also include the ways in stakeholders will participate in the evaluation. Mapping is important for ensuring evaluations are inclusive of the range of duty bearers and rights holders, including marginalized and vulnerable populations. A template for use in this analysis is provided in Annex 7(d).

Sampling Approaches
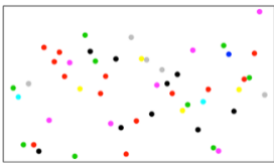
There are two broad categories of sampling approaches - probability and non-probability sampling, with examples of each discussed below.

Probability sampling is considered the most methodologically robust because it has the element of randomness, meaning that every case or person within the target area or population (known as the universe) has the same chance of being selected. In order for such a sample to be statistically representative, the number of units required in the sample must be calculated. Sampling error (the degree to which the sample may differ from the total population) can then also be derived[49]. On the other hand, a purposeful process is used in **non-probability sampling**. This has efficiency-related benefits as it generally requires fewer resources and enables evaluators to select the cases that are likely to be more relevant and information-rich. However, a purposeful approach can introduce a high level of bias. It is up to evaluators to determine the size of the sample, and it needs to be large enough to be credible given the purpose of the evaluation.

The sampling process should be structured to ensure the sample of evaluation participants is gender-balanced to the extent feasible, and adequately representative of the range of stakeholders. It is also important for evaluators to specify the sampling method used for each type of method and data source. This includes the selection of sites for field visits, the selection of informants for interviews and group discussions, and the selection of those receiving invitations to online surveys.

The most common types of probability sampling used in evaluations are a:

> Random sample - selection is by chance. All units having the same likelihood of being chosen thereby avoiding selection bias. This type of sampling is used in evaluations that have an experimental design.

> Stratified random sample - this is used when there are different groups that need to be represented in the sample based on factors such as gender, location, stakeholder group, budget size, economic situation, etc.

Typical types of non-probability sampling include:

---

[49] There are many open access online resources to do these calculations, such as Calculator.net

Convenience sample - participants are chosen based on factors such as their proximity, availability, or accessibility.

Opportunistic sample - as new opportunities arise in the course of data collection, data sources are strategically added to the sample. Such situations may occur during field visits when evaluators encounter individuals or situations that are particularly relevant and useful for inclusion in the study.

Purposive sample - the evaluator makes a judgement as to who to include as an evaluation participant from a certain population. This can be appropriate when only a limited number of people can serve as primary data sources.

Snowball sample - this strategy builds the sample size by asking existing evaluation participants to recommend other potential informants who would be useful to engage based on their knowledge or experience with the intervention. This is an effective way of finding key persons who are not otherwise known to the evaluators.

Further information on probability and non-probability sampling can be found at BetterEvaluation.

8.6 DATA ANALYSIS PROCESSES

Data analysis is used to make sense of the data. It involves systematically identifying trends, clusters, or other relationships amongst the typically large amount of information that has been collected. This aggregation and synthesizing of evidence is an essential part of the evaluation process with some of the more common methods being discussed here.

The Inception Report should lay out a clear plan for how the different types of data collected will be organized, triangulated, analysed, and presented. It is also important for the methodology section of the Evaluation Report to include a description for how each type of data was analysed. A common shortcoming is for this information to not be provided in the Evaluation Report.

Qualitative Data

QUAL data generally needs to be reviewed and coded before it can be analysed. Coding involves systematically organizing text-based information from field notes, interviews, documents, open-ended survey questions, etc. into manageable categories. A code is a label or value found in these texts that is usually related to the outcome indicators (such as key words or themes). The coding process can be done in steps that progressively develop unsorted data into more refined categories, themes, and concepts.

Content Analysis – this is the most common approach to analysing qualitative data. Once coded, the data is reviewed for trends, patterns, similarities, etc. Coding and content analysis can be done using software programmes such as Nvivo, Atlas and Dedoos, as well as databases such as Excel and Access. The process can also be done manually using, for example, different coloured markers or shading to highlight different themes in documents.

SWOT Analysis – this can be used as a very basic method of analysis. It involves assigning qualitative evidence into the four broad categories of strengths and weaknesses (internal factors which project implementors have some control over) and opportunities and threats (external factors that can significantly affect the success of the intervention). This process can also be used as a participatory tool for gathering data, whereby respondents are asked for their perspectives on an intervention according to each category, and as a tool for evaluators to draw conclusions from evaluation findings.

Quantitative Data

QUANT data needs to be cleaned before it is analysed. Cleaning is a quality control process that involves identifying and correcting or removing inaccurate, incomplete, or irrelevant records from the data that has been collected. This can be done manually or through computer-assisted processes.

Statistical Analysis – is the most common way of summarizing and analyzing quantitative data, and is used in particular for data obtained from surveys. Descriptive statistics are used to understand characteristics of the sample studied (i.e., average age, percentage exhibiting a certain behaviour, income range). Inferential statistics are used for testing hypotheses and drawing conclusions about a larger population set based on the sample by using processes such as T-Tests, confidence intervals and regression. Descriptive statistics are automatically generated by online survey tools. Excel or more robust statistical packages such as SPSS, Stata and SAS can be used for both inferential and descriptive analysis.

Converting Qualitative Data - in many cases, it is possible for numerical codes to be assigned in the process of coding qualitative data to produce quantitative values for statistical analysis. For example, identifying the number or percent of interviewees who perceived the management of the intervention favourably. This is a more precise way of presenting evidence compared to using vaguer statements such as 'most' or 'several' interviewees had favourable perceptions. Such a strategy provides further rigor in reporting of findings from qualitative methods.

8.7 DEALING WITH COMPLEXITY

Complexity is inherent in most of the work that UNOCT undertakes and has implications for its evaluations. Few interventions are simple and have all of the ideal elements that would enable an evaluation to be relatively straightforward – such as a clearly defined outcome, a single causal pathway, SMART indicators, robust monitoring data, and a stable implementation process.

Interventions that have multiple components are sometimes referred to as 'complicated' rather than 'complex'. Interventions can be **complicated** for many reasons. They may have multiple levels of implementation (global, regional &/or national), multiple implementing agencies with different agendas, long causal chains with many intermediate outcomes, and/or outcomes requiring multiple interventions. Although they may have many moving parts, the parts usually come together in predictable ways to produce a result. In such cases, evaluations need to be sufficiently sophisticated to capture all of the parts and interactions, and make clear the limitations of the process.

On the other hand, **complex** interventions are those which the results are less predictable because they involve emergent and causal processes that cannot be easily controlled. Features of complexity will be familiar to many UNOCT programme managers and include, for example, having to respond to constantly changing and possibly unsafe environments, uncertain funding streams, and new opportunities and challenges. When many different dynamics are at play, the pathways of change and intended results may be less clear. In such cases, evaluation processes may need to be more flexible with evaluators being attentive to emergent partners, strategies, and outcomes rather than solely focusing on the outcomes and targets identified during the intervention's planning stage.

Evaluation Approaches and Tools to Address Complexity

Evaluations will generally be more successful in addressing complexity if they go beyond the standard practice of interviews, group discussions, online surveys, and non-structured approaches to analysis. Greater robustness can be achieved by drawing on more sophisticated methodologies and by using

subject-specific tools to gather change-related data. Such practices also support the triangulation of findings.

Evaluators have a range of resources to draw from in designing evaluations to assess UNOCT interventions, including approaches that have been used for some time and their application to CT/PCVE is becoming more evident. Examples include:[50]

Contribution analysis - is useful for examining the contribution of an intervention to changes in complex environments and policy systems. It is designed to be used alongside theories of change and looks particularly at the pathways between the different levels of intended results.

Outcome harvesting – an approach whereby evidence of change is collected, and evaluators work backwards to determine whether and how an intervention has contributed to those changes. It can be applied to evaluations of programmes as well as to policy and institutional change, including when there are multiple variables that may be contributing to that change.

Most Significant Change (MSC) Analysis - a participatory evaluation methodology that involves collecting and selecting critical stories of change that are produced and analysed by a broad range of stakeholders. It emphasizes the 'lived experiences' and perspectives of individuals and communities most affected by the intervention.

Process Tracing - a qualitative analysis methodology. It looks at potential causes that have influenced a specific change or changes, and the strengths of evidence linking the cause to the change. It also involves testing alternative ideas about how change might have come about.

Tools and indicators that can measure change at an individual level - there is an increasing number of methodological approaches and tools being developed specifically for assessing CT/PCVE, including:

Building Resilience Against Violent Extremism (BRAVE) measure is a self-reporting tool that considers the key risk and protective factors associated with resilience to violent extremism. It is used in evaluations of strategies and programmes for strengthening resilience to violent extremism amongst young people and helps to show the effectiveness of culturally and contextually sensitive interventions. (Resilience Research Center, 2022)

PVE Change Capturing Tool provides a process and matrix to examine how changes brought about by an intervention have affected vulnerability or resilience, and the significance of this change. (International Alert and UNDP, 2018)

Countering Violent Extremism Evaluation Tool provides examples of programmes, evaluations and indicators for high-level outcomes related to resilience, diversion, disengagement, and capability. (New South Wales Government's Department of Communities & Justice, 2019)

Risk Analysis and Management

It is good practice for all evaluations processes to include an assessment of risks in undertaking the evaluation and to develop a risk management plan. This is even more important when dealing with

complex situations, particularly those involving sensitive subjects or that take place in conflict-affected areas where the safety, security and well-being of all evaluation participants and evaluation team members need to be given special consideration.

---

[50]A source for more information about each methodology is INTRAC, a non-government organization that focuses on building skills and knowledge of civil society including in the area of monitoring, evaluation, and learning (MEL).

Risk management involves proactively anticipating and understanding the risks that already exist and those that may emerge during the course of the evaluation and developing strategies for avoiding or minimizing them. This is important for helping to ensure the smooth running of the evaluation, and even more critical for ensuring the protection of human rights. The assessment should flag key ethical risks that may be posed by the conduct of the evaluation – with the United Nations Evaluation Group guidelines on the ethical conduct of evaluations being very clear on this point:

> *"It is necessary to achieve a compromise between the risks an evaluation exposes stakeholders to, on the one hand, and maintaining the social change objectives of the evaluation, on the other. **Every possible measure should be undertaken to ensure that no stakeholder be put in danger through an evaluation**. There are many types of harm to anticipate and consider in evaluations. Examples include discomfort, embarrassment, intrusion, devaluation of worth, unmet expectations, stigmatization, physical injury, distress and trauma. Political and social factors may also jeopardize the safety of participants before, during or after an evaluation. While 'do no harm' applies to all settings and all stakeholder groups, it is a particularly important concept in conflict settings and when working with the least powerful. In these circumstances, a double safety net needs to be in place."*

United Nations Evaluation Group, *Ethical Guidelines for Evaluation*, page 11

The table below provides an example of basic template for structuring a Risk Management Plan. In addition to ethical issues pertaining to data collection highlighted above, risks to address will be dependent on the type of evaluation and context but need to include data safety and security and could include interruptions to travel, national holidays that may limit access to stakeholders, difficulties in recruiting qualified data collectors, etc.

Table 8.1: Example of a Risk Management Plan Template

Foreseen Risk	Likelihood of occurrence	Potential Impact	Mitigation Strategies

Data Safety – In line with UNEG Ethical Guidelines, UNOCT requires evaluators to follow the principle of Do No Harm in collecting data. As part of the 2030 pledge to leave no one behind, this includes ensuring there is adequate representation from all groups of the population and adequate disaggregation in the analysis and presentation of data. Do no harm extends to adequate training of data collectors; obtaining informed consent from all respondents and being clear that they can opt out of answering any questions or continuing with the interview at any time; providing relevant physical, psychological and medical support for any vulnerable or at-risk populations identified during the evaluation; that safety and security includes zero tolerance for sexual harassment, abuse and exploitation, and; adherence to social distancing requirements, etc.

Evaluators must also ensure security of collected information before and after the evaluation and have protocols to ensure anonymity and confidentiality of sources of information. The information, knowledge and data gathered in the evaluation process must also be solely used for the evaluation and not for other uses.

Stakeholders' Limited Understanding of the Intervention

It is often the case that some key stakeholders, such as government officials and other partners, will only have a limited understanding of the entirety of the intervention being evaluated. This is even more likely when the intervention has multiple components or is being implemented in multiple countries or regions. Staff turnover can also be a factor in key contacts not knowing the background, scope or intended outcomes. This can make it difficult to obtain important data on some of the broader evaluation questions, particularly those related to relevance, efficiency, and sustainability.

In such situations it can be useful for the evaluation manager or the evaluators to prepare a briefing note on the intervention that can be shared with evaluation participants prior to or at the beginning of interviews and group discussions. Such a note could also be drawn upon for the introduction to a survey. It is important to tailor the description to the target audience and to avoid including information that could bias the recipients' responses.

Learning From Crisis Situations

Clearly, events such as the COVID-19 pandemic, wars that cause widespread economic and food supply crisis, and the increasing severity of climate change, have all created additional levels of complexity for programming and evaluations of these. COVID-19, in particular, forced the field of evaluation to adapt and much of the learning and amended practices that came from this time can be applied to other complex situations. Examples include the importance of:

- Revisiting the ToC, and in particular its assumptions, to ensure they reflect changed circumstances, and of developing a ToC if one does not already exist.

- Capturing and documenting changes that occurred in implementation areas, including for the purposes of creating new baselines, if necessary, to provide a basis for subsequent evaluations.

- Investing in support to evaluation capacity at the national level; that can be backed up remotely by international evaluators when needed.

- Seeking out and using the abundant amount of data sets that already exist and are rapidly expanding within and outside of the UN system (as per the "Data Strategy of the Secretary-General for Action by Everyone, Everywhere 2020-22")

- Carefully using remote data collection processes. Digital tools such as online meeting and survey platforms, smartphone and tablet apps, videos, and more all became firm features of the evaluation landscape during the pandemic. They continue to be highly useful in terms of cost savings and reducing the carbon footprint of evaluation activities. However, evaluators also have to consider when these are and are not appropriate or have to be used with caution. For example, thought needs to be given to whether online conversations can be held safely and confidentially (relating to issues of privacy and data security), accessibility of the technology (such as phone and SMS access/coverage being more prevalent that internet availability), infrastructure constraints (such as reliable electricity), and equity considerations (who might be excluded because they do not have phones or other devices). Attention also has to be paid to technical oversight and support, data collector training, and quality assurance processes. A useful resource for Information and Communication Technology (ICT) tools and their constraints is Capacity4dev.

- No matter the circumstance, ensuring all evaluations and evaluation activities have a do no harm approach and adhere to UNEG Ethical Guidelines for Evaluations. This includes ensuring the health and physical safety for stakeholders, national and international evaluators, and staff throughout the evaluation process.

8.8 FURTHER RESOURCES

Evaluation Practice in General

- Bamberger, M., J. Rugh, and L. Mabry. 2006. *RealWorld Evaluation: Working under Budget, Time, Data, and Political Constraints*. Thousand Oaks, CA: Sage Publications, 2006. This is one of the most widely used step-by-step evaluation textbooks on mixed-methods evaluation practice.

- Bamberger, M., J. Vaessen, and E. Raimondo, *Dealing with Complexity in Developmental Evaluation: a Practical Approach*. Thousand Oaks, CA: Sage Publications, 2015. This book offers practical guidance for policymakers, managers, and evaluation practitioners on how to design and implement complexity-responsive evaluations. In addition to a comprehensive overview of common evaluation tools and methodologies, the content covers more cutting-edge material, including approaches stemming from systems thinking and complexity sciences, the use of ICT in data collection and analysis, leveraging big data, and realist syntheses.

- BetterEvaluation. This online resource is the result of a global collaboration dedicated to improving the practice and theory of evaluation by sharing information about tools, methods, and approaches. It has information on hundreds of processes and tools and is the go-to site for many evaluation practitioners looking to quickly access information on a wide range of evaluation topics.

- Office of Internal Oversight Services (OIOS), Methodological Guidelines on Evaluation. The OIOS has produced practical and detailed guidance on several topics that are useful for evaluation managers and evaluators which are listed below[51]:

    1. Developing an Evaluation Terms of References (TOR), 2022
    2. How to Conduct Surveys, 2022
    3. How to Conduct Interviews, 2022
    4. Developing a Theory of Change, 2023
    5. Mainstreaming Gender Equality, Disability Inclusion, the Environment, and Human Rights in Evaluations, 2023

- United Nations Office on Drugs and Crime (UNODC), Guidance Note for Managers and Evaluators: Planning and Undertaking Evaluations in UNODC During the COVID-19 Pandemic and Other Crises, 2021. Although this was developed for UNODC, much of the content has broader applicability including a decision-tree for evaluation planning during a crisis, information on risks and mitigation strategies, and a section on data collection options during a crisis.

- Vaessen, J., S. Lemire, and B. Befani, *Evaluation of International Development Interventions: An Overview of Approaches and Methods*: Independent Evaluations Group of the World Bank, 2020. This guide provides easy reference to a broad range of evaluation approaches and methods. It is intended for managers and practitioners, including those conducting policy-related evaluations.

---

[51] These guidance documents can be accessed through the OIOS Evaluation Knowledge Platform – a SharePoint site dedicated and accessible to all UN Secretariat staff.

Evaluation Practice in the UN System

The final page of Chapter Two has links to guidance documents produced by UNEG and UN entities including on evaluation norms and standards, the integration of human rights and gender equity, disability inclusion, leave no one behind, intersectionality, and ethical principles for evaluation.

CT/PVCE Evaluation Practice and Tools

The following publications provide more in-depth information specific to evaluating CT/PCVE interventions:

- Clément, P., P. Madriaza and D. Morin, "Constraints and opportunities in evaluating programs for prevention of violent extremism: how the practitioners see it", UNESCO Chair in Prevention of Radicalization and Violent Extremism (UNESCO-PREV Chair): Sherbrooke University, 2021. The content is based on the experiences of 57 professionals involved in PVE efforts across the globe. Key recommendations emerging from this report are provided in the section on Unique Aspects of Evaluating PVE and CT in Chapter 3.

- United Nations Development Programme (UNDP) and International Alert, Improving the Impact of Preventing Violent Extremism Programming: a toolkit for design, monitoring and evaluation, Oslo, 2018. This toolkit includes a range of practical information including an indicator bank with numerous options for measuring programme achievement.

- United Nations Office on Drugs and Crime, UNODC Toolkit for Evaluating Interventions on Preventing and Countering Crime and Terrorism, (Vienna, 2021). In addition to methodological information this resource includes guidance and a useful checklist for conducting evaluations on sensitive subjects and in complex and conflict-affected environments.

- United Nations, Learn Better, Together: Independent Meta-Synthesis Under the Global Counter-Terrorism Strategy, (Vienna, 2021). Lessons from evaluating counter-terrorism related programming can be found in this meta-synthesis report. It was commissioned by the Sub-Group on Evaluation of the Resource Mobilization, Monitoring and Evaluation Working Group of the United Nations Global Counter-Terrorism Coordination Compact and provides collective results of 118 evaluations and oversight reports from Compact entities.