UNODC TOOLKIT for Evaluating Interventions on Preventing and Countering Crime and Terrorism
UNODC Toolkit
for Evaluating Interventions on Preventing and Countering Crime and Terrorism
The present evaluation toolkit was prepared by a team of evaluation experts under the guidance and supervision of the Independent Evaluation Section of the United Nations Office on Drugs and Crime (UNODC). The team of external experts consisted of Heather Ann Sutherland (team leader; evaluation expert), Mark Brown (crime prevention expert), Ashley Hollister (evaluation, inclusivity and gender expert) and Sarang Mangi (research and design).

The Independent Evaluation Section of UNODC may be contacted at the following:

United Nations Office on Drugs and Crime  
Vienna International Centre  
P.O. Box 500  
1400 Vienna, Austria  
Telephone: (+43-1) 26060-0  
Email: unodc-ies@un.org  
Website: www.unodc.org/unodc/evaluation.html

Disclaimer
The content of this publication does not necessarily reflect the views or policies of UNODC or contributory organizations, nor does it imply any endorsement.
## Contents

**INTRODUCTION** ........................................................................................................... V  

1. **THE MANDATE, STRUCTURE AND THEMATIC AREAS OF WORK OF THE UNITED NATIONS OFFICE ON DRUGS AND CRIME** ............................................. 1  
   One United Nations and the 2030 Agenda for Sustainable Development .......................... 4  
   Challenges and good practices in evaluating crime-related initiatives ................................. 7  
   Understanding the terminology ......................................................................................... 8  

2. **EVALUATION WITHIN THE UNITED NATIONS OFFICE ON DRUGS AND CRIME AND THE UNITED NATIONS SYSTEM** ......................................................... 9  
   Types of evaluations ......................................................................................................... 10  
   Evaluation standards ......................................................................................................... 11  
   Ethical practice in conducting UNODC evaluations .......................................................... 13  
   External quality assessment of evaluations ....................................................................... 14  

3. **EVALUATION DESIGN** ............................................................................................... 15  
   Alignment with the purposes of the evaluation ................................................................ 17  
   Developing or reformulating the theory of change ............................................................ 18  
   Preparing the evaluation matrix ....................................................................................... 19  
   Choosing evaluation questions .......................................................................................... 20  
   Selecting relevant indicators ............................................................................................ 20  
   Considerations for independent project evaluations ....................................................... 21  

4. **METHODOLOGY** ....................................................................................................... 23  
   Methods ............................................................................................................................ 24  
   Data tools and sources for different contexts ..................................................................... 30  
   Data coding and analysis ................................................................................................. 33  
   Ensuring human rights and gender equality responsiveness, and disability inclusion ........ 34  
   Resources ......................................................................................................................... 36  

5. **ADDITIONAL CONSIDERATIONS FOR EVALUATION IN COMPLEX ENVIRONMENTS** ................................................................. 37  
   Fragile contexts and conflict sensitivity analysis .............................................................. 38  
   Resources ......................................................................................................................... 39  
   Limited availability of data and hard-to-reach populations .............................................. 39  
   Resources ......................................................................................................................... 40
Engagement of youth........................................................................................................ 40
Language challenges .................................................................................................... 41
Stakeholders’ limited understanding of the intervention............................................... 42
Risk management in relation to security and the protection of human rights ............... 42

6. PRESENTATION OF RESULTS .................................................................................. 45

ANNEXES .................................................................................................................. 47

Annex 1. Areas of work of the United Nations Office on Drugs and Crime: key outcomes and outputs ................................................................. 48
Annex 2. United Nations Office on Drugs and Crime organizational chart and field offices .......... 49
Annex 3. Human rights-related instruments and provisions relevant to the work of the United Nations Office on Drugs and Crime ....................................... 51
Annex 7. Example theory of change: Regional Programme for South-East Asia and the Pacific, 2021–2026 ................................................................. 60
Annex 8. Indicative evaluation questions by criterion .................................................... 61
Annex 9. Checklist for conflict-sensitive evaluations and measurement on sensitive subjects ........ 63
Annex 10. Advantages and disadvantages of common data collection tools and approaches in the context of crime prevention ............................................. 66
Annex 11. Practical application of evaluation methods relevant to the areas of work of the United Nations Office on Drugs and Crime ................................................ 69
Introduction

The United Nations Office on Drugs and Crime (UNODC) Strategy 2021–2025 will "advance the central mission of UNODC – to contribute to global peace and security, human rights and development by making the world safer from drugs, crime, corruption and terrorism – and tailor our support in the pandemic and beyond. It will act as a lodestar as we forge new partnerships and strengthen existing ones; empower women and youth; and harness a culture of learning, evaluation and innovation, helping to build inclusive, equitable and just societies resilient to the threats of today and ready to meet the challenges and opportunities of tomorrow."

UNODC Strategy 2021–2025

The Independent Evaluation Section of UNODC contributes to the accountability and evaluation-based decision-making of the Office in its response to drugs, crime, corruption and terrorism by providing a wide variety of services and products, including the conducting of participatory, independent, inclusive and gender-responsive evaluations at the corporate, policy, programme and project levels, reporting on completed independent evaluations directly to the Executive Director of UNODC and to Member States. The Section cooperates with other evaluation functions within the United Nations system, in particular through the United Nations Evaluation Group, ensuring that UNODC evaluations fully meet the norms and standards of the Evaluation Group and are focused on utilization.

Against this backdrop, the present toolkit is intended to guide the design and implementation of independent evaluations of the work of UNODC in countering and preventing crime and terrorism. The toolkit is primarily aimed at evaluation experts who have limited familiarity with UNODC and its mandated areas of work, but also at those with substantive expertise in the Office's areas of thematic focus who could benefit from guidance in evaluation practice, as well as the evaluation community at large.

This toolkit is relevant for the two types of evaluations conducted by UNODC: (a) in-depth and strategic evaluations; and (b) independent project evaluations. It provides strategies for evaluating both the normative support provided by UNODC (e.g., support in developing legislation) and the Office's technical assistance functions (e.g., capacity-building). The toolkit is also intended to benefit other types of evaluation undertaken within and outside of UNODC.

Chapter 1 provides an overview of UNODC, including its five thematic areas of work and its different types of programmes, and how these connect with the overarching goals and priorities of the United Nations. This chapter links readers to international conventions and other instruments relevant to the Office’s mandate. It also highlights key challenges and good practices in evaluating initiatives to counter crime and terrorism.

Chapter 2 introduces the evaluation function and related expectations of UNODC. It provides links to the key resources for ensuring that evaluations are aligned with good evaluation practice in the United Nations system and the specific requirements of UNODC. The chapter describes ethical issues that need to be considered to ensure that the evaluation processes are carried out with respect and integrity, and in accordance with the principle of "do no harm".
Chapter 3 focuses on the design of the evaluation process. It highlights the foundational aspects of a good design, including having a theory of change and an evaluation matrix with sharply focused questions and indicators.

Chapter 4 suggests a number of methods and tools relevant for evaluating crime and terrorism prevention initiatives. Evaluators may find these useful for ensuring that data collection and analysis processes are adequately robust and capture perspectives from the Office’s diverse range of programme stakeholders.

Chapter 5 further examines how evaluations can be effectively carried out in complex environments. It covers fragile contexts and conflict sensitivity analysis, addressing factors such as limited availability of data and hard-to-reach populations, as well as language-related challenges and stakeholders with limited knowledge of the intervention being evaluated. It also highlights the importance of preparing a solid risk management plan at the design stage of the evaluation.

Chapter 6 focuses on the evaluation report. It guides the reader in establishing links between the key findings, conclusions, recommendations and lessons learned. The chapter also includes suggestions on ways to clearly communicate the main results of the evaluation work as a whole to ensure that the evaluation is as useful to the intended audience as possible.

Throughout this toolkit, references are provided to additional resources. These include the tools and additional information provided in the annexes, links to supplementary tools and templates available online, and links to guidance and reference documents available from UNODC and other organizations. The references are indicated by the three icons shown below.
CHAPTER 1.

The mandate, structure and thematic areas of work of the United Nations Office on Drugs and Crime

This chapter provides an overview of UNODC, including its five thematic areas of work, its different types of programmes, and how these connect with the overarching goals and priorities of the United Nations. The chapter provides readers with links to international conventions and other instruments relevant to the Office’s mandate. It also highlights key challenges and good practices in evaluating initiatives to counter crime and terrorism, and addresses relevant terminology.
UNODC was established in 1997 and adopted its present name in 2002. In terms of organization, UNODC is part of the United Nations Secretariat and thus acts on behalf of the Secretary-General and provides substantive services to the General Assembly, the Economic and Social Council, and various committees and conferences relevant to its thematic areas of work. It is governed by the Commission on Narcotic Drugs and the Commission on Crime Prevention and Criminal Justice. Its mission is to contribute to global peace and security, human rights and development by making the world safer from drugs, crime, corruption and terrorism by working for and with Member States to promote justice and the rule of law and build resilient societies. To achieve these aims, UNODC has developed three primary work streams: norms and standards, technical support, and research and analysis. Each work stream contributes to the Office’s five thematic areas of activity, as shown in figure I below. The key results for each area are shown in annex 1 and described in more detail in the UNODC Strategy 2021–2025.

When beginning an evaluation with UNODC, it is important to understand how the work under each of these five thematic areas is organized, where activities are undertaken and contributions are made (e.g., at field offices or headquarters) and towards which goals. This will assist the evaluator in mapping the sometimes complex linkages set out in programme documents.

---

To begin, it is important to realize how the work of UNODC is informed by its role as custodian of a variety of international instruments, standards and norms, and as the entity responsible for the implementation of certain United Nations resolutions. In relation to the thematic area of preventing and countering organized crime, for example, UNODC serves as the custodian of the United Nations Convention against Transnational Organized Crime and its three Protocols and, in addition, is in charge of implementing the international drug control conventions of 1961, 1971 and 1988. Implementation of these conventions is guided by the outcome document of the thirtieth special session of the General Assembly, entitled “Our joint commitment to effectively addressing and countering the world drug problem”, while monitoring of Member States’ progress in regard to such implementation is undertaken by the International Narcotics Control Board. Recalling the three work streams described in the UNODC Strategy 2021–2025, activity in this thematic area may therefore involve, for example:

(a) Normative work to assist Member States in the implementation of the Organized Crime Convention and the international drug control conventions, such as assistance in developing national legislation compliant with obligations under the Convention (e.g., the criminalization of certain acts or the establishment of appropriate punishments);

(b) Technical assistance to enhance the capacity of Member States in countering transnational organized crime and illicit drug use, such as assistance in improving customs and border inspection regimes (e.g., developing a canine drug detection capacity at airports and other transit points);

(c) Research and analytical work to understand the dynamics of organized crime, such as the identification of trends in trafficking or changes in the volume and value of drugs being moved along transit routes (e.g., the development of knowledge products such as the UNODC World Drug Report).

UNODC AND THE COMMISSIONS

In addition to operating as the governing bodies of UNODC, the Commission on Narcotic Drugs and the Commission on Crime Prevention and Criminal Justice are the foremost bodies within the United Nations system in the areas of drugs and crime, establishing norms and standards, discharging treaty-based normative functions and bringing Member States together for the purposes of policymaking. In that connection, every five years, Member States come together to discuss progress and share experiences and insights at a United Nations Congress on Crime Prevention and Criminal Justice. In March 2021, the Fourteenth Crime Congress was held in Kyoto, Japan. At the Congress, Member States adopted the Kyoto Declaration on Advancing Crime Prevention, Criminal Justice and the Rule of Law: Towards the Achievement of the 2030 Agenda for Sustainable Development, in which they reaffirmed their commitment to tackling issues of crime prevention and criminal justice in pursuit of the 2030 Agenda for Sustainable Development.

*www.unodc.org/documents/commissions/Congress/21-02815_Kyoto_Declaration_ebook_rev_cover.pdf*

---


3 Note that these are illustrative examples only, developed specifically for the purposes of this toolkit.
UNODC is headquartered in Vienna but most of its programmatic work is delivered through a network of eight regional offices, two liaison offices, two liaison and partnership offices, seven country offices, the Office of the Gulf Cooperation Council Region, and, at the time of writing, 94 programme offices. A map of field offices can also be found in annex 2, along with the UNODC organizational chart.

As an evaluator, it is important to understand where the programme you are charged with evaluating is situated within this structure and how it is supported by other departments and sections within the Office, as well as across the wider United Nations system. UNODC carries out projects and programmes at the country, subregional, regional and global levels. Therefore, evaluations at UNODC cover a range of different implementation modes, from smaller, individual projects at the country level to integrated programmes at the country and regional levels, as well as comprehensive global programmes. It is thus important for evaluators to recognize that other programmes might be working in the same thematic and geographical area as the one they are evaluating.

ONE UNITED NATIONS AND THE 2030 AGENDA FOR SUSTAINABLE DEVELOPMENT

In recent years, the United Nations has taken action to achieve even more coherence in its work. Much of this is channelled through a “One United Nations” strategy, so understanding how the programme you are evaluating connects with and contributes to the “One United Nations” approach will be important. The “One United Nations strategy is linked to the 2030 Agenda for Sustainable Development and its 17 associated Sustainable Development Goals. Each goal, such as Goal 16 (peace, justice and strong institutions), has associated targets (e.g., target 16.1, to significantly reduce all forms of violence and related death rates everywhere), and each of the targets has assigned indicators. UNODC serves as custodian of 16 of these indicators, across five separate goals. How does all this work in practice? In a given Member State, the United Nations country team brings all United Nations entities working in the country together with government counterparts to plan the Member State’s road map towards its commitments under the 2030 Agenda. The resulting agreement between the Member State and the United Nations country team is called the United Nations Sustainable Development Cooperation Framework, formerly referred to as the United Nations Development Assistance Framework. The Framework is used as a basis for aligning the programme you are evaluating with the needs and priorities of the Member State and with the “One United Nations” approach to programme delivery in the country. Such alignment is important for ensuring and maintaining the coherence of individual programmes and the overall mission of the United Nations.

Evaluations play a role in the achievement of the 2030 Agenda for Sustainable Development by helping to ensure that policies and programmes are aligned with the Sustainable Development Goals and corresponding targets, and their achievement. Evaluators should, as appropriate and realistic, aim to:

• Include performance indicators that connect with or strive to measure at least one Sustainable Development Goal indicator
• Gauge the connection of the intervention to Sustainable Development Goals in the assessment of the relevance criterion
Click on the following link to learn more about the connection between these goals and the work of UNODC.

UNODC tools and publications relevant to the Sustainable Development Goals

“Universal values are what enable the Sustainable Development Goals to be truly transformative, by placing the person and their inherent dignity at the heart of development efforts, empowering all people to be active partners in this endeavour.”

United Nations Sustainable Development Group

Underlying the Sustainable Development Goals is a set of universal values that all United Nations programmes and evaluations are expected to be aligned with. Recognition of these values is aimed at ensuring that the Sustainable Development Goals are implemented in accordance with international human rights law, eliminating gender inequalities and all forms of discrimination, and reaching out first to those who are the most in need, to ensure that no one is left behind. The values are reflected in the following three principles:

- Take a human rights-based approach
- Leave no-one behind
- Promote gender equality and the empowerment of women
Human rights law in particular places a number of obligations upon so-called “duty bearers”, that is, governments, agencies and other actors who must ensure that rights are respected or that conditions conducive to the enjoyment of those rights are maintained. Evaluators should be familiar with the rights relevant to specific programmatic areas, such as law enforcement or drug prevention, treatment and care, as well as those rights attaching to specific groups of rights holders, such as women, or children and youth.

Annex 3 provides a non-exhaustive compilation of key human rights instruments and relevant entitlements and protections relating to the mandated areas of work of UNODC.

RESOURCES: RELEVANT CONVENTIONS AND INSTRUMENTS

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

THE WORLD DRUG PROBLEM
The international drug control conventions of 1961, 1971 and 1988; Ministerial Declaration on Strengthening Our Actions at the National, Regional and International Levels to Accelerate the Implementation of Our Joint Commitments to Address and Counter the World Drug Problem of 2019

ORGANIZED CRIME

CORRUPTION AND ECONOMIC CRIME
United Nations Convention against Corruption; Political declaration adopted at the special session of the General Assembly against corruption held in 2021

TERRORISM
The 19 international legal instruments related to the prevention and suppression of international terrorism

CRIME PREVENTION AND CRIMINAL JUSTICE
Compendium of United Nations Standards and Norms in Crime Prevention and Criminal Justice; 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 list of conventions and other resources related to the cross-cutting issues of human rights, gender, youth and disability can be found in the following annexes:

Annex 1. Areas of work of the United Nations Office on Drugs and Crime: key outcomes and outputs

Annex 2. United Nations Office on Drugs and Crime organizational chart and field offices

Annex 3. Human rights-related instruments and provisions relevant to the work of the United Nations Office on Drugs and Crime
CHALLENGES AND GOOD PRACTICES IN EVALUATING CRIME-RELATED INITIATIVES

Challenges and good practices in evaluating crime-related initiatives, particularly those of UNODC, are the primary areas of focus in this toolkit. Notable challenges faced in conducting evaluations of UNODC interventions include lack of baseline data, missing or poorly constructed theories of change and indicators due to inherent challenges in measuring crime prevention work (especially interventions focused on change at the normative level), lack of common definitions and understanding of interventions and expected outcomes in specific areas of work, such as terrorism and violent extremism, and external constraints on the programme specific to the local conditions where UNODC works, including rapidly changing contexts and political and security priorities.  

One major obstacle that is not unique to UNODC is the limited ability to measure the impact of the crime prevention intervention. This is because interventions are typically highly dynamic and intersect multiple sectors, thematic areas, regions, programmes and activities, which can be represented by different divisions, branches and sections of UNODC. As such, it can be difficult to determine exactly what brought about change, especially when long-term outcomes are focused on the absence of criminal activity, as in the case of prevention, as opposed to an observable and measurable change that would typically be assessed in other types of development programmes, such as income.

Further challenges, and good practices and mitigation strategies for overcoming them, are described in this toolkit in chapter 3, on evaluation design and theories of change; chapter 4, on methodologies; and chapter 5, on conducting evaluations in complex environments. An additional resource recommended for review, available from the first link shown below, describes specific challenges and strategies for monitoring and evaluating programmes aimed at preventing violent extremism, including as related to monitoring (at p. 90), collecting data on violent extremism (at p. 104) and selecting survey methods in the context of prevention of violent extremism (at p. 111). Another useful reference is available at the second link below. It discusses the challenges and opportunities of evaluation from the perspective of practitioners in the area of preventing violent extremism, and their implications for evaluators.


Pierre-Alain Clément, Pablo Madriaza and David 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)

*As identified in the UNODC publication entitled *Meta-Synthesis of Evaluation Reports 2017–2018* and in quality assessments of UNODC evaluation reports.*
UNDERSTANDING THE TERMINOLOGY

Many subject matter-related terms used in UNODC programming have specific definitions that the evaluator will need to be familiar with. In most cases, such definitions can be found in the relevant sections of the UNODC website, and evaluators are urged to review this information carefully. For example, the following links may be useful:

- Terminology and information on drugs
- Transnational organized crime glossary of terms
- Working definitions of terrorism in the United Nations

Similarly, for evaluators who are subject matter experts but who are less familiar with evaluation practices within the United Nations system, the guides and handbooks cited throughout this toolkit provide precise and important definitions of key evaluation concepts and processes.
CHAPTER 2.

Evaluation within the United Nations Office on Drugs and Crime and the United Nations system

This chapter is an introduction to the evaluation function and related expectations of UNODC. It provides links to the key resources for ensuring that evaluations are aligned with good evaluation practice in the United Nations system and the specific requirements of UNODC. This includes ethical issues that need to be considered to ensure that evaluation processes are carried out with respect and integrity, and in accordance with the principle of “do no harm”.
The Independent Evaluation Section of UNODC is responsible for conducting evaluations in the Office’s mandated areas of work and ensuring that evaluation results are used to guide and improve programming and decision-making. The Section also aims to improve evaluation practice by:

- Providing updated tools, templates and guidance for use in the evaluation process
- Providing oversight and guidance, and ensuring that quality standards are met for all evaluation products and deliverables
- Managing and conducting independent evaluations in collaboration with external evaluators
- Ensuring that all evaluation processes, products and deliverables take into account human rights and gender equality standards
- Promoting a culture of evaluation and learning throughout UNODC, as well as throughout the broader United Nations system, and beyond

**TYPES OF EVALUATIONS**

UNODC carries out projects and programmes at the country, subregional, regional and global levels. Therefore, evaluations at UNODC cover a range of different implementation modes, from smaller, individual projects at the country level to integrated country and regional programmes, as well as comprehensive global programmes. According to the UNODC Evaluation Policy (2015) there are two distinct types of evaluations:

- **In-depth and strategic evaluations**, which are evaluations of relevance to the organization, such as evaluations of country, regional, global or thematic programmes, or cross-cutting issues, for example, gender, human rights or corporate policies
- **Independent project evaluations**, in which the object of analysis is an individual project designed to achieve specific objectives within specified resources, in an adopted time frame and following an established plan of action

---

Further information on the Independent Evaluation Section of UNODC, as well as templates, guidelines and other evaluation tools, can be found at [www.unodc.org/evaluation/index.html](http://www.unodc.org/evaluation/index.html).
The Independent Evaluation Section is also responsible for managing the evaluation of UNODC projects funded through the United Nations Development Account. The Development Account is a capacity development programme of the United Nations Secretariat aimed at building capacities of developing countries in the priority areas of the United Nations development agenda. Evaluators of Development Account projects need to be mindful of the different focuses of the various entities involved and aware that the expectations for evaluations of both must be met. Under most circumstances, if the UNODC evaluation requirements are met, the Development Account standards will be as well.

**United Nations Development Account Evaluation Framework**

**United Nations Development Account Project Evaluation Guidelines**

### Evaluation Standards

All evaluations undertaken within the United Nations system need to take into account, in addition to the 2030 Agenda for Sustainable Development, specific principles, criteria, norms, standards and other considerations aimed at ensuring that evaluations are conducted in an ethical and inclusive manner. These rubrics were established by the United Nations Evaluation Group and the Development Assistance Committee of the Organization for Economic Cooperation and Development (OECD). Links to the key documents and requirements that all evaluators should be familiar with are listed below.

### Resources

- United Nations Evaluation Group norms and standards
- Evaluation criteria of the OECD Development Assistance Committee (2019)
- United Nations Disability Inclusion Strategy (2020)
- Guidance for mainstreaming gender in UNODC evaluations

---

6 The United Nations Evaluation Group is an inter-agency professional network of units and offices of the United Nations system responsible for evaluation, including United Nations departments, specialized agencies, funds and programmes, and affiliated observers.
ETHICAL CONSIDERATIONS

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. The United Nations Evaluation Group publication entitled *Ethical Guidelines for Evaluation* sets out four guiding ethical principles for those organizing and conducting evaluation activities (see figure III). The publication also includes a checklist of ethical issues to consider during each evaluation phase, as well as a Pledge of Ethical Conduct that all members of evaluation teams are encouraged to review and sign. The Pledge and more details about the four ethical principles can be found in the annexes.


**Figure III. United Nations Evaluation Group guiding ethical principles for evaluation**

**Integrity** is the active adherence to moral values and professional standards, which are essential for responsible evaluation practice.

**Accountability** is the obligation to be answerable for all decisions made and actions taken; to be responsible for honouring commitments, without qualification or exception; and to report potential or actual harms observed through the appropriate channels.

**Respect** involves engaging with all stakeholders of an evaluation in a way that honours their dignity, well-being and personal agency while being responsive to their sex, gender, race, language, country of origin, status as lesbian, gay, bisexual, transgender, queer or intersex, age, background, religion, ethnicity and ability, and to cultural, economic and physical environments.

**Beneficence** means striving to do good for people and planet while minimizing harms arising from evaluation as an intervention.

ETHICAL PRACTICE IN CONDUCTING UNODC EVALUATIONS

UNODC evaluators are expected to adhere to the highest standards of conduct and to uphold and respect the purposes and principles set out in the Charter of the United Nations, including the determination to reaffirm faith in fundamental human rights, in the dignity and worth of the human person and in the equal rights of men and women.

It is mandatory that the United Nations Evaluation Group ethical guidelines be followed in UNODC evaluations and that the methodology section of both the inception and evaluation reports includes an explanation of the ethical aspects that have been taken into consideration. This means it is necessary to go beyond simply stating that United Nations Evaluation Group guidance was followed by being explicit about how the relevant issues from the ethical guidelines were addressed. This could include noting and attaching the signed Pledge of Ethical Conduct, as well as describing how issues such as data security and highly sensitive information, informed consent, confidentiality, and the principle of “do no harm” were addressed. In UNODC evaluations, the stakeholders are often high-level government officials who are being asked about issues that are potentially 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. 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.

Useful tips for avoiding conflict of interest and bias can be found in the International Development Evaluation Association (IDEAS) publication *Evaluation in Contexts of Fragility, Conflict and Violence: Guidance from Global Evaluation Practitioners* (Exeter, United Kingdom, 2021), p. 28.

Ethical considerations also extend to how we understand and refer to the different stakeholders involved in the intervention being reviewed. UNODC encourages the use of rights-based terminology that reflects the human rights-based approach to its work. For this reason, those whom the intervention is intended to assist are referred to as “rights holders” and the institutions obligated to fulfil the holders’ rights are referred to as “duty bearers”.

Annex 3 of the *United Nations Evaluation Group Ethical Guidelines for Evaluation*: checklist of ethical issues for those who conduct evaluations to consider during each evaluation phase.
More information about the human rights-based approach and rights-based language can be found at the following links:

- UNODC, UNODC Toolkit for Mainstreaming Human Rights and Gender Equality into Criminal Justice Interventions to Address Trafficking in Persons and Smuggling of Migrants
- OECD, “The human rights-based approach to development cooperation: towards a common understanding among the United Nations agencies”
- United Nations Sustainable Development Group, “Human rights-based approach”
- UNFPA, “HRBA checklist of questions”

EXTERNAL QUALITY ASSESSMENT OF EVALUATIONS

To help ensure that evaluations meet the full range of agreed standards, all evaluations published by UNODC are externally assessed for quality. The results of the assessments are made available on the UNODC website alongside the evaluation reports. It is highly recommended that evaluators review the evaluation quality assessment template (available at the below link), particularly during the design and report-writing stages, as it is a useful reminder of how to apply good practice throughout the evaluation process.

- UNODC evaluation quality assessment template
CHAPTER 3.

Evaluation design

This chapter focuses on designing the evaluation process. It highlights the foundational aspects of a good design, including the importance of aligning the evaluation framework with the purpose of the evaluation, as well as having a theory of change and an evaluation matrix with sharply focused questions and indicators. It also addresses the different expectations for evaluation processes that are smaller in scale.
The UNODC Evaluation Handbook is a good starting point for understanding the fundamental components in designing an evaluation that meets UNODC standards. Evaluators more accustomed to drawing on methods such as randomized controlled designs, quasi-experimental designs or the use of complex multivariate statistical controls to isolate programme effects will especially benefit from understanding how mid-term and final evaluations within UNODC utilize a different range of techniques grounded in triangulation logic to understand programme outcomes and impacts and any relevant factors affecting programme performance.

Evaluations within the United Nations system are normally required to apply a broad lens to assessment, guided by the set of six criteria of the OECD Development Assistance Committee, as well as principles of human rights, gender equity, disability inclusion and “leave no one behind”. The present chapter covers key considerations that evaluators need to take into account during the design stage. This includes the importance of aligning the evaluation criteria and the key questions to be answered with the purpose of the evaluation, assessing and utilizing a theory of change, and then using these tools to develop an evaluation matrix that provides the road map for data collection and analysis. The components are then incorporated into the inception report.

Additional resources for guidance on evaluation design and methods include the following:

- BetterEvaluation.org
- IDEAS, Evaluation in Contexts of Fragility, Conflict and Violence: Guidance from Global Evaluation Practitioners (2021)
- InterAction, Gender-based Violence Prevention: A Results-based Evaluation Framework (May 2021)

INCEPTION REPORTS

Although the terms of reference for an evaluation normally include a proposed evaluation design, final decisions about the approach and methodological process of the evaluation are made during the inception phase. Depending on the type of evaluation, this is done in consultation with the project or programme manager, the Independent Evaluation Section, and, if applicable, the evaluation reference group.

Any changes to the evaluation design as described in the terms of reference need to be noted in both the inception and evaluation reports. The evaluation report should also capture any modifications made after the inception report was approved.
ALIGNMENT WITH THE PURPOSES OF THE EVALUATION

Seasoned evaluators are likely to have come across evaluation terms of reference that include an extensive list of criteria and questions to be addressed. Although useful for communicating the general direction the evaluation should take, following these directly can lead to an unwieldy process in which numerous issues are explored but the depth of analysis is compromised. In contrast with past practice, whereby evaluations were expected to consider all of the standard OECD Development Assistance Committee criteria, as well as any agency-specific and gender mainstreaming criteria, expectations have now shifted. Although most in-depth and strategic evaluations will cover all criteria, there is more flexibility with other types of evaluations commissioned by UNODC. The guidance issued by the OECD Development Assistance Committee in 2019, “Applying Evaluation Criteria Thoughtfully”, emphasizes the importance of the following:

- **Context.** Consider the context of the individual evaluation, the intervention and the stakeholders involved. The evaluation questions (what you are trying to find out) and what you intend to do with the answers should inform how the criteria are interpreted and analysed.

- **Purpose.** The criteria should not be applied mechanistically. Instead, they should be used according to the needs of the relevant stakeholders and the context. Depending on the purpose, more or less time and resources may be devoted to the evaluation of the individual criteria. The availability of data, resource and time constraints and methodological considerations may also influence how and whether a particular criterion is used.

---

**Figure IV.** OECD Development Assistance Committee criteria that serve as the basis for evaluative judgments

Source: [www.OECD.org](http://www.OECD.org)
It is equally important for the evaluation questions to be restricted to those that align with the main purpose of the evaluation and the availability of data, resources and time. The questions will then determine the methods available to the evaluators. One of the key outcomes of the inception phase of an evaluation should be agreement with UNODC on the essential criteria and questions to be used. As a general rule, the fewer the number of questions, the more focused the evaluation can be. The final evaluation of the implementation of the Global Programme on the Implementation of the Doha Declaration: Towards the Promotion of a Culture of Lawfulness (GLOZ82), published in 2020, provides an illustrative example of a distilled set of questions that helped to ensure a focused evaluation process.

**Annex 6. Example of a succinct set of United Nations Office on Drugs and Crime evaluation questions**

**DEVELOPING OR REFORMULATING THE THEORY OF CHANGE**

At the inception phase, evaluators will normally need to allocate time to the theory of change for the intervention. This could entail developing a theory of change if one is not available, assessing the adequacy of an existing one, or creating a unified theory of change, if conducting a broad thematic or strategic evaluation involving multiple projects. Similar to a logical framework and aligned with a results-based management approach, a theory of change shows the connections between the intervention’s activities and the results they are intended to produce. However, unlike the table format of logical frameworks, theories of change are presented in diagram and/or narrative form and can more clearly depict different levels of change, different actors and multiple causal pathways. Theories of change generally include contextual factors, such as the role of the State versus civil society or the different values people hold with regard to the topic that help or hinder the envisioned change and the assumptions on which it is built (e.g., the conditions that are necessary for the change to occur but that are not under the control of the implementers).

For an evaluation to proceed, it is important to have a solid theory of change to work from because it helps the evaluator to know what to look for, including what the intent of the intervention is and what its parameters are. However, there are instances in which the intervention may not have been as carefully planned as it might have been or in which circumstances have changed and the logic no longer holds up or, perhaps, the intended results as stated are not measurable. Therefore, a key and often omitted step during the inception phase is to assess the theory of change to see if adjustments need to be made to ensure that it can provide a solid framework on which the evaluation can be based. The review of programme documents and interviews at the inception phase can be structured to help evaluators assess the current theory of change, as well as to reformulate and test revised versions. Any changes to the theory of change should be agreed upon with the key staff and other stakeholders (including the evaluation reference group, if one exists) prior to the start of the evaluation. Moreover, unless the process of consultation concerning the revisions is conducted broadly, the new version of the theory of change should only be used for the purposes of the evaluation and should be more thoroughly discussed by the range of stakeholders before being formally adopted.
Resources and guidance for developing theories of change can be found in the UNODC Evaluation Handbook and in the toolkit highlighted in the box below.

**UNDP AND INTERNATIONAL ALERT, IMPROVING THE IMPACT OF PREVENTING VIOLENT EXTREMISM PROGRAMMING: A TOOLKIT FOR DESIGN, MONITORING AND EVALUATION (OSLO, 2018)**

This toolkit contains useful step-by-step guidance for developing theories of change (see pages 52–58), with examples relevant to the work of UNODC and a checklist of elements to consider in the design process. Page 73 of the toolkit includes suggestions for different levels of change one might expect to see in a project working with returned foreign fighters in prisons, along with suggested indicators to use for measuring achievement.

**THEORIES OF CHANGE FOR UNITED NATIONS OFFICE ON DRUGS AND CRIME INTERVENTIONS**

The theory of change diagram for the Regional Programme for South-East Asia and the Pacific, 2021–2026, can be found in annex 7. It provides a broad overview of the intended chain of results of a multifaceted programme and key underlying assumptions. Its stated outcomes, along with measurable indicators from the logical framework, provide the foundation for future evaluations.

Annex 7. Example theory of change: Regional Programme for South-East Asia and the Pacific, 2021–2026

Further examples can be found in the following links to UNODC evaluation reports. They exemplify a range of ways in which theories of change can be formulated. Good practice is to include both a visual representation of the change process and a narrative description with a more detailed explanation.

- **Evaluación independiente de proyecto/evaluación en profundidad de “Fortalecimiento para la Seguridad de Grupos en situación de vulnerabilidad” (MEXZ93-2018-005), p. 54 (in Spanish)**
- **Final Independent In-depth Evaluation of the Global Action against Trafficking in Persons and Smuggling of Migrants (GLO.ACT) (GLOZ67) (December 2019), pp. 8–9**
- **Independent Project Evaluation: Improving the Criminal Justice Response to Violence against Women in Egypt (EGYZ33) (March 2019), pp. 5–6**
**PREPARING THE EVALUATION MATRIX**

Another key piece in the preparation of the inception report is the development of an evaluation matrix. The matrix is a planning and analysis tool used to ensure that the evaluation process addresses the key questions in a sufficiently robust manner. Also referred to as an evaluation framework, the evaluation matrix should clearly show what the basis is for judging the extent of achievement (the indicators) for all evaluation criteria and key questions, how the data will be collected and how triangulation between data sources will be accomplished. The matrix is additionally used as the basis for designing the data collection instruments, because the table format makes it relatively easy to see all the different types of information that need to be collected from each stakeholder group. In addition to being integrated into the inception report, good practice calls for the matrix to also be attached to the final evaluation report. In the final report, some evaluators will include a summary of data collected for each evaluation question to clearly illustrate the link between the data and the findings.

The in-depth thematic cluster evaluation of UNODC law enforcement and border control projects in Central Asia: XAC/Z60, TAJ/E24, TAJ/H03, RER/H22 and XAC/K22 (November 2018) provides a useful example of an evaluation matrix that clearly specifies the indicators for each question and how the data will be collected. It also includes a brief summary of findings for each question (see pages 97–109).

**CHOOSING EVALUATION QUESTIONS**

Although the terms of reference document will suggest a set of questions to be answered by the evaluation, it is normally the evaluator’s responsibility to review them carefully during the inception stage and to suggest any needed revisions. The questions establish the trajectory the study will take, so it is important that: (a) they fit the evaluation’s purpose and learning objectives; (b) they are clear and specific; (c) they are feasible to answer in a meaningful way with the resources available and within the page limit for the report; and (d) everyone agrees on the final set. The annexes include a list of indicative questions for each of the evaluation criteria.

Annex 8. Indicative evaluation questions by criterion

**SELECTING RELEVANT INDICATORS**

One or more indicators will be needed to answer each evaluation question. Indicators serve as the basis for determining the extent of an intervention’s achievements and point to the type of data that should be collected. If the intervention’s logical framework has been well developed, it will include measurable indicators that will be used in the evaluation process, particularly for the questions under the criterion of effectiveness. Evaluators will need to review those indicators to make sure they remain relevant and are feasible to use in the evaluation process, particularly in conflict-affected or fragile settings. Normally, additional indicators will need to be developed. In some cases, these can be drawn from existing sources, a benefit being that the data...
produced can be more robust, as they can be compared to those of other studies. This is particularly important for UNODC, as evaluations are an important basis for building aggregate evidence of progress towards its strategic objectives. Ideally, all evaluations will use some indicators that align with the outcomes and outputs in the UNODC Strategy 2021–2025.

Useful resources for developing indicators and finding existing ones relevant to the work of UNODC include the following:

- InterAction, *Gender-based Violence Prevention: A Results-based Evaluation Framework* (May 2021), module 3.2
- European Crime Prevention Network, “Criteria for the evaluation of crime prevention practices” (October 2016)
- United Kingdom Gender Action for Peace and Security (GAPS) network and others, “Beyond consultations: a tool for meaningfully engaging with women in fragile and conflict-affected States” (2020)
- Center on Global Counterterrorism Cooperation, “Evaluating countering violent extremism programming: practice and progress” (2013), p. 8

**CONSIDERATIONS FOR INDEPENDENT PROJECT EVALUATIONS**

Independent project evaluations are expected to be smaller in scale compared with in-depth and strategic evaluations, while still meeting United Nations Evaluation Group norms and standards. As such, the terms of reference for independent project evaluations will generally set out a narrower scope for the evaluations and provide a more defined methodological approach. Table 1 highlights the characteristics of a process, as well as the standards that independent project evaluations are nevertheless expected to meet, despite being of a more limited scope.
TABLE 1. CHARACTERISTICS OF SMALLER-SCALE EVALUATIONS

<table>
<thead>
<tr>
<th>ASPECT OR COMPONENT OF THE EVALUATION</th>
<th>TYPICAL CHARACTERISTICS IN INDEPENDENT PROJECT EVALUATIONS COMPARED WITH IN-DEPTH EVALUATIONS</th>
<th>STANDARDS TO BE MET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inception report</td>
<td>Briefer and primarily focused on methodological approach</td>
<td>UNODC evaluation norms and standards for independent project evaluations</td>
</tr>
<tr>
<td>Design and methodology</td>
<td>• Comprehensive theory of change not required</td>
<td>• Basic results chain still needs to be described</td>
</tr>
<tr>
<td></td>
<td>• Fewer OECD Development Assistance Committee criteria and evaluation questions to be covered; limited to those most critical to the evaluation purpose</td>
<td>• Human rights, gender equality and disability to be included as cross-cutting issues</td>
</tr>
<tr>
<td></td>
<td>• Basic results chain still needs to be described</td>
<td>• United Nations Evaluation Group ethical standards for evaluations to be adhered to and their application explained</td>
</tr>
<tr>
<td></td>
<td>• Human rights- and gender equality-responsive evaluation practices to be adhered to and their application explained</td>
<td>• Human rights- and gender equality-responsive evaluation practices to be adhered to and their application explained</td>
</tr>
<tr>
<td>Data collection and analysis</td>
<td>More focused, while still ensuring rigorous and solid data collection and analysis</td>
<td>United Nations Evaluation Group norms and standards for evaluations</td>
</tr>
<tr>
<td>Evaluation report</td>
<td>• Shorter reports, generally not exceeding 15 pages, excluding the executive summary and annexes</td>
<td>• United Nations Evaluation Group norms and standards for evaluations to be upheld</td>
</tr>
<tr>
<td></td>
<td>• No evaluation briefs</td>
<td>• Structure to follow independent project evaluation report template</td>
</tr>
</tbody>
</table>
CHAPTER 4.

Methodology

This chapter suggests a number of methods and tools relevant to evaluating crime and terrorism prevention initiatives. Their use can help to ensure that data collection and analysis processes are adequately robust and capture perspectives from the Office’s diverse set of programme stakeholders. Further details on each can be found in the annexes. The chapter also highlights several examples of good evaluation practice, as well as software options for data coding and analysis. It concludes with a discussion on ways to ensure that your evaluation is human rights- and gender equality-responsive and disability inclusive.
The purpose of the present chapter is to provide evaluators with an introduction to various evaluation methods and data collection tools that will define the overall methodology in any given evaluation for UNODC. The United Nations Evaluation Group norms and standards lay the foundation for the defining of evaluation methods, specifically requiring that the credibility of evaluations be established through the use of sufficiently rigorous methodologies that ensure that the evaluation responds to the scope and objectives, is designed to answer evaluation questions and leads to a complete, fair and unbiased assessment (standard 4.5, ”Methodology”).

Typically, the terms of reference for the evaluation will provide preliminary guidance on the preferred methodologies for ensuring a sufficient level of rigour, which will vary depending on the evaluation type (i.e., in-depth and strategic evaluations, independent project evaluations or cluster evaluations) However, evaluation teams are provided a further opportunity to elaborate on the methodologies and tools during the inception phase. For this reason, it is the responsibility of evaluators to select multiple methods, typically including both quantitative and qualitative methods, to allow for the collection, analysis and triangulation of data across methods, as well as across valid and reliable sources.

**METHODS**

In addition to the type of evaluation, its scope and the questions guiding the selection of methods, the method of evaluation is also an important consideration, as particular evaluation methods may be more useful than others in the context of the evaluand’s crime prevention strategy and associated outcomes. For example, if an objective of the evaluation is to determine the effect of an intervention on the incidence of offences (or reoffences) among perpetrators in organized crime cases, or rates of reduced recidivism or drug use among offenders, obtaining quantitative data using an experimental or quasi-experimental design would be the most appropriate. However, ethical, data, time, resource or other constraints may make it unsuitable to select such methods. In addition, in undertaking impact evaluations in crime prevention, challenges are typically presented in the assessment of attribution, as it is rarely possible to isolate the effects of one project, programme or policy from other anti-crime initiatives. This is especially true in the context of the United Nations, which typically works at the national, regional or global levels with multiple and diverse stakeholders, rather than in isolated communities.

---

**THOUGHTS ON CONTRIBUTION AND ATTRIBUTION**

“A systems approach to effectiveness also helps us to move beyond the ‘attribution’ versus ‘contribution’ challenge from a monitoring and evaluation perspective. Rather than asking whether a different ‘end state’ at the macro level has been created (and can be attributed to a specific programme), it asks whether and how change (facilitated by a specific programme) in one part of the system influences – or fails to influence – other parts. This focuses the question of accountability for impact not on the achievement of a planned end state, but rather on whether and how programmes have leveraged change in the system, and therefore are likely to affect how the system behaves – hopefully in positive ways.”


For UNODC evaluations, the methods will, generally, be non-experimental and will be adapted to the necessary process for gathering and analysing data in order to respond to evaluation questions. Table 2 below provides the most common evaluation methods used in programmes with thematic focuses similar to those of
UNODC; however, most have yet to be systematically applied in UNODC evaluations, and thus the examples presented in the table are largely drawn from publications issued by external organizations. In addition, annex 11 provides links and citations to evaluations in which the various methods have been successfully used to measure outcomes corresponding to those often measured in UNODC evaluations or among populations of concern to UNODC.

### TABLE 2. EVALUATION METHODS RELEVANT TO UNODC AREAS OF WORK

<table>
<thead>
<tr>
<th>METHOD</th>
<th>ADVANTAGES AND LIMITATIONS FOR USE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONTRIBUTION ANALYSIS</strong></td>
<td>Advantages:</td>
</tr>
<tr>
<td></td>
<td>• It can be used to assess the process or mechanism by which an intervention led to or contributed to outcomes</td>
</tr>
<tr>
<td></td>
<td>• It is useful for complex social, economic and political systems in which assessment of sole attribution is difficult; it does not require baseline data</td>
</tr>
<tr>
<td></td>
<td>• It is particularly useful for work focused on research, normative work and policy influencing, and capacity development</td>
</tr>
<tr>
<td></td>
<td>• The counting or identification of instances of congruence of evidence with a programme’s theory of change is easily facilitated through qualitative analysis software, such as NVivo, when set up appropriately during the inception phase</td>
</tr>
<tr>
<td></td>
<td>Limitations or special considerations:</td>
</tr>
<tr>
<td></td>
<td>• It requires a clear and confirmed (or revised) theory of change</td>
</tr>
<tr>
<td></td>
<td>• It is typically retrospective and prone to confirmation bias</td>
</tr>
<tr>
<td></td>
<td>• It is most useful when it involves extensive consultation with stakeholders and iteration with core team members at each step; it is therefore time consuming, and also requires sufficient access to policymakers</td>
</tr>
</tbody>
</table>

| **QUALITATIVE COMPARATIVE ANALYSIS** | Advantages:                                                                                                                                                                                                                  |
|                                      |  • It can be used to assess the difference made (effect) and the circumstances under which it was made, even in the absence of baseline data                                                                                     |
|                                      |  • It is useful for complex social, economic and political systems, especially where there are other known factors or interventions causing change                                                                              |
|                                      |  • It enables users to extrapolate or generalize findings beyond the evaluation sample                                                                                                                                        |
|                                      |  • It can be used to probe for unintended impacts, both positive and negative                                                                                                                                                 |
|                                      |  • It is useful for global programmes for which a field office could be a unit of analysis                                                                                                                                     |
|                                      | Limitations or special considerations:                                                                                                                                                                                           |
|                                      |  • It is an iterative process that can be quite time-consuming, requiring engagement and consensus-building among stakeholders on multiple outcomes of interest, despite knowledge of an outcome or its success or failure |
|                                      |  • It can be seen as being at odds with evaluations that have an accountability purpose, as the process is more focused on learning                                                                                            |
|                                      |  • Evaluators not familiar with the method may find it difficult to learn, particularly the methods and software used for analysis                                                                                         |
**PROCESS TRACING**

A qualitative analysis methodology used to test whether, and how, a potential cause or causes influenced a specified change or set of changes. The methodology works backwards on the results chain: it focuses on the use of clues (i.e., the results of a set of formal tests within a case) to examine the strength of evidence and adjudicate between alternative possible causal explanations contributing to development outcomes.

**Advantages:**
- It can be used to assess the process or mechanism by which an intervention led to or contributed to outcomes, even in the absence of baseline or monitoring data, and in consideration of diverse local contexts.
- It can be used to probe for unintended impacts, both positive and negative.
- It is particularly useful for work focused on policy influencing and capacity development, where assessment of attribution is difficult.

**Limitations or special considerations:**
- It is most useful when it involves extensive consultation with diverse stakeholders.
- It requires extensive evidence to be generated and/or collected.
- It is challenging and more time-consuming to apply if a theory of change is not available and the outcomes are not fully known.
- It holds the possibility of confirmation bias and/or misidentification of the significance of a process or causal explanation, especially for complex programmes and outcomes at higher levels of the causal chain.

**OUTCOME HARVESTING**

A non-experimental, participatory evaluation methodology with action learning at its core. Similar to process tracing, it works backwards in the causal chain and is used to identify, describe, verify and analyse development outcomes and what has contributed to them. It is used in complex situations where cause and effect relationships are not fully understood or are not defined using a theory of change. It starts with the design of a harvesting question to be responded to, followed by the drafting of outcome descriptions drawn from document reviews and the engagement of key informants, then ends with the selection of outcomes to be substantiated and analysed in the data collection phase. An example of a harvesting question related to the crime prevention work of UNODC would be: "How did the Government change its policies related to the prevention of violent extremism to reflect the concerns of communities affected by such extremism?"

**Advantages:**
- It can be used to assess which outcomes of the programme or intervention are most relevant across stakeholder types.
- It is useful for complex social, economic and political systems, especially for situations of uncertainty where activities were adaptive, fluctuating and/or relatively undefined.
- It is applicable to the assessment of work with Governments to integrate or implement legislation, policies and action plans.
- It can be used to probe for unintended impacts, both positive and negative.

**Limitations or special considerations:**
- It is prone to confirmation bias, and to confirming outcomes that are considered to be "low-hanging fruit" and overlooking outcomes that are more difficult to measure.
- It requires significant stakeholder participation.
- The focus is mostly on outcomes rather than activities or outputs, or the identification of processes and causal pathways.
- It has the potential for an unwieldy number of outcomes, depending on the scope of the evaluation, which might make the report cumbersome and/or not useful for the purpose of learning.
<table>
<thead>
<tr>
<th>METHOD</th>
<th>ADVANTAGES AND LIMITATIONS FOR USE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MOST SIGNIFICANT CHANGE</strong></td>
<td>A participatory evaluation methodology that involves the collection and selection of critical stories of change produced and analysed by a broad range of stakeholders. It emphasizes the centrality of culture and the “lived experience” of communities in the design and conduct of the evaluation.</td>
</tr>
</tbody>
</table>
| Advantages: | • It can be used to assess which outcomes of the programme or intervention are most relevant across stakeholder types  
• It is best for identifying difficult-to-measure, higher-level impacts, both intended and unintended, and both positive and negative  
• It can be used to evaluate human rights and democracy assistance interventions that lack clear outcomes and indicators  
• It is primarily a qualitative method, but can be adapted to include quantitative measurements |
| Limitations or special considerations: | • It is a simple approach and is easy to use in its purest form but, as with most participatory methods, can be time-consuming  
• It is typically applied as a method for monitoring change over time, but can be utilized in evaluations |
| **MAPPING/SOCIAL NETWORK ANALYSIS** | The mapping methodology, used in the design or implementation phase, helps to identify multiple stakeholders, their roles and relationships, and the drivers of and/or obstacles to change. Social network mapping is specifically designed to help map social networks and to identify and analyse the relationships within and between different actors. The results are presented with complex diagrams and maps of networks, capturing communication flows, business transactions, social media interactions or other relational interactions between individuals, groups or organizations |
| Advantages: | • It is useful for assessing normative change, policy influencing and mobilization, and to map and analyse knowledge networks and communities of practice  
• It is good for programmes that include activities and results (outputs and outcomes) related to partnerships, networks and cooperation |
| Limitations or special considerations: | • The findings are best presented visually, which requires some knowledge of software for developing diagrams (some examples include UCINET, integrating NetDraw, Gephi, and NodeXL).  
• It is highly participatory and also requires high response rates and/or complete data, therefore requiring sufficient time to engage stakeholders  
• It is most useful when the normative work has involved many stakeholders over a long period of time |
| **PARTICIPATIVE RANKING** | Participative ranking methodology is a mixed-method approach to data collection whereby facilitators guide participants in generating responses to specific questions, drawing upon qualitative and quantitative data derived from the responses. The methodology involves the contextualization, counting, ranking and comparison of responses across or within groups of participants. The approach is participatory in that it quickly generates findings while also permitting individuals to identify, support or refute the ideas, strategies, risks and changes that are most relevant, and then to come to consensus on key priorities as a group |
| Advantages: | • It can be used to assess which outcomes of the programme or intervention are most relevant across stakeholder types  
• It can be used to probe for unintended impacts, both positive and negative, as well as key good practices and lessons learned  
• It is useful for generating both qualitative and quantitative data on outcomes |
| Limitations or special considerations: | • It can be seen as being at odds with evaluations that have a purpose of accountability, as the process is more focused on learning  
• Although used in evaluation, it is more often applied in the conduct of needs assessments for programme design purposes  
• It is interactive and is typically conducted in the context of in-person focus group discussions, and therefore requires a more innovative approach when conducted remotely |
A cost-benefit analysis is a rigorous economic evaluation of all the costs and benefits associated with an intervention, including in terms of financial, environmental and social aspects, as well as productivity. A cost-effectiveness analysis determines how much is spent on an intervention in order to produce a particular outcome. The main difference between the two types is that in a cost-benefit analysis, benefits and costs are defined in comparable monetary terms, whereas in a cost-effectiveness analysis, benefits are identified but not compared with or defined by monetary values or costs.

**Advantages:**
- Either method is useful for answering questions related to efficiency, or for comparing the costs of two different interventions operating in similar conditions (useful for in-depth thematic or regional evaluations, for example).

**Limitations or special considerations:**
- There is a greater desire from donors for this type of analysis, but it can be difficult to use under typical evaluation constraints; it requires sufficient time and expertise in data analysis.
- Accurate financial data is needed.
- It is difficult to accurately define monetary values for factors or outcomes that cannot be easily quantified (e.g., reduced fear or trauma from crime and/or victimization).
- Evaluators must be careful when interpreting the results of the analysis, because the benefits of normative work are often difficult to quantify.

*The content in this table is largely drawn from BetterEvaluation.org, the United Nations Evaluation Group report “Compendium of evaluation methods reviewed – volume 1” and the United Nations Evaluation Group Handbook for Conducting Evaluations of Normative Work in the United Nations System, as well as the sources available at the links provided for each method listed in the table.

*The example question is drawn from the UNDP and Internal Alert publication entitled Improving the Impact of Preventing Violent Extremism Programming: A Toolkit for Design, Monitoring and Evaluation (Oslo, 2018)."

For any of these methodologies, it is important to consider that all team members, especially those engaged in data collection, will need to be adequately informed about the approach so that there is consistency in its application. This may be particularly challenging for evaluations that engage a geographically dispersed group of evaluators, as in-person consultation, training and pilot programmes may not be possible or effective. Given the increasing demand for external evaluators to apply more rigorous methodologies, there is an inherent trade-off between time and finances that must be considered in consultation with the Independent Evaluation Section and the project or programme commissioning the evaluation.
Evaluating global advocacy work initiatives requires special, non-traditional methods that are flexible and responsive to changing environments and shifting outcomes. Methods are often primarily participatory and/or qualitative in nature, although evaluators continue to innovate and test new methods in order to generate more credible findings. Among them are policy mapping, process tracing and network mapping.

Evaluating programmes that support democratization, the rule of law, anti-corruption efforts and human rights also requires flexible and participatory approaches that include a mix of quantitative and qualitative measures, such as contribution analysis or qualitative comparative analysis. Sentinel indicators may also be of use when assessing issues of governance, as the United Nations has developed, and has access to, potentially useful indices, such as the UNDP human development index and the Institute for Economics and Peace Global Terrorism Index. As with the other approaches listed, there are limitations to sentinel indicators. For example, although a country’s performance during the course of the United Nations support may have improved according to a given index, the improvement is unlikely to have been caused by the United Nations support alone.

Six steps to produce a credible contribution story:

1. Set out the attribution problem
2. Develop ToC and risks
3. Gather evidence
4. Assemble and assess
5. Seek out more evidence
6. Revise the story


Evaluating normative products and services such as publications, training courses, databases and knowledge networks also requires a combination of the special methods listed above, although peer reviews can also be helpful in assessing the technical quality of normative materials and longitudinal web analytics on the use of materials.

Innovative practice

UNODC encourages evaluators to introduce innovation into their evaluation process. This can be done by adding data collection methods that go beyond the standard fare of interviews, group discussions and surveys to include, for example, the use of approaches that are not based on the measurement of perceptions and by structuring different ways for people to provide input. Such practices support more robust triangulation of findings. Innovation can also be achieved by means of high levels of stakeholder engagement throughout the evaluation process, such as the co-development of data collection tools and mini-workshops to validate findings, or by using creative ways to share evaluation results. Demonstration of innovative practices is one element in the UNODC evaluation quality assessment grid.
DATA TOOLS AND SOURCES FOR DIFFERENT CONTEXTS

UNODC evaluations should use both primary and secondary data to answer evaluation questions, collecting data from existing sources, as well as generating new data that can be used to verify and extrapolate from the existing data. This is a form of triangulation, that is, the collection and verification of data drawn from multiple tools and sources. Most tools are able to generate both quantitative and qualitative data, although it is important to consider the availability and contexts of responding stakeholders in selecting and designing data collection tools. For example, in uncertain contexts, such as the COVID-19 pandemic, evaluators have more often opted for remote means of data collection, integrating more systematic reviews of secondary data and using online platforms to reach stakeholders and carry out interviews and surveys.

Further to this point, it is not always possible to reach specific stakeholders in person, and therefore remote means may be used to collect data. In the context of the COVID-19 pandemic, remote means of data collection were extensively tested, and good practices in that regard were compiled in a guidance note by the Independent Evaluation Section (see the link below). The key stakeholders in any UNODC evaluation process are, among others, project or programme managers, senior managers, government counterparts from different ministries, representatives of judicial bodies and law enforcement, direct beneficiaries, civil society organizations, donors and academic institutions.

UNODC Independent Evaluation Section, “Guidance note for managers and evaluators: planning and undertaking evaluations in UNODC during the COVID-19 pandemic and other crises”

Table 3 presents some common data collection tools that can be used to produce a credible and complete evaluation. Annex 10 provides further information on the advantages and limitations of using each tool.

---

### TABLE 3. COMMON DATA COLLECTION TOOLS

<table>
<thead>
<tr>
<th><strong>DESKTOP REVIEW</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Desk review includes scanning, and, in some cases, more rigorously coding, documentation of internally- and externally-sourced background information. Such information may include monitoring data and annual reports (e.g., annual programme or project progress reports or donor reports for UNODC), as well as research products of UNODC and other United Nations entities, and national or regional strategies. Typically, programme or project managers collect extensive evidence on how activities were implemented (the process), what was achieved (the results) and whether the approach taken was systematic or not. Desk review should be integrated into any evaluation in order to generate evidence across evaluation questions. Good practice is to integrate across phases, thereby helping to tailor the selection of evaluation questions and methodologies during the inception phase, and then to respond to subsidiary questions or assumptions and enrich findings from primary sources in the data collection and analysis phase.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>REVIEW AND MINING OF SECONDARY DATA</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>The review of secondary data can include the more systematic mining of secondary source documents, such as national household surveys and crime statistics databases, or literature or evaluations regarding similar types of support for quantitative and/or qualitative data that have already been collected. Existing socioeconomic datasets can be used as proxies for other measures, such as outcomes or impact indicators, and can also provide a means to triangulate information regarding the relevance of a project or programme, based on changing country contexts.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>INTERVIEWS WITH KEY INFORMANTS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluations typically involve qualitative interviews with key informants, that is, people who know details about the activity, such as staff members, government officials, representatives of United Nations agencies, donors and community leaders. The interviews can be conducted one-on-one or in groups and usually include structured or semi-structured open-ended questions. When needed and feasible, the interviews can also be conducted by means of live telephone calls, using computer-assisted telephone interviewing techniques.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>FOCUS GROUP DISCUSSIONS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unlike group interviews, in which participants are asked to respond to the same set of questions as posed in individual interviews, focus group discussions are exploratory discussions with a group of stakeholders, typically structured to allow for open, in-depth answers. Focus groups provide the opportunity to include diverse groups and perspectives, especially under time constraints. Typically, a focus group discussion involves homogenous groups of 6 to 10 people, for example, representatives from a specific field office or participants in training events. Cultural considerations should also be taken into account; in order to better facilitate open discussion, for example, it may be more effective if men and women are interviewed in separate groups.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>SURVEYS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveys provide medium to high volumes of quantitative information through structured questions. Questions are typically closed-ended with a predefined answer, such as “yes”, “no” or “don’t know”, or “complete”, “partial”, “none” or “don’t know”. Open-ended questions can sometimes be used, although it is suggested that they be targeted and sparing in order to increase survey response rates. Surveys are best undertaken using available software, such as SurveyMonkey, SurveyCTO or KoboToolbox, to facilitate automated reminders, analysis functions and reporting. When necessary and feasible, for example, owing to time constraints or the remoteness of respondents, surveys can be carried out using, among other options, text messaging applications on mobile phones or interactive voice response technology to survey hard-to-reach populations. Surveys are usually tailored to one or all of the evaluation criteria, either gathering broad feedback on perceptions, experiences or expectations, or narrowing the focus to a specific criterion, such as effectiveness, to explicitly measure progress towards achievement of programme outcomes. Response rates in surveys typically plunge after 10 minutes of participation, so it is best to keep surveys short when they are self-administered and/or conducted remotely.</td>
</tr>
</tbody>
</table>
OBSERVATIONS

A coded assessment checklist or grading matrix may be used to record observable events or behaviour such as attendance at, participation in and outcomes of training events. Observation functions as an open-ended recording and/or visual intake of an experience, typically focused on participants or beneficiaries of a programme activity. Innovation in this area has included the use of remote sensing and satellite imagery or other geospatial technology.

SELF-REPORTED CHECKLISTS AND SCORECARDS

Checklists and scorecards can be used to gather participant perspectives on the basis of a preselected list aligned with programme or project indicators. They can be used to collect quantitative information (e.g., frequency or satisfaction) or qualitative information (e.g., descriptions of the experience).

CASE STUDIES

Case studies are empirical inquiries and collections of narratives from individuals, such as recipients of assistance, about their experiences. They examine, usually in great depth, the theme or context (e.g., geographical area) from a range of different perspectives (qualitative data). Choosing four to five case units and analysing them within and across units can generate qualitative and quantitative information about the conditions needed for interventions to work. An example in the context of a programme aimed at preventing and countering violent extremism would be the conducting of a set of case studies that examine a small sample of programme participants to better understand whether the programme facilitated the departure of persons from radicalized groups in different regions.

BENCHMARKING

Benchmarking is a method used to compare the performance of one individual, project or programme, organization or institution, or policy with that of an appropriately similar one, such as a similar organization, best practice or industry standard.

GOOD PRACTICE

A high-quality evaluation is one that ensures rigour in the evaluation process and methodology, with a view to producing reliable findings, typically by incorporating multiple and mixed methods into the evaluation design rather than relying on one method alone. Selecting the appropriate types and number of tools and ensuring that they can accommodate a sufficiently large and varied group of stakeholders helps evaluators to mitigate potential limitations and risks, such as having unresponsive or inaccessible stakeholders, by enhancing the collection of strong evidence and counteracting the possibility of inherent bias in a particular method or source. For example, an evaluator might triangulate three types of perception-related data (e.g., data derived from desk reviews, interviews and surveys), and further strengthen the assessment of the intervention’s contribution through the use of statistical evidence from a systematic data mining or document analysis process (e.g., gathering evidence on the number of women reporting on violent crimes longitudinally, or on the cost of new psychoactive substances at the street level over the course of a programme).
EXAMPLE TOOLS

CRIME PREVENTION

QUALIPREV: A TOOL FOR SCORING THE QUALITY OF CRIME PREVENTION PROGRAMMES

**Description.** Developed into a Microsoft Excel sheet with comprehensive scoring manual, the QUALIPREV tool supports rigorous, comparable and translatable evaluation findings. It is grounded in a scientific systematic review of the literature on crime prevention effectiveness and evaluation.

**Why and when to use.** When seeking a robust and empirically grounded assessment of the quality of the crime prevention programme being evaluated. The tool directs evaluators to attend to programme characteristics that research has shown are associated with effectiveness.

**Tips for increasing usefulness.** The tool may be used both for programmes that have received some form of earlier (e.g., mid-term) evaluation, for which an extended suite of scoring options is available, or for programmes that have not been previously evaluated.

JUDICIAL REFORM

MEASURING PERCEPTIONS OF FAIRNESS: AN EVALUATION TOOLKIT, FOR USE IN COURT PROCESSES

**Description.** An evaluation toolkit designed to support evaluations of court reforms, providing a measurement tool to elicit perceptions of the four key dimensions of procedural fairness: voice, respect, neutrality and understanding, as well as a fifth dimension, helpfulness.

**Why and when to use.** When seeking data on the final beneficiaries of programming aimed at enhancing judicial capacity to function independently, impartially and with integrity, and to uphold human rights and norms in criminal cases.

**Tips for increasing usefulness.** The toolkit provides a tool to measure user perceptions that may be adapted to many situations where perceptions of procedural fairness and justice are concerned, including special measures to enhance gender equality and satisfaction with justice.

DATA CODING AND ANALYSIS

The process for coding and analysing each type of data collected also needs to be set out in the inception report, implemented and then described in the methodology section of the evaluation report. The type of analysis undertaken will depend on the method followed to obtain the data, as well as the amount of data to be dealt with. Although not a requirement, analysis software is increasingly used for both qualitative and quantitative data, and a selection of the more common packages is listed in the table below. Those working with smaller data sets may still choose more basic processes such as manual coding and descriptive analysis using Microsoft Excel or other spreadsheet applications. However, analysis software offers the benefits of efficiency and accuracy (if used properly), especially when dealing with large amounts of data, and in providing a range of options for the visual depiction of results.

<table>
<thead>
<tr>
<th>QUALITATIVE</th>
<th>QUANTITATIVE</th>
<th>VISUAL/GEOGRAPHIC INFORMATION SYSTEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVivo</td>
<td>Microsoft Excel</td>
<td>Esri ArcGIS</td>
</tr>
<tr>
<td>Atlas.ti</td>
<td>Stata</td>
<td>Google Earth Pro</td>
</tr>
<tr>
<td>MAXQDA</td>
<td>SPSS</td>
<td>BatchGeo</td>
</tr>
</tbody>
</table>
ENSURING HUMAN RIGHTS AND GENDER EQUALITY RESPONSIVENESS, AND DISABILITY INCLUSION

Equality and inclusion are critical to programme or project design and evaluation ethics, and gender and human rights have been standard cross-cutting criteria for assessment in other United Nations entities that have a more direct service delivery role, such as UNICEF, the United Nations Population Fund and the World Food Programme. However, UNODC, by placing more focus on normative work, continues to evolve and adapt higher standards for equality and inclusion; for example, in 2019, UNODC still remained below the average rating of the second generation of the United Nations System-wide Action Plan for Gender Equality and the Empowerment of Women (UN-SWAP 2.0) for the United Nations Secretariat by 6 per cent, and below the average for all other entities in the United Nations system by 3 per cent. UNODC strategies are currently focused more on field-level work, and areas of intervention, such as terrorism prevention, are being further defined and required to take a “whole-of-society” approach and/or to integrate the “leave no one behind” principle of the United Nations. As a result, gender equality, human rights and considerations relating to the principle of “leave no one behind” are likely to take on a more prominent role in evaluation design, methodologies, analysis and reporting and dissemination in the coming years.

HUMAN RIGHTS AND GENDER EQUALITY

All United Nations evaluations are required to recognize and mainstream human rights and gender equality as part of the United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women. This means that:

- Evaluations must consider the extent to which human rights and gender equality principles were meaningfully incorporated into the intervention’s design and implementation and include appropriate gender-related recommendations.
- The evaluation process itself must be conducted in a manner that is human rights- and gender equality-responsive. For example, rights holders should have the opportunity to fully participate in the evaluation process, accommodations should be made for their full participation and data should be disaggregated by gender and by relevant disadvantaged and advantaged group.

---

Useful resources for achieving these aims include the following:

- United Nations Office at Vienna (UNOV)/UNODC Strategy for Gender Equality and the Empowerment of Women, 2018–2021
- UNODC and human rights
- Guidance for mainstreaming gender in UNODC evaluations
- UNODC and gender
- Convention on the Elimination of All Forms of Discrimination against Women (1979)
- United Kingdom Gender Action for Peace and Security (GAPS) network and others, “Beyond consultations: a tool for meaningfully engaging with women in fragile and conflict-affected States” (2020)
- United States Department of State, “INL guide to gender in the criminal justice system” (2016)

**INNOVATIVE STRATEGIES MAY BE NEEDED WHEN GENDER-RELATED DATA ARE NOT CLEARLY APPARENT.**

For example, in the 2017 mid-term evaluation of the UNODC Sahel Programme 2013–2023 (XAMZ17), which is aimed at strengthening criminal justice and law enforcement capacity, the evaluators noted that there had been limited availability of information and baseline data on human rights and gender equality to form the basis for conducting an analysis and developing recommendations for future programming, as had been requested in the terms of reference. To mitigate this situation, the evaluators devised several proxy indicators, including by examining how capacity-building activities had been conducted (which involved reviewing letters of invitation and conducting mini case studies), how human rights and gender equality had been considered in the development of national strategies and international conventions, and the proportion of women within law enforcement and the judiciary for each country of the Sahel region.

**DISABILITY INCLUSION**

In alignment with the United Nations Disability Inclusion Strategy, evaluations within the United Nations system are now to consider disability-related issues as a specific cross-cutting theme and, like other entities, UNODC is required to report on its progress towards achieving the Strategy on an annual basis. There are several ways in which the evaluation process can be disability-inclusive, the most common being to have questions that shed light on the extent and quality of disability inclusion (see annex 8 for indicative questions) and to present the evidence in the report findings. Other strategies are to involve persons with disabilities or organizations of persons with disabilities in stakeholder mapping and data collection processes, as well as within evaluation teams, to collect and present data that are disaggregated by disability where feasible, and to reflect data that have been collected on disability inclusion in the conclusions and recommendations of the report.
Disability inclusion refers to “the meaningful participation of persons with disabilities in all their diversity, the promotion and mainstreaming of their rights into the work of the Organization, the development of disability-specific programmes and the consideration of disability-related perspectives, in compliance with the Convention on the Rights of Persons with Disabilities”.

United Nations Disability Inclusion Strategy (www.un.org/en/content/disabilitystrategy/)

**RESOURCES**

- Convention on the Rights of Persons with Disabilities
- United Nations Disability Inclusion Strategy
CHAPTER 5.

Additional considerations for evaluation in complex environments

This chapter looks further into how evaluations can be effectively carried out in complex environments. It covers fragile contexts and conflict sensitivity analysis, dealing with limited availability of data and hard-to-reach populations, and addressing language challenges and stakeholders with limited knowledge of the intervention being evaluated. It also highlights the importance of preparing a solid risk management plan when the evaluation is being designed.
The present chapter goes into further detail on the features of the locations, environments and sometimes rapidly changing conditions where UNODC work is carried out, and consequent considerations for evaluation design. These include considerations relating to the security of the evaluators and the scope and viability of field missions, as well as solutions for overcoming difficulties in the countries concerned, including with regard to the capacity to enter or move around within the country for purposes of data collection and the evaluator’s capacity to work in both official and vernacular languages, and how to ensure data quality when using remote means of communication.

**FRAGILE CONTEXTS AND CONFLICT SENSITIVITY ANALYSIS**

Many UNODC programmes are delivered in fragile or conflict-affected States or regions. Internationally agreed norms encompassed in the Principles for Good International Engagement in Fragile States and Situations (2007) of the OECD Development Assistance Committee and the New Deal for Engagement in Fragile States (2011) help international actors, such as UNODC, apply values such as “leave no one behind” and “do no harm”. Fragility and conflict present particular difficulties for the design, delivery and evaluation of programmes. A key concept in these processes is conflict sensitivity. This refers, firstly, to the way conflict or fragility factors have been mapped and understood in the design phase and how the programme has been suitably adapted to positively manage them, which is typically assessed in evaluations under the "relevance" criterion and in the assessment of unintended impacts, whether positive or negative. Secondly, evaluators need to understand how these same conflict and fragility dynamics might affect their own planning, such as by placing limitations on fieldwork or the types of data or indicators it might be possible to collect or obtain access to. Specifically in fragile contexts, there may be missing, unreliable or poor quality data; insecurity, making it difficult to access regions or specific populations (owing to either political reasons or physical remoteness); lack of trust on the part of evaluation stakeholders or participants; and fast-changing and unpredictable environments that are not fully understood by evaluation practitioners.

In such cases, it is important to recognize how both the programme and the evaluation (and evaluator) may interact with the fragile or conflict-affected environment, and to recall the “do no harm” principle that should guide all activity. The publication entitled “Doing research in fragile contexts: literature review” (provided as a resource below), notes that participants could perceive themselves at a disadvantage in the evaluation process because evaluators generally hold the power to interpret the data collected; those interpretations could exert considerable influence over policies and programmes that directly affect participants’ lives or touch on politically sensitive issues that are perceived as private, not public, information. In circumstances such as these, drawing on the skills and knowledge of national evaluators and experts is particularly useful. They are frequently more able to move around within a country, identify and navigate risks of different sorts, from security to logistical risks, and understand cultural norms. Also useful are efforts to evaluate and consider innovative, often remote, data collection techniques. Annex 9 provides a checklist of elements to consider in relation to conflict-sensitive evaluations at each stage of the evaluation process.
RESOURCES

- "Human rights due diligence policy on United Nations support to non-United Nations security forces: guidance note and text of the policy"

- United Kingdom Stabilisation Unit, "Conflict sensitivity tools and guidance" (June 2016) (see tool 5, "Checklist for conflict sensitivity within monitoring and evaluation plans")


- Iffat Idris, "Doing research in fragile contexts: literature review", GSDRC Research Report No. 1444 (Birmingham, GSDRC, University of Birmingham, 2019)

- Annex 9: Checklist for conflict-sensitive evaluations and measurement on sensitive subjects

CHATHAM HOUSE RULE

There may be situations where it is helpful for respondents to be anonymous. In fragile and conflict-affected environments, it is permissible for off-the-record conversations to be used in an evaluation for safety and security reasons. The Chatham House Rule is used to encourage inclusive and open dialogue. It enables the use of information received but requires that the identity and affiliation of the speaker, or other participants in the session, not be revealed.


LIMITED AVAILABILITY OF DATA AND HARD-TO-REACH POPULATIONS

Identifying the contribution of programme activities to change at the outcome level or, in particular, the longer-term outcome or impact level, requires thoughtful selection of measures and modalities of data collection. This can be a great deal more complex when the environment itself is complex, as in the case of fragile or conflict-affected settings. The evaluation may face a variety of restrictions based on security and logistical factors and, sometimes, simply financial factors. Furthermore, many of the populations targeted by UNODC work may be difficult to reach for additional reasons. They might, as in the case of intravenous drug users or undocumented migrants, for example, be wary of formal-looking investigators, have limited trust and be concerned about the risks that providing information might expose them to. Other groups, including final beneficiaries or rights holders, might be difficult to reach owing to their remoteness, such as poppy or coca farmers, or because the Member State restricts access to them either physically, such as in border regions, or completely, as might occur in the case of convicted criminal offenders or even victims or witnesses. Yet obtaining good quality data is more than possible by combining direct and indirect measures, and giving voice to difficult-to-reach populations is critical to the universal goals of the United Nations to leave no one behind and to reach the furthest behind first.
Evaluations that include interactions with survivors of gender-based violence need to pay particularly close attention to ethical and safety issues. Although it is often desirable to receive input from the range of rights holders served by the intervention, the benefits for respondents of collecting data on gender-based violence need to be greater than the risks posed to them. The publication by InterAction entitled *Gender-based Violence Prevention: A Results-based Evaluation Framework*, listed in the resources below, offers useful guidance on respectful data collection approaches and tools. In addition, annex 1 outlines steps to take in conducting safe and ethical research on gender-based violence in conflict and humanitarian settings.

“As we embark on this great collective journey, we pledge that no one will be left behind. Recognizing that the dignity of the human person is fundamental, we wish to see the Goals and targets met for all nations and peoples and for all segments of society. And we will endeavour to reach the furthest behind first.”

*Transforming our World: the 2030 Agenda for Sustainable Development*

### RESOURCES

- UNDP, “What does it mean to leave no one behind?: a UNDP discussion paper and framework for implementation” (July 2018)

### ENGAGEMENT OF YOUTH

Through its work, UNODC is increasingly reaching young people, directly or tangentially. Successfully engaging this population in evaluation processes needs to be well thought out. It often requires evaluators to unpack commonly held assumptions and recognize, for example, that adolescents and young people are not a homogenous group, so methods may need to be customized accordingly, that most young people will never participate in a youth programme, so issues of exclusion and alienation need to be taken into account, and that many young people feel frustrated by Governments and other institutions not having the capacity or determination to provide youth services, in spite of the rhetoric about doing so. Strategies to increase engagement include involving young people at each stage of the evaluation – from the design and protocol development stage to validating results and formulating recommendations – and using more novel methods of data collection, such as photovoice and other forms of storytelling.

---

**Photovoice** is a method for storytelling, typically used to capture the perspectives of youth and other marginalized or disempowered groups, that can add a dynamic element to evaluations, especially when desired groups are unlikely or unable to respond to surveys, or when only limited translation services are available. Participants are typically asked to represent their point of view, either openly or in response to specific questions, by taking photographs, discussing them together, and developing narratives to go with their photos. The photos and narratives are typically coupled with a subsequent phase of outreach or other action. This is an innovative method for gathering data on community-level outcomes and impacts, whether intended or unintended, of policies, laws and development plans based on international norms.

**RESOURCES**

- UNODC youth-related resources
- "The Convention on the Rights of the Child and the international legal framework on children’s rights"
- Act for Youth, "Resources for youth participatory evaluation"

**LANGUAGE CHALLENGES**

In many cases, the evaluation will be conducted, and data must be collected, in contexts in which the language skills of the evaluator are not adequate. It is not uncommon that more senior government representatives of the Member State will speak English, but even then, there may be reasons of policy or concerns about precision of expression that lead them to prefer to speak in a language other than English. Furthermore, secondary data sources, such as government statistical collections or civil society reports, will often be in the national language. Stakeholders with whom the evaluators will need to engage may even speak a less standard vernacular. In Pakistan, for example, where the national language is Urdu, final beneficiaries might speak only one of the dozen or more regional or indigenous languages of the country.

Carefully considering questions of language, in consultation with the Independent Evaluation Section and the UNODC country office, will help to identify such challenges. One effective way around them involves recruitment of one or more national evaluators. Owing to their intimate knowledge of local political and security contexts, national evaluators bring substantial benefits to an evaluation by rendering language assistance, whether in relation to in-person fieldwork or remote data collection, such as in offering translation support during a voice or video interview via Microsoft Teams. As another example, questionnaires and surveys may need to be translated into one or more local languages and then translated back to the original language to confirm their accuracy using translation software, evaluations teams or by relying on internal capacity within the commissioning department at UNODC. Results will also need to be translated back into English. Furthermore, and importantly, engaging national evaluators helps to build national capacity and evaluation expertise, consistent with General Assembly resolution 69/237, on building capacity for the evaluation of development activities at the country level, and target 17.9 of the Sustainable Development Goals.
STAKEHOLDERS’ LIMITED UNDERSTANDING OF THE INTERVENTION

A frequent challenge faced by UNODC evaluations is that key stakeholders, such as government officials and other partners, often have only a limited understanding of the entirety of the programme being evaluated. Sometimes this is due to the programme having multiple components that are implemented in different locations, or even in different countries, as in the case of regional programmes. Frequent staff turnover can also be a reason why key contacts do not know the background, scope or intended outcomes of the intervention. This can make it difficult to obtain important data on some of the broader evaluation questions, particularly those related to relevance, efficiency and sustainability.

One suggestion for addressing these issues is to prepare a briefing note on the programme that can be shared with evaluation participants prior to or at the beginning of interviews or group discussions. Such a note could also be drawn upon for the introduction to a survey. It is important to keep in mind the need to tailor the description to the specific audience and to avoid including information that could bias the recipients’ responses. It can also be helpful to share the interview questions in advance with stakeholders who are new to their positions. This can enable them to consult with others who may have had a deeper connection to the programme.

RISK MANAGEMENT IN RELATION TO SECURITY AND THE PROTECTION OF HUMAN RIGHTS

It is good practice for all evaluation processes to include a risk management plan. Such a plan should be used to identify potential risks and develop mitigation strategies to enable completion of the evaluation as agreed. This is even more important when dealing with sensitive subjects or in conflict-affected areas where the safety, security and well-being of all evaluation participants, data collectors and evaluators need to be given special consideration.

Risk management involves proactively anticipating and understanding risks that already exist or that may emerge during the course of the evaluation, and developing strategies for avoiding or minimizing them. It is likely that many of these strategies will reflect the additional considerations discussed above. For example, conflict sensitivity analysis may identify the risk that the evaluation process itself could impact local social dynamics and safety. Take the example of a remote drug cultivation region where UNODC has engaged with small farmers on crop substitution under an alternative livelihoods programme. To what extent might the
presence of international evaluators interviewing final beneficiaries of the programme affect local social
dynamics (such as through perceptions of favouritism) or security (e.g., if such contact attracts attention and
exposes farmers to potential violence from drug gangs)? In such cases, risk mitigation measures might involve
working with village leaders to identify suitable candidates for participation in the evaluation and utilizing
national evaluators rather than international team members to conduct fieldwork in such a remote and
insecure environment. Of course, risks to evaluators themselves must also be considered, therefore the risk
management plan will need to logically calculate all possible risks and mitigation options. In some cases, such
as those involving security risks, it might be decided that no feasible mitigation measures exist and that some
other approach is necessary, for example, the use of remote sensing data to estimate changes in crop
cultivation patterns, rather than interviews with smallholder farmers themselves.

As this example illustrates, undertaking a risk analysis and the process of devising a risk management plan are
important not only because they help to ensure the smooth running of the evaluation, but also because they
flag to the evaluator the key ethical risks that may be posed by the conduct of the evaluation. United Nations
Evaluation Group guidelines on the ethical conduct of evaluations are explicit 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 connection between proper ethical conduct and risk management should therefore be clear. It is
important for the evaluator to consider at each stage of the evaluation process how that which is being
planned might invite ethical or other risks. Therefore, another way of thinking of these risks is that potentially
negative outcomes for participants or evaluators themselves, ranging from violence to breach of privacy
or personal integrity, are also threats to their human rights. A human rights-compliant evaluation is therefore one
in which the risk management planning has given explicit consideration to and made explicit accommodation
for situations where human rights issues might be involved. Informed consent for participation is of course
required as a key ethical principle in every evaluation.

Risks relating to safety, security, ethical conduct and human rights might be described as the hard end of
potential risks. At the other end of the risk continuum are a variety of risks that evaluators should try to
identify that affect the timely completion of the evaluation. Factors such as national elections, national
holidays, religious festivals, rainy or hot seasons and the like may all affect the potential availability of groups
with whom the evaluator wishes to engage. They may also affect the feasibility of specific evaluation activities,
for example, the undertaking of field work that includes travel within a country, which may be affected by
factors such as road closures during the rainy season.
RESOURCES


Overseas Development Institute, “Tips for collecting primary data in a COVID-19 era”
CHAPTER 6.

Presentation of results

This final chapter focuses on the evaluation report. It guides the reader towards establishing the links between the key findings, conclusions, recommendations and lessons learned. It also contains suggestions for clearly communicating the main results of the work so that the evaluation can be as useful to the intended audience as possible.
More than anything else, the main output of the evaluation process, the evaluation report, needs to be useful for its intended audience. To meet this aim, the report must be viewed as a credible product and, unlike a research paper, it should be easily accessible to a range of readers, many of whom will not have time to read the report in full.

Although report writing can be a daunting process, the UNODC report template, evaluation handbook and evaluation quality assessment grid all provide useful guidance for this task.

The credibility of the report is enhanced by having a well-described and adequately robust methodology, a logical flow from the evidence through to the recommendations (as shown in figure V), and a well-presented document. A key element of usability is the way in which the recommendations are developed and structured. Good practice is for recommendations to be co-developed with key stakeholders, including rights holders, and this is often done as part of a data validation workshop. Where this is not feasible, stakeholders should at least have the opportunity to comment on the draft recommendations to help ensure they are well targeted and feasible to implement. Another consideration is how the recommendations are drafted. UNODC prefers to have a limited number of overall recommendations (there can be subsets within each) and for them to be “SMART”, that is, specific and clear, measurable, achievable, relevant – which can be ensured by co-developing them with stakeholders – and time-bound, and ranked according to priority. Recommendations in UNODC evaluations should exclusively be targeted at UNODC, and not Member States. Recommendations can be directed at other United Nations entities or other organizations in cases where they are taking part in a joint evaluation process.

Suggestions for increasing the usefulness and accessibility of the report include the following:

- Include a concise and well-written executive summary
- Use well-formatted visual aids that clearly communicate the evaluation process and the key data and results
- Include illustrative quotes from evaluation participants, but only if consent was obtained and there are no identifiers that might put the source at risk
- Illustrate the linkages between the main conclusions and recommendations by numbering each and referencing the relevant conclusion number or numbers for each recommendation
Annexes
ANNEX 1. AREAS OF WORK OF THE UNITED NATIONS OFFICE ON DRUGS AND CRIME: KEY OUTCOMES AND OUTPUTS

ANNEXES 48

UNODC FIELD OFFICES

UNODC operates in more than 80 countries around the world through its network of 115 field offices, with approximately 2,400 UNODC personnel globally.
ANNEX 3. HUMAN RIGHTS-RELATED INSTRUMENTS AND PROVISIONS RELEVANT TO THE WORK OF THE UNITED NATIONS OFFICE ON DRUGS AND CRIME

The references listed below include both binding and non-binding standards relating to international human rights. The references are non-exhaustive and are provided as examples only. The list is not intended as a definitive statement as to the extent of international human rights law.

NON-DISCRIMINATION

- Prohibition on discrimination – International Covenant on Civil and Political Rights, article 2
- Right to freedom of thought, conscience and religion – International Covenant on Civil and Political Rights, article 18
- Right to exercise economic, social and cultural rights without discrimination of any kind – International Covenant on Economic, Social and Cultural Rights, article 2
- Direct discrimination – Human Rights Committee General Comment No. 18, paragraph 7
- Indirect discrimination – Committee on the Elimination of Racial Discrimination General Recommendation No. 14, paragraph 2
- Legitimate and proportional distinctions – Human Rights Committee General Comment No. 18, paragraph 13
- Protection from mandatory HIV testing – International Guidelines on HIV/AIDS and Human Rights, guidelines 3, paragraph 20 (b), and 5, paragraph 22 (j)
- Separation of prisoner categories – Standard Minimum Rules for the Treatment of Prisoners, paragraph 8
- International Convention on the Elimination of All Forms of Racial Discrimination (1965)
- International Covenant on Economic, Social and Cultural Rights (1966)
- International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families (1990)

RIGHTS OF WOMEN

- Development and advancement of women – Convention on the Elimination of All forms of Discrimination against Women, article 3
- Obligation to take all appropriate measures to suppress all forms of traffic in women and exploitation of prostitution of women – Convention on the Elimination of All forms of Discrimination against Women, article 6
- Honour defences – CEDAW/C/JOR/CO/4, paragraph 24
- Elimination of prejudices – Convention on the Elimination of All forms of Discrimination against Women, article 5
- Responses to violence against women – Economic and Social Council resolution 2010/15
- Women prisoners – United Nations Rules for the Treatment of Women Prisoners and Non-custodial Measures for Women Offenders (the Bangkok Rules)
RIGHTS OF THE CHILD

- Obligation for the best interests of the child to be a primary consideration – Convention on the Rights of the Child, article 3
- Non-discrimination with respect to child rights – Convention on the Rights of the Child, article 2, paragraph 1
- Best interests of the child – Convention on the Rights of the Child, article 3, paragraph 1
- General principles for juvenile justice – Convention on the Rights of the Child, article 40, paragraph 1
- Aims of juvenile justice – United Nations Standard Minimum Rules for the Administration of Juvenile Justice (Beijing Rules), article 5.1
- Detention to be a disposition of last resort – Convention on the Rights of the Child, article 37 (b)
- Protection of children from illicit drugs – Convention on the Rights of the Child, article 33
- Right of the child accused of having infringed the penal law to be treated in a manner consistent with the promotion of the child’s sense of dignity and worth – Convention on the Rights of the Child, article 40

LAW ENFORCEMENT

- Right to liberty, security of person, and non-arbitrary arrest or detention – International Covenant on Civil and Political Rights, article 9
- Right to be informed of reasons for arrest – International Covenant on Civil and Political Rights, article 9, paragraph 2
- Right to be brought promptly before a court – International Covenant on Civil and Political Rights, article 9, paragraph 3
- Right to life – International Covenant on Civil and Political Rights, article 6
- Right not to be subjected to enforced disappearance – International Convention for the Protection of All Persons from Enforced Disappearance, article 1
- Right not to be subjected to torture or to cruel, inhuman or degrading treatment or punishment – International Covenant on Civil and Political Rights, article 7
- Right not to be subjected arbitrary or unlawful interference with privacy, family or home – International Covenant on Civil and Political Rights, article 17
- Right to freedom of association – International Covenant on Civil and Political Rights, article 22
- Right to liberty of movement – International Covenant on Civil and Political Rights, article 12
- Obligation to take appropriate measures to investigate acts of enforced disappearance – International Congress on Population Education and Development, article 3
- Powers of seizure and confiscation – European Court of Human Rights appeal No. 19955/05; E/CN.4/1995/19/Add.1
- Evidence-based searches and arrests – Human Rights Committee Communication No. 1493/2006
  — Intelligence gathering and the right to privacy – Report of the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism, Martin Scheinin (A/HRC/13/37)
- General principles – Code of Conduct for Law Enforcement Officials, General Assembly resolution 34/169, annex, article 2
• Prohibition on torture – Universal Declaration of Human Rights, article 5; Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, article 2; Code of Conduct for Law Enforcement Officials, article 5
• Use of force – Basic Principles on the Use of Force and Firearms by Law Enforcement Officials
• International Covenant on Civil and Political Rights (1966)

PROSECUTION AND COURT

• Right to a fair trial – International Covenant on Civil and Political Rights, article 14
• Right to be presumed innocent – International Covenant on Civil and Political Rights, article 14, paragraph 2
• Prohibition on retroactive criminal offences – International Covenant on Civil and Political Rights, article 15
• Adequate time and facilities for defence – International Covenant on Civil and Political Rights, article 14, paragraph 3 (b)
• Prohibition on use of evidence obtained by torture – Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, article 15
• Timeliness of criminal proceedings – International Covenant on Civil and Political Rights, article 9, paragraph 3; Body of Principles for the Protection of All Persons under Any Form of Detention or Imprisonment, article 38
• Protection of counsel – Human Rights Committee General Comment 13, paragraph 9
• Non-prosecution of trafficked persons – Recommended Principles and Guidelines on Human Rights and Human Trafficking, guideline 2
• Equality of arms – European Court of Human Rights, Ofer, No. 524/59 and Hopfinger, No. 617/59
• Impartiality of judiciary – Basic Principles on Independence of the Judiciary, article 2

SENTENCING AND PRISONS

• Treatment of prisoners – International Covenant on Civil and Political Rights, article 10
• Prohibition of the death penalty – International Covenant on Civil and Political Rights, paragraph 2
• Restrictions on application of the death penalty – International Covenant on Civil and Political Rights, article 6, paragraph 2); Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Philip Alston (A/HRC/4/20), paragraph 53; Economic and Social Council resolution 1984/50
• Prohibition on imprisonment related to contractual obligations – International Covenant on Civil and Political Rights, article 11
• Provision of adequate facilities for prisoners – Standard Minimum Rules for the Treatment of Prisoners, paragraphs 9–26
• Communication with family and visits – Standard Minimum Rules for the Treatment of Prisoners, paragraphs 37–39
• Proportionality of penalties – A/CONF.144/28/Rev.1, at p. 164
• Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (1984)
DRUG PREVENTION, TREATMENT AND CARE

- Right to the highest attainable standard of physical and mental health – International Covenant on Economic, Social and Cultural Rights, article 12
- Obligation to protect children from illicit use of narcotic drugs and psychotropic substances – Convention on the Rights of the Child, article 33
- Right not to be subjected to cruel, inhuman or degrading treatment or punishment – International Covenant on Civil and Political Rights, article 7
- Right not to be subjected without free consent to medical or scientific experimentation – International Covenant on Civil and Political Rights, article 7
- Right not to be subjected to arbitrary or unlawful interference with privacy, family or home – International Covenant on Civil and Political Rights, article 17
- Right to an adequate standard of living - International Covenant on Economic, Social and Cultural Rights, article 11
- Right to appropriate serves in connection with pregnancy – Convention on the Elimination of All forms of Discrimination against Women, article 12
- Right to non-discrimination – International Covenant on Civil and Political Rights, article 2
- Drug dependence as a health disorder – Report of the Special Rapporteur on torture and other cruel, inhuman or degrading treatment or punishment, Manfred Nowak (A/HRC/10/44), paragraph 71
- Evidence based treatment for drug dependence – Report of the Special Rapporteur on the right of everyone to the enjoyment of the highest attainable standard of physical and mental health (A/65/255), paragraphs 50–55; A/64/92-E/2009/98, Political Declaration, paragraph 21
- Proportionality of responses to drug laws – E/INCB/2007/1, chapter 1, paragraphs 13 and 60
- Criminal law should not impede access to drug treatment – Report of the Special Rapporteur on the right of everyone to the enjoyment of the highest attainable standard of physical and mental health (A/65/255), paragraph 62
- Detained drug users and access to services – A/HRC/10/44, paragraph 59; E/C.12/UKR/CO/5, paragraph 51; A/64/92-E/2009/98, Plan of Action, paragraph 16 (c)
- Drug treatment should be voluntary – International Covenant on Economic, Social and Cultural Rights, article 12; Committee on Economic, Social and Cultural Rights General Comment No. 14, paragraph 34; Report of the Special Rapporteur on the right of everyone to the enjoyment of the highest attainable standard of physical and mental health (A/64/272), paragraphs 88–91

ALTERNATIVE DEVELOPMENT AND SUSTAINABLE LIVELIHOODS

- Adequate standard of living – International Covenant on Economic, Social and Cultural Rights, article 11
- Right to Development – A/RES/41/128; African Charter on Human and Peoples’ Rights, article 22
- Requirement for good faith consultations – Indigenous and Tribal Peoples Convention, 1989 (No. 169), article 6
• Conditionality of development assistance – Commission on Human Rights, Working Group on the Right to Development, High-level task force on implementation of the right to development, E/CN.4/2005/WG.18/TF/3, paragraph 32
• Involvement of women in the development process – Subcommission on Human Rights resolution 1999/15

HIV/AIDS

• Right to non-discrimination – International Covenant on Civil and Political Rights, article 2
• Criminal law should not be an impediment to HIV/AIDS services – International Guidelines on HIV/AIDS and Human Rights, guideline 4 (paragraph 21 (d))
• Criminal law, sex workers and HIV/AIDS – International Guidelines on HIV/AIDS and Human Rights, guideline 4 (paragraph 21 (c))
• Mandatory HIV testing – International Guidelines on HIV/AIDS and Human Rights, recommendations, paragraph 120
• HIV/AIDS and prisoners – International Guidelines on HIV/AIDS and Human Rights, guideline 4 (paragraph 21 (e))
• HIV-specific criminal laws – International Guidelines on HIV/AIDS and Human Rights, guideline 4 (paragraph 21 (a))

### ANNEX 4. UNITED NATIONS EVALUATION GROUP ETHICAL GUIDELINES FOR EVALUATION

<table>
<thead>
<tr>
<th>GUIDELINES FOR IMPLEMENTING THE PRINCIPLES (CONT’D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>Integrity</strong></td>
</tr>
<tr>
<td>- Alignment with ethical principles: Ensure that your conduct and work are guided by and aligned with ethical principles and professional standards. This requires ensuring that you are fully aware of ethical guidelines and professional standards of conduct, review them regularly and reflect continuously on how best to live up to them.</td>
</tr>
<tr>
<td>- Communication: Communicate truthfully and openly with clients and relevant stakeholders concerning aspects of the evaluation, such as findings, procedures, limitations or changes that may have occurred. Negotiate honestly when estimating the necessary amount of work, related payment and workload actually performed.</td>
</tr>
<tr>
<td>- Professional development: Continually undertake professional development and exchange to strengthen qualification and expertise for competent and ethical evaluation practice. This can be achieved through formal training, professional exchange, supervision or informal collegial discussions, particularly in relation to ethical challenges in evaluation.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GUIDELINES FOR THOSE WHO CONDUCT EVALUATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>- <strong>Accountability</strong></td>
</tr>
<tr>
<td>- Redress: Ensure participants know how to seek redress for any perceived disadvantage or harm suffered from the evaluation and how to register a complaint concerning the unethical conduct of those involved in the conduct or organization of the evaluation.</td>
</tr>
<tr>
<td>- Evaluative judgements: Demonstrate in the inception and evaluation reports that the evaluation is conducted in a rigorous, fair and balanced way and that any judgements made are based on sound and complete evidence that can be verified.</td>
</tr>
<tr>
<td>- <strong>Respect</strong></td>
</tr>
<tr>
<td>- Inform: Ensure prospective participants in the evaluation know what they are being consulted on and why; what the intended outputs are; and have sufficient and adequate information for informed consent. This includes explicitly noting their right to skip questions or withdraw consent at any stage of the evaluation process without fear of penalty.</td>
</tr>
<tr>
<td>- Confidentiality: Respect participants’ right to provide information in confidence and ensure that participants fully understand the scope and limits of confidentiality. Evaluators must ensure that sensitive information cannot be traced to its source so that individuals are protected from reprisals.</td>
</tr>
<tr>
<td>- Inclusion and non-discrimination: Ensure equitable participation and treatment of all participants and their opportunity to voice their perspectives. Respect differences in culture, local customs, religious beliefs and practices, personal interaction, gender roles, ability, age and ethnicity, and be mindful of the potential implications of these differences when carrying out and reporting on evaluations. Where the evaluation involves the participation of members of vulnerable groups, evaluators must be aware of and comply with international and national legal codes governing respecting and protecting the rights of these groups (e.g., guidelines on researching and interviewing children and young people).</td>
</tr>
<tr>
<td>- Show empathy: Empathize and work collaboratively with all stakeholders, treating evaluation participants, evaluators and evaluation commissioners in a way that honours their professional expertise and personal dignity.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GUIDELINES FOR IMPLEMENTING THE PRINCIPLES (CONT’D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- <strong>Beneficence</strong></td>
</tr>
<tr>
<td>- Clear benefits: As far as possible, position the evaluation as an intervention working towards the realisation of human rights, gender equality and achievement of the SDGs and be clear about the implications of this positioning. Push for and fully implement participatory and empowerment approaches and other forms of stakeholder consultation to maximize potential benefits to the evaluation and those involved in it.</td>
</tr>
<tr>
<td>- Addressing power imbalances: Make sure the voices of the most vulnerable are heard, including when data collection is remote. Recognize, report on and attempt to address or mitigate potential power imbalances within the evaluation approaches adopted.</td>
</tr>
<tr>
<td>- Do no harm: Ensure that ongoing risk assessments clarify and mitigate potential and actual harms that may arise and that go beyond what participants have consented to. This risk assessment must be ongoing and evaluations should not proceed where mitigation (through, for example, use of alternative methods) is not possible and harm will ensue. Where unanticipated harm has been identified, ensure redress channels are triggered. (See Box 1 for examples of harm.)</td>
</tr>
</tbody>
</table>

---

Beneficence means that 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. This involves the usual considerations plus additional consideration to avoid further perpetuation of exclusion, unmet expectations and distress.

Beyond harm to participants, the ‘do no harm’ principle also requires consideration of potential harm to evaluators themselves, particularly in terms of safety, potential trauma, culture shock and availability of emotional support. Conversely, there may be situations where powerful stakeholders seek to divert evaluator attention away from potentially confronting or examining uncomfortable areas or truths under the guise of ‘do no harm’. Evaluators need to apply professional scepticism and watch out for risks, but also proceed without fear or favour and carefully, respectfully and intelligently uncover those truths. In turn, evaluators must ensure that they use the principle appropriately and not in order to shy away from difficult conversations.

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.
ANNEX 5. UNITED NATIONS EVALUATION GROUP PLEDGE OF ETHICAL CONDUCT IN EVALUATION

Note that the following pledge of commitment to ethical conduct in evaluation can also be downloaded as a separate document from the United Nations Evaluation Group website.

ANNEX 1: PLEDGE OF ETHICAL CONDUCT IN EVALUATION

By signing this pledge, I hereby commit to discussing and applying the UNEG Ethical Guidelines for Evaluation and to adopting the associated ethical behaviours.

INTEGRITY

I will actively adhere to the moral values and professional standards of evaluation practice, as outlined in the UNEG Ethical Guidelines for Evaluation and as per the values of the United Nations. Specifically, I will be:

- Honest and truthful in my communication and actions.
- Professional, engaging in credible and trustworthy behaviour, alongside competence, commitment and ongoing reflective practice.
- Independent, impartial and incorruptible.

ACCOUNTABILITY

I will be answerable for all decisions made and actions taken, responsible for honoring commitments, without qualification or exception, and will report potential or actual harms observed. Specifically, I will be:

- Transparent regarding evaluation purpose and actions taken, establishing trust and increasing answerability on performance to the public, particularly those populations affected by the evaluation.
- Responsive as questions or events arise, adapting plans as required and referring to appropriate channels where corruption, fraud, sexual exploitation or abuse or other misconduct or waste of resources is identified.
- Responsible for meeting the evaluation purpose and for actions taken, and for ensuring redress and recognition as needed.

RESPECT

I will engage with all stakeholders of an evaluation in a way that honours their dignity, well-being, personal agency and characteristics. Specifically, I will ensure:

- Access to the evaluation process and products by all relevant stakeholders—be they powerless or powerful, with due attention to factors that may impede access such as sex, gender, race, language, country of origin, LGBTQ status, age, background, religion, ethnicity and ability.
- Meaningful participation and equitable treatment of all relevant stakeholders in the evaluation processes—from design to dissemination. This includes engaging different stakeholders, particularly affected people, so they can actively inform the evaluation approach and products rather than being solely a subject of data collection.
- Fair representation of different voices and perspectives in evaluation products (reports, webinars etc.)

I commit to playing my part in ensuring that evaluations are conducted according to the Charter of the United Nations and the ethical requirements laid down above and contained within the UNEG Ethical Guidelines for Evaluation. Where this is not possible, I will report the situation to my supervisor, designated focal points or channels, and will actively seek an appropriate response.

BENEFICENCE

I will strive to do good for people and planet while minimizing harm arising from evaluation as an intervention. Specifically, I will ensure:

- Explicit and on-going consideration of risks and benefits from evaluation processes.
- Maximum benefits at systemic (including environmental), organizational and programmatic levels.
- No harm. I will not proceed where harms cannot be mitigated.
- Evaluation makes an overall positive contribution to human and natural systems and the mission of the United Nations.

(Signature and Date)
ANNEX 6. EXAMPLE OF A SUCCINCT SET OF UNITED NATIONS OFFICE ON DRUGS AND CRIME EVALUATION QUESTIONS

The following is an excerpt from the final in-depth evaluation of the United Nations Office on Drugs and Crime (UNODC) Global Programme on the Implementation of the Doha declaration: Towards the Promotion of a Culture of Lawfulness (GLO/Z82) (November 2020)

DESIGN AND RELEVANCE

Design and relevance questions broadly pertain to the strategic fit achieved by the programme in helping UNODC deliver on its mandates. The greater the fit between strategy and environment, the greater is the continued relevance of the programmes and policies to UNODC and its key stakeholders.

Evaluation question: To what extent are the outputs, outcomes and objectives of this Global Programme relevant to its key stakeholders, implementing the Sustainable Development Goals and related national strategies? To what extent do its objectives and design continue to respond to their needs, policies and priorities, under changed circumstances such as the COVID-19 pandemic?

EFFICIENCY

Programme efficiency is concerned with the relationship between inputs and outputs.

Evaluation question: To what extent is the relationship between inputs and outputs timely, cost-effective and meeting expected standards?

EFFECTIVENESS

The effectiveness criterion is concerned with the relationship between outputs and outcomes and is the focus of the evaluation.

Evaluation question: To what extent have the objectives and outcomes of the programme been achieved or can be expected to be achieved in all streams of work? What have been the facilitating or hindering factors in achievement of results in all streams of work?

COHERENCE

Coherence seeks to understand how well the intervention fits with other interventions in a country, sector or institution.

Evaluation question: Were appropriate synergies sought, created, maintained and used, including within UNODC, with United Nations agencies, civil society organizations and academia?

IMPACT

Impact refers to the achievement of objectives pertaining to long-term benefits to targeted beneficiaries, including institutional, policy and social transformations.

Evaluation question: What were the overall intended and unintended impacts or long-term results, if any, of the Global Programme for the promotion of the rule of law and a culture of lawfulness? What contribution did each of the four substantive components make to the attainment of the Sustainable Development Goals?

SUSTAINABILITY

The sustainability-related questions focus on both the long-term financial sustainability of the projects and the self-sustainability of the results achieved as a result of the programme, with or without further support. Even though sustainability is essentially a process variable, it can be measured on the basis of an analysis of how the programme has delivered its outcomes, its relationship with donors and beneficiaries and how it is placed within wider organizational strategies and embedded in the organizational structure.

Evaluation question: To what extent: will the change achieved for the partner institutions and beneficiaries, if any, be long-term and persist after the end of the Global Programme?

HUMAN RIGHTS, GENDER EQUALITY AND LEAVING NO ONE BEHIND

Evaluation questions:
To what extent were human rights considerations, especially with regard to underrepresented and vulnerable groups, included in the programme design and implementation?
To what extent were gender equality considerations included in the programme design, implementation and monitoring, and reporting on results?
ANNEX 7. EXAMPLE THEORY OF CHANGE: REGIONAL PROGRAMME FOR SOUTH-EAST ASIA AND THE PACIFIC, 2021–2026

**Key assumptions (not exhaustive)**
1. Member States and relevant regional bodies (e.g., ASEAN) are willing to work with UNODC and to cooperate with each other on the issues falling under UNODC mandates.
2. Governments engage on commitments to policy/operational reforms with UNODC.
3. UNODC supports and promotes UNODC’s mandates in relevant regional bodies and events.
4. Engaged stakeholders have sufficient financial and human capacity to engage meaningfully.
5. Public and private stakeholders understand the benefits to engage on workstreams developed by UNODC and its partners.

**External factors**
- Political and financial commitments
- Geographical and economic factors in partner countries
- Government commitment to reduce transnational organized crime, trafficking in persons, terrorism and other forms of organized crime
- Reintegration of victims and victims of human trafficking
- Inclusion of vulnerable groups, such as persons with disabilities
- Women and non-state actors

**Rapid Response**

**Thematic Areas**
- Transnational Organized Crime
- Corruption
- Terrorism
- Criminal Justice
- Drugs and Health, and Alternative Development

**Components**
- Partnerships and Advocacy
- Policy and Regulatory Frameworks
- Capacity Development
- Availability and Access
- Data and Evidence

**Outcomes**
- Engaged public and private stakeholders share and promote innovative solutions and delivery models.
- Governments develop policy initiatives in line with international laws, good practices and standards.
- Participants become active agents of change towards safer, healthier, fairer and greener societies.
- People in Southeast Asia and the Pacific benefit from the UNODC’s value proposition in the region.

**Ultimate Impacts**
- Member States accelerate on their pathways to address challenges against crime in all its dimensions, the drug problem, and to prevent terrorism in South East Asia and the Pacific.
- UNODC and its partners significantly contribute to safer, fairer and more protected environments, as key enablers for a peaceful, prosperous and sustainable region for all.
- Reduce all forms of violence and related death rates, especially of women and girls, and improve population safety and well-being, leaving no one behind.
- Promotion and protection of Human Rights is at the core of the drugs, crime and terrorism related national and regional agendas.
ANNEX 8. INDICATIVE EVALUATION QUESTIONS BY CRITERION

INDICATIVE QUESTIONS: RELEVANCE

• To what extent does the intervention align with and support national policies and priorities?
• To what extent does the intervention’s needs analysis utilize reliable and appropriate data, including United Nations Office on Drugs and Crime (UNODC) knowledge products (e.g., the World Drug Report)?
• To what extent were the implemented activities suited to the priorities and policies of the justice system actors, law enforcement officers and civil society?
• Is the design of the intervention appropriate for addressing cross-cutting issues such as equity and vulnerability, gender equality and human rights and disability inclusion?

INDICATIVE QUESTIONS: COHERENCE AND PARTNERSHIP

• To what extent is the process for cooperation and partnership in place and appropriate?
• Is the set of partnerships appropriate? Are there other potential partners that UNODC should seek to engage?

INDICATIVE QUESTIONS: EFFECTIVENESS

• To what extent have the intervention’s intended objectives and outcomes been achieved?
• What were the main internal and external factors affecting the achievement or non-achievement of the intervention’s intended and unintended results?
• To what extent did the facilitation of the training component align with good practice, and how could it have been improved?
• To what extent was the results monitoring system able to capture shorter- and longer-term outcomes, as distinct from outputs (e.g., number of training sessions delivered)?

INDICATIVE QUESTIONS: EFFICIENCY

• To what extent was the intervention implemented in a cost-efficient manner?
• What measures were taken during the implementation of activities to ensure that the resources were used appropriately and efficiently, and converted into outputs in a timely manner?
• Are risks being adequately identified and addressed?
• To what extent is the intervention’s governance structure adequate?

INDICATIVE QUESTIONS: IMPACT1 AND SUSTAINABILITY

• To what extent are the actions within this intervention likely to reduce [intervention target, e.g., serious organized crime]?
• Is the current set-up of the intervention sustainable in the long term, including in respect to the ownership of implementing partners?
• To what extent did the project contribute to achieving the relevant Sustainable Development Goals?
• To what extent has UNODC assistance helped to generate confidence and commitment of the justice system actors in the work of the service providers, and cooperation among the actors that is likely to continue after the intervention ends?

1 Because of the funding and administrative constraints of UNODC, evaluations at UNODC typically address questions regarding impact in terms of visible or likely contribution, as per the OECD Development Assistance Committee’s definition (“Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended;” see www.oecd.org/dac/evaluation/2754804.pdf). This differs from the more restricted approach to impact evaluation common in other institutions, such as the Abdul Latif Jameel Poverty Action Lab, the World Bank and the Inter-American Development Bank, among others, where the impact attributable to the intervention is estimated by means of randomized control trials or quasi-experimental methods, which involve collecting baseline and follow up data from “treatment” and “comparison” or “control” groups.
INDICATIVE QUESTIONS: HUMAN RIGHTS AND GENDER EQUALITY

• To what extent are gender equality and human rights considerations integrated into the design and implementation of the intervention?
• What is preventing duty bearers from meeting their obligations?
• Is disaggregated data being produced and/or collected?
• Is inequality and discrimination addressed from the start? Are all duty bearers actively and meaningfully participating?

INDICATIVE QUESTIONS: “LEAVE NO ONE BEHIND”

• Are UNODC programmes looking at whether laws, statistics and research consider inequalities among all population groups?
• To what extent is the intervention informed and updated based on the needs of diverse stakeholders?
• Which groups are left behind, which are furthest left behind within vulnerable populations, and to what extent was the intervention tailored to meet their needs?
• Does UNODC make information accessible to everyone, considering issues such as disability, language, and literacy levels, for example?

INDICATIVE QUESTIONS: DISABILITY INCLUSION

• To what extent was there consideration of persons with disabilities and of the range of needs that different populations have?
• What proportion of rights holders were persons with disabilities?
• What barriers did persons with disabilities face?
ANNEX 9. CHECKLIST FOR CONFLICT-SENSITIVE EVALUATIONS AND MEASUREMENT ON SENSITIVE SUBJECTS

The purpose of this checklist is to bring together good practices from other toolkits and resources relating to the conducting of evaluations on sensitive subjects and in complex and conflict-affected environments. The source documents are provided at the end of the table.

### REFLECT AND PLAN

- Define the purpose of data collection and the extent to which conflict, gender and human rights are incorporated into the evaluation scope and programme or evaluand design.

- Define the stakeholders to be engaged, paying special attention to specific vulnerabilities related to the contexts or identities, and ensure those most affected by rights violations are included in the stakeholder analysis. In the context of crime prevention programmes, this might include victims, witnesses and/or perpetrators in the context of, for example, trafficking in persons.

- Ensure that the population to be engaged does not include children under the age of 18, as this requires more specific protocols for ensuring safety and informed consent.

- Has a conflict analysis and/or gender analysis been conducted by the programme or project at any point, either before, during or after?

- Do the benefits of collecting sensitive data outweigh the risks? Determine the challenges in evaluating human rights and gender equality that are to be addressed during the evaluation, upholding the principle of “do no harm”.

### DESIGN

- Determine stakeholder accessibility and how stakeholders are to be engaged in the evaluation. In conducting evaluations in insecure areas or on highly sensitive subjects, sometimes approaches that are fit for purpose and/or innovative are required to engage diverse stakeholders while upholding the principle of “do no harm”. These can include drawing on secondary sources, identifying proxies (either in indicators and/or in representatives speaking for conflict-affected groups, such as community-based organizations), changing the geographical location of data collection or switching to remote means.

- Given the potential safety risks and the important survival tasks occupying respondents in conflict settings, ensure data collection activities are designed in a manner that ensures that they are of high quality, not time-intensive, and safety-centred.

- Assess how human rights, gender and conflict analysis are integrated into the evaluation criteria:
  - Under the criterion of relevance or a cross-cutting criterion, such as human rights and gender equality, in the evaluation matrix, include indicators and/or assumptions that assess whether there are or were plans to conduct or update the conflict and human rights and gender equality analyses, including in relation to key events, and use this information to inform any adjustments needed to the programme.
  - In addition, include assumptions and/or indicators in the evaluation matrix that specifically assess whether beneficiary feedback mechanisms exist and whether they track conflict sensitivity concerns.
  - Under efficiency, it is also good practice to assess the extent to which human rights and gender equality and conflict has been considered in the programme results framework, including whether indicators that can track relevant conflict trends are included.
• Evaluations are also typically required to assess whether there have been any unintended impacts, and for this reason, evaluation matrices should include both positive and negative indicators and conflict interaction indicators which track key risks of harm that might be attributable to the programme.
• Other possible indicators and/or assumptions to be assessed:
  – Has conflict sensitivity been included as a line of enquiry within annual reviews and other evaluations? If so, what were the findings and how were they responded to?
  – Is there a section within implementing partner reporting templates for reporting on conflict sensitivity concerns? How have partners monitored and managed conflict?

Related to the above, ensure a mixed-method approach, as well as a mix of quantitative indicators (which measure whether something exists and in what quantity) and qualitative indicators (which measure quality or perceptions of processes and/or efforts), that can be disaggregated.

Ensure the evaluation methodology favours stakeholders’ right to participation, including those most vulnerable.

---

**PREPARE**

- Establish a line of communication with community-based organizations and advocates that supports the population of interest (such as local women’s, victims’ and disability inclusion advocates or organizations) to gauge stakeholder accessibility.
- Engage the national evaluators, interpreters or other on-the-ground professionals and organizations that are most familiar with national contexts.
- Seek guidance from other United Nations entities operating in the area on protocols and good practice within the specific context, country or region of data collection.
- For evaluations that collect information on sensitive issues, such as gender violence or mental health, identify (or develop) a local referral system to ensure the safety of respondents.
- Identify the most appropriate and contextually relevant tool and/or questions for gathering necessary data, and do not include any unnecessary or desirable but optional questions on sensitive issues; all information gathered should be critical to improving the situations of respondents.
- Familiarize yourself with ethical protocols for protecting respondents and their data (e.g., with regard to confidentiality or informed consent).
- Develop data security and communication and dissemination plans for data collection and reporting, and include all necessary and required information in the informed consent documents.
- Ensure that all members of the evaluation team are well informed, have appropriate skill sets and are trained on ethical protocols and evaluation methods and tools. For sensitive issues such as gender violence, it may be necessary to specifically engage a protection professional to conduct the ethics and referrals training.
- Ensure consent documents and data collection tools are written and provided in accessible languages.

---

**COLLECT**

- When collecting data on sensitive subjects, it is important to go through several rounds of internal reviews, as well as to pilot the data collection tools for the purpose of making final improvements.
- Identify a space for interviews and surveys that guarantees privacy. If done remotely, request that respondents be in a private place and/or ask whether they may experience challenges in accessing a private and secure place. Do not put any respondent at risk in data collection processes.
Closely follow logistics and scripts in the data collection plan and tools.

Team leaders and senior evaluators should provide regular oversight, feedback and support to team members collecting data on sensitive subjects.

Ensure the collection of disaggregated data, for example, by ethnic group, sex, age, religion or disability status, so that differential impacts on different groups can be measured.

**ANALYSE AND REPORT**

- Ensure that all stakeholder groups are consulted, or at least represented in some capacity, at the end of the data collection stage to discuss findings and hear their views on the conclusions and recommendations of the evaluation.

- Be mindful of how evaluation findings are portrayed, ensuring they do not unintentionally identify any one individual or organization and create or exacerbate conflict in any way; evaluation reports are commissioned with the expectation that they will be actionable and do good.

- As with any evaluation, apply triangulation of different types of data and sources to ensure robust findings not influenced by any one person or group.

- De-identify all data once reporting is done and strictly follow all data protection plans.

- Include the disaggregated data in the report findings, and develop recommendations to address any differential impacts.

- Ensure that at least one recommendation addresses identified human rights and gender equality issues.

**DISSEminate**

Ensure awareness of how the recommendations affect different stakeholder groups and ensure that there are clear plans in place for disseminating to a wide group of stakeholders who are interested and/or affected by conflict and human rights and gender equality issues; this will typically be done by the independent evaluation office or commissioning body directing the evaluation, although it is often helpful for evaluation teams to provide guidance and consultation on an appropriate dissemination plan.

This checklist draws from the following:

- United Kingdom Stabilisation Unit, “Conflict sensitivity tools and guidance” (June 2016) (see tool 5, “Checklist for conflict sensitivity within monitoring and evaluation plans”)

- Global Women’s Institute, Gender-based Violence Research, Monitoring and Evaluation with Refugee and Conflict-affected Populations (George Washington University, 2017)


ANNEX 10. ADVANTAGES AND DISADVANTAGES OF COMMON DATA COLLECTION TOOLS AND APPROACHES IN THE CONTEXT OF CRIME PREVENTION

DESKTOP REVIEW

Advantages:
- Time and cost efficiencies, using readily available data
- Internal reporting can provide the benefit of in-depth insights gathered by project staff and partners who have already built trust with difficult-to-reach stakeholders who are less accessible in one-time evaluation efforts

Disadvantages:
- Potential for bias towards positive results; highly dependent on project staff to collect data and analyse and report on both successes and limitations or weaknesses
- Data are not always systematically collected on project indicators

REVIEW AND MINING OF SECONDARY DATA

Advantages:
- Useful for assessing normative work, especially the development of an international norm, standard or convention
- May reveal lessons about successful and unsuccessful practices under similar contexts and circumstances that can be compared to the evaluation subject
- Can also be used to reconstruct baseline data sets when baseline assessments were not conducted

Disadvantages:
- Time-consuming to gather sufficient levels of evidence
- Can require extensive stakeholder engagement and consents if not all sources or databases are immediately available online or are openly available

INTERVIEWS WITH KEY INFORMANTS

Advantages:
- Used to collect in-depth information about a specific topic from an expert or experts
- Allows anonymity through a one-on-one or small group interview format

Disadvantages:
- There is a possibility that response bias will emerge from findings if the sample of respondents is not large enough
- Time-consuming to schedule, especially for more senior officials with limited availability

FOCUS GROUP DISCUSSIONS

Advantages:
- Discussions generate a wide range of opinions and provide the opportunity to exchange ideas (e.g., youth can learn from police officers and vice versa)
- Allows the opportunity for immediate cross-checking of information if it is unclear
- Especially in conjunction with surveys, it can be used to reach a large number of stakeholders, in particular most vulnerable populations, in a short amount of time

Disadvantages:
- Facilitators may bring personal bias or perspectives to the group discussion and thereby skew results
- Social desirability bias occurs where participants answer questions dishonestly in order to be viewed favourably by other participants or the facilitator
- Participants may not share sensitive information, as it is not experienced as anonymous when others are present, especially when rigid hierarchies among participants are present
- Can be difficult to arrange, as it may involve bringing together larger groups of people with competing schedules
SURVEYS

Advantages:

• Time and cost efficiencies through use of technology and generation of larger volumes of quick responses
• Enables respondents to be anonymous and therefore potentially elicits more truthful responses
• Can be used to collect information in a short period of time from a larger group of stakeholders who may otherwise be inaccessible, especially when self-administered

Disadvantages:

• Unless conducted in-person through fieldwork, access is limited to those who are conversant with Internet or phone access; in contrast, fragile security contexts, or time or resource constraints may hinder fieldwork and the obtaining of responses from a sufficiently large and representative sample
• Response rates are typically low when conducted remotely
• Does not allow for responses to clarifying or follow-up questions, so questions must be very clear and precise to avoid misinterpretation
• Results are open to bias when participants are self-selected; random samples, although still having limits, provide greater confidence in their representativeness

OBSERVATIONS

Advantages:

• Can be used to identify why an intervention is not working well (intervention quality). For example, it can be useful for understanding the process used by Member States in developing a norm or standard, stakeholder participation, the quality of partnerships and normative leadership
• Can be used when formalized data collection methods are viewed negatively or when a population is resistant to being interviewed

Disadvantages:

• Difficult to arrange and gain access to activities, owing to security and human rights protocols surrounding the issues of interest
• Time-consuming to arrange and observe
• Typically done without a systematic tool or process for measurement, which is not good practice and makes the method too subjective
• May encourage “false positives”, whereby participants may behave differently, knowing that the evaluator is present

SELF-REPORTED CHECKLISTS AND SCORECARDS

Advantages:

• Can be integrated into observations or surveys

Disadvantages:

• Self-assessments are often positively biased, so it is important to collect multiple lines of evidence and triangulate well
CASE STUDIES

Advantages:
- Useful when the primary evaluation objective is to gain in-depth understanding of the context for normative work and the rationale for the successful integration of international standards or to determine the kind of United Nations support needed to assist Governments at the national level in implementing normative instruments.
- The unit of analysis can be a country, a programme, a theme or a beneficiary or stakeholder group, making case studies particularly useful for global and regional programmes with intersecting subprojects, or for a thematic area addressed by UNODC.

Disadvantages:
- Time- and resource-intensive.
- Poses an inherent risk of generalizing findings across a smaller selection of cases (fewer than five, for example).
- They need to be focused on information-rich cases to justify costs and efforts, although cases about which there is a lack of information might be of greater interest.

BENCHMARKING

Advantages:
- Within normative work, case studies are helpful in comparing the policies, legislation or action plans of Governments against a benchmark or with an ideal example that meets all United Nations criteria.

Disadvantages:
- Highly contextual and require clear understanding and analysis of the development context and extenuating circumstances in some countries, such as limited legislative drafting capacity or lack of political will, which may make it difficult to achieve benchmarks.
ANNEX 11. PRACTICAL APPLICATION OF EVALUATION METHODS RELEVANT TO THE AREAS OF WORK OF THE UNITED NATIONS OFFICE ON DRUGS AND CRIME

CONTRIBUTION ANALYSIS


- Outcomes and/or populations of interest: the influence of research and knowledge exchange on public health policy and programmes and the impact on health at the population level


- Outcomes and/or populations of interest: assessment of causal links between advocacy efforts on the adoption of specific policies to increase the transparency, efficiency and effectiveness of United States foreign assistance

QUALITATIVE COMPARATIVE ANALYSIS


- Outcomes and/or populations of interest: assessment, or disentanglement, of the relationships between contextual conditions and outcomes; in the example provided, this was specifically sought to answer how citizens, through media, provoke responses from powerful actors or duty bearers (e.g., politicians and businesses) in Kenya and the United Republic of Tanzania

PROCESS TRACING

Alix Wadeson, Bernardo Monzani and Tom Aston, “Process tracing as a practical evaluation method: comparative learning from six evaluations” (March 2020); available at http://mande.co.uk

- Outcomes and/or populations of interest: Presents six case studies in which process tracing was variably applied across different contexts. Specifically, three evaluations employed contribution tracing and three used process training in combination with other methods such as contribution analysis and outcome harvesting; areas of focus spanned from infrastructure, value chains and education to climate change, but all focused on governance and advocacy. The most relevant example of an assessed contribution claim for United Nations Office on Drugs and Crime evaluations was on the organization’s leadership role in successfully influencing Governments to uphold their commitment to the Paris Agreement.

OUTCOME HARVESTING


- Outcomes and/or populations of interest: Identifies and documents more than 300 outcomes from 111 countries related to sustainable livelihoods, gender justice and strengthening the voice of marginalized people through the promotion of democratic principles and practices in global governance arrangements

Kornelia Rassmann and others, “Retrospective ‘outcome harvesting’: generating robust insights about a global voluntary environmental network” (2013); available at http://betterevaluation.org

- Outcomes and/or populations of interest: This paper describes the use of the outcome harvesting methodology adopted by the global Biodiversity Action Network to assess observable changes in behaviour, relationships and actions of individuals, groups and organizations with whom the network works towards biodiversity conservation and food security. It also discusses measures taken to increase validity and the mitigation of challenges.


- Outcomes and/or populations of interest: This report documents 10 pilot experiences, where teams retrospectively applied outcome harvesting to identify new learnings from multi-stakeholder projects. The areas covered include public sector reform and leadership, health, governance and education.
### MOST SIGNIFICANT CHANGE


- **Outcomes and/or populations of interest:** This evaluation assessed and captured institutional change through use of the “most significant change” method. Staff were asked to identify the most significant change in their thematic area over the period 2008–2019, as well as the outcomes associated with that change.

### MAPPING AND SOCIAL NETWORK ANALYSIS

Alecia Kallos, “Social network analysis and evaluation: learnings from the evaluator and the client”; available at www.evalacademy.com

- **Outcomes and/or populations of interest:** This article describes lessons learned from applying social network analysis for a community development initiative. While the actual evaluation report is not provided, an example of graphed network data is provided, as well as other resources, examples and good practices to consider.

### PARTICIPATORY RANKING

Louise Knight and others, “Intervention with Microfinance for AIDS and Gender Equity (IMAGE): women’s engagement with the scaled-up IMAGE programme and experience of intimate partner violence in rural South Africa” *Prevention Science*, vol. 21, No. 1 (2020), pp. 268–281

- **Outcomes and/or populations of interest:** This is a rigorous evaluation to assess the Intervention with Microfinance for AIDS and Gender Equity (IMAGE) programme of South Africa. It used participatory methodologies such as participatory wealth ranking; community profiling; monitoring of Sisters for Life training sessions; key informant, household and loan group monitoring; and a community-based participatory rural appraisal project.


- **Outcomes and/or populations of interest:** This report uses group participatory ranking methodology to inform and document prevalence of gender-based violence among internally displaced persons in northern Uganda in 2006. Through use of structured participatory methods, participants ranked and narrated accounts of problems faced by women in the camps.

### COST-BENEFIT ANALYSIS AND COST-EFFECTIVENESS ANALYSIS


- **Outcomes and/or populations of interest:** This toolkit, developed by the Australian Institute of Criminology in partnership with the Community Crime Prevention Unit of the Department of Justice and Community Safety of the State of Victoria, Australia, specifically seeks to help public institutions to evaluate public safety infrastructure projects. The ready-made Microsoft Excel file to support cost-benefit analysis can be used to estimate the financial benefits of a project within the first year.

The Vera Institute for Justice hosts a Cost-Benefit Knowledge Bank for Criminal Justice, which includes a toolkit for conducting cost-benefit analysis, including a step-by-step set of instructions for evaluators, as well as a background paper on the relevance of cost-benefit analysis to criminal justice planning.