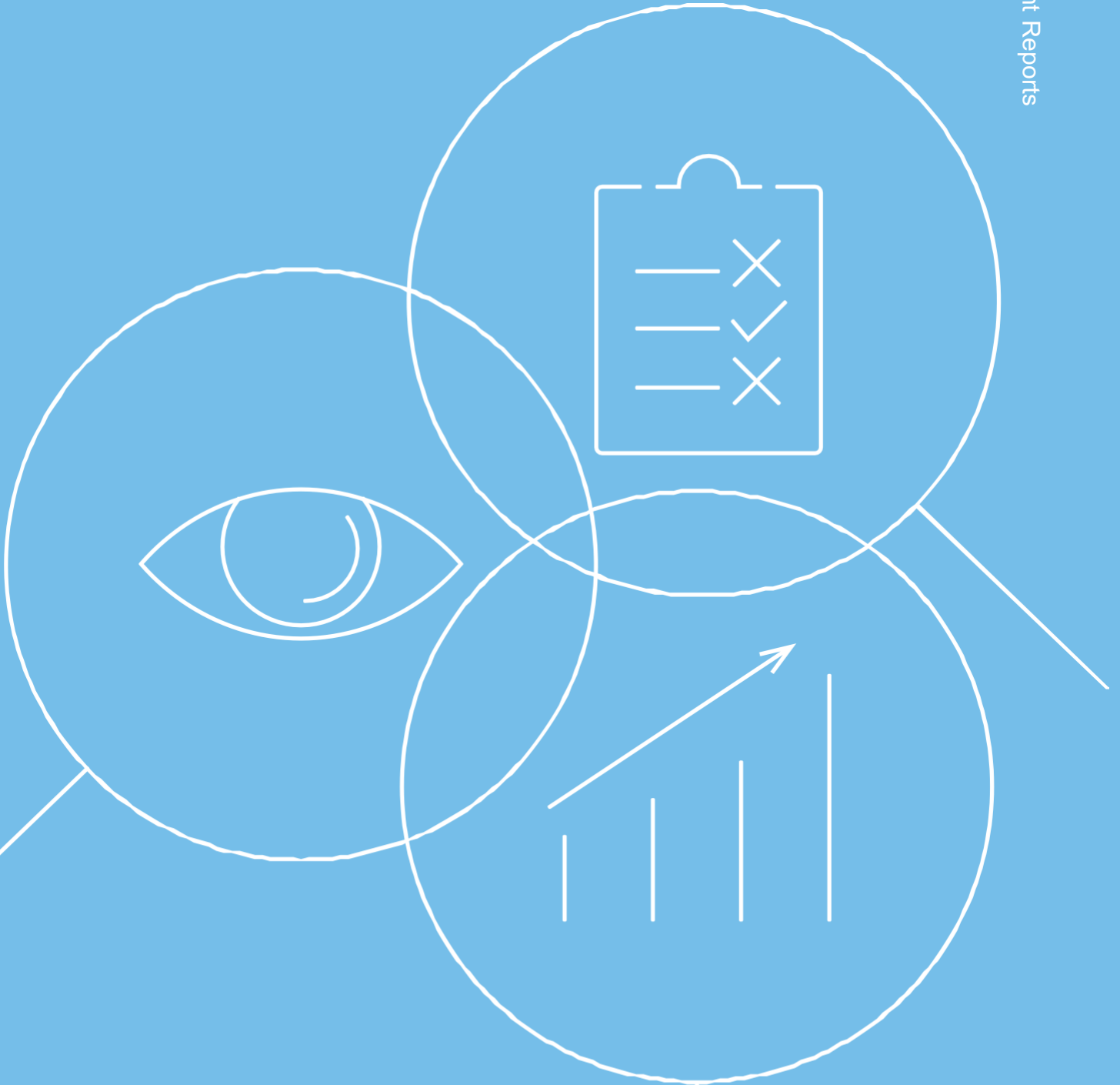


Evaluation of the Use and Impact of IOD Evaluation Section Recommendations

Internal Oversight Reports



IOD Ref: EVAL 2020-03
September 10, 2021
Evaluation Section

This guide results from a collaboration between iNudgeyou and the World Intellectual Property Organization, Evaluation Section. iNudgeyou is a pro-social company specializing in applied behavioral research. Its primary purpose is to develop and communicate evidence-based strategies with pro-social purposes.

© The World Intellectual Property Organization, 2021

This work is licensed under the Creative Commons Attribution-Noncommercial No-Derivatives 3.0 IGO License. The user is allowed to reproduce, distribute, and publicly perform this publication without explicit permission, if the content is accompanied by an acknowledgment that WIPO is the source. No part of this publication can be used for commercial purposes or adapted/translated/ modified without the prior permission of WIPO.

To view a copy of the license, please visit <http://creativecommons.org/licenses/by-ncnd/3.0/igo/>. Please write to iod@wipo.int to obtain permission. When content, such as an image, graphic, data, trademark, or logo, is attributed to a third party, the user is solely responsible for clearing the rights with the right holders.

Suggested citation: WIPO (2021). *Evaluation of the Use and Impact of IOD Evaluation Section Recommendations*, Geneva.

TABLE OF CONTENTS

1. INTRODUCTION.....	6
WHY THIS GUIDE?.....	6
WHOM IS THIS GUIDE FOR?	7
WHAT IS EVALUATION?	9
WHY ARE BEHAVIORAL INSIGHTS IMPORTANT FOR EVALUATIONS?	9
EVALUATION CRITERIA.....	10
WHAT ARE THE KEY SUCCESS FACTORS FOR IMPLEMENTING A BEHAVIORAL INSIGHTS APPROACH IN EVALUATIONS?.....	11
WHY ARE EVALUATION AND BEHAVIORAL INSIGHTS IMPORTANT?	12
WHAT ARE THE TYPES AND DIMENSIONS OF EVALUATION?	13
HOW ARE EVALUATION AND BEHAVIORAL SCIENCE DIFFERENT?	14
2. HUMAN RIGHTS, GENDER EQUALITY AND BEHAVIORAL INSIGHTS EVALUATIONS.....	18
HUMAN RIGHTS AND GENDER EQUALITY CONSIDERATIONS IN EVALUATION	18
3. THE EVALUATION PROCESS INTEGRATING BEHAVIORAL INSIGHTS.....	21
PHASE 1: COMMITMENT	24
STEP 1- PREPARATION: SOCIAL EXPECTATIONS	25
STEP 2 – LAUNCH: MAKE IT EASY AND SOCIAL	26
PHASE 2: DIAGNOSTIC.....	32
STEP 3 – DESIGN – MAKE IT INTUITIVE.....	33
STEP 4 – FIELDWORK – SUPPORT JUDGMENT	39
PHASE 3: IMPACT	45
STEP 5 – REPORTING – PLAN FOR INATTENTION.....	46
STEP 6 – FOLLOW-UP – CHANGE PROCESS.....	48
ANNEXES.....	53

LIST OF ACRONYMS

BE	Behavioral Economics
BI	Behavioral Insights
BIE	Behavioral Insights Evaluation
EoI	Expression of Interest
EQM	Evaluation Question Matrix
GE	Gender Equality
HR	Human Rights
IOD	Internal Oversight Division
IP	Intellectual Property
JIU	Joint Inspection Unit of the United Nations System
OIOS	Office of Internal Oversight Services
OECD	Organization for Economic Co-operation and Development
OECD-DAC	OECD – Development Assistance Committee
SMART	Specific, Measurable, Achievable, Realistic and Time-bound
ToR	Terms of Reference
UN	United Nations
UNEG	United Nations Evaluation Group
UN SWAP	UN System-wide Action Plan on Gender Equality and the Empowerment of Women
WIPO	World Intellectual Property Organization



INTRODUCTION



1. INTRODUCTION

WHY THIS GUIDE?

The assumption in a rationalistic organizational model is that the organization as a system always acts rationally to evaluation feedback (evidence-based findings, conclusions and recommendations). Individuals in the organization make rational decisions on behalf of their organizations.

These rational behaviors consist in the use of evidence from evaluations to maximize the benefit for the organization. However, if this is the case, why evaluation recommendations do not reach a 100 per cent rate of implementation? Why evaluation findings are not always taken into consideration when making decisions?

What is wrong? Why are managers, decision makers and other stakeholders not rationally using hard-earned evaluative evidence, findings, conclusions and recommendations for the benefit of the organization?

The answer to these questions comes from a new branch of economics - behavioral economics (BE) developed by Nobel Prize economists Daniel Kahneman and Richard Thaler.

The main assumptions of BE is that our rationality is bounded. The reasoning involved in human judgment and decision-making is constrained and moderated by the cognitive limitations as well as the finite amount of time and information available when making decisions.

It is a well-documented fact that individuals use typical heuristics “rules-of-thumb” full of cognitive biases to make decisions whether professional or personal.

It is hard to recognize that human brains do not equal to perfect rational economic behavior. Managers and decision makers

accept findings and recommendations from evaluations but in the majority of occasions do not act rationally on them. Sometimes they use instead other heuristics and come back to what they already knew or were doing before the evaluation was conducted despite evidence showing the opposite.

While it is recognized across this guide that not all problems affecting evaluations arise from the failure of human behavior, it also acknowledges that individuals with different behaviors drive interventions and participate in evaluations.

The goal of this guide is to recognize, identify and address human bias when engaging in evaluation by using the knowledge produced by behavioral insights and “nudges” to increase the usefulness and impact of evaluations.

More specifically this guide aims at:

- (a) Assisting evaluators, evaluation commissioners, managers, and implementers of interventions in making evaluations more effective and impactful using behavioral insights (BI) tools;
- (b) Supporting decision-makers and implementers in enhancing the effectiveness and impact of their interventions through **behavioral changes**;
- (c) Positively impacting evaluation outcomes using a behavioral approach; and
- (d) Defining the behavioral insights evaluations using the ENABLE WIPO Evaluation Approach which has been designed by the IOD Evaluation Section to be performed when conducting evaluations with BI.

WHOM IS THIS GUIDE FOR?

The World Intellectual Property Organization (WIPO) Evaluation Section has designed this Behavioral Insights guide and the ENABLE WIPO Evaluation Approach as a practical resource for a wide range of stakeholders interested in integrating BI to their evaluation or aiming to strengthen the impact of evaluations by identifying strategies to address behavioral challenges affecting evaluations. This guide could be used in particular by:

- (a) Evaluation commissioners, practitioners, and evaluation consultants;
- (b) Interventions' stakeholders, including staff involved in an evaluation process;
- (c) Evaluation users interested in using BI; and

- (d) Behavioral insights evaluations to improve interventions.

This guide is structured in three parts with its corresponding annexes. The first part is of introductory nature and describes key notions of the evaluation in the institutional context. It also justifies why BI is important to improve evaluation.

The second part of the guide covers key considerations to produce evaluations that integrate gender and human rights.

The third part of the guide is devoted to explaining how to integrate BI through the evaluation process in a step-by-step fashion.

Finally, the guide is completed with annexes and useful templates supplementing and operationalizing the guidance contained in the document.

Evaluation's role in the context of WIPO



Roles and responsibilities for an effective evaluation function

An effective evaluation function requires the cooperation and commitment of many actors as evaluation, in this sense, is a shared responsibility distributed across several types of stakeholders.

Organizations should ensure that:

- (a) Independent evaluations are conducted by a central evaluation function;
- (b) An appropriate percentage of the evaluation funds is reserved and used efficiently for evaluations;
- (c) The evaluation function performs independently from other management functions to ensure impartial reporting;
- (d) There is a clear differentiation between other disciplines and evaluation;
- (e) Key stakeholders understand the benefits from having their field of work evaluated and plans of action developed accordingly to tackle the improvement opportunities identified;
- (f) Evaluation function staff and externally contracted evaluation consultants are protected against undue influence to enable them to express their judgements and opinions objectively and in an impartial manner;
- (g) Evaluations are conducted by staff who have the relevant educational background, qualification, and training, as well as professional work experience;
- (h) Evaluation policies, procedures, methodologies, and guidelines are in line with the United Nations Evaluation Group (UNEG) Standards; they are applicable to the whole organization and are being updated regularly;

- (i) Terms of Reference for evaluations are developed for each evaluation, with the support of key stakeholders;
- (j) Before finalization of evaluation reports, evaluation findings, conclusions, and recommendations are discussed with Sectors leads;
- (k) Independent evaluation reports are sent to high-level decision-makers;
- (l) The outcomes of evaluation reports are included in the internal oversight reports;
- (m) User-friendly mechanisms for the publication and dissemination of findings and lessons from evaluations are applied;
- (n) BIE communication plan is developed for each evaluation;
- (o) Adequate support is given to the intervention's decision-makers for planning and conducting decentralized evaluations; and
- (p) Evaluation staff contributes to the organization's results-based management processes, as required.

An example of how to draft formal procedures for evaluation functions can be found in the Evaluation Policy and the Evaluation Manual of the WIPO Internal Oversight Division (IOD).

Sectors' roles and responsibilities in decentralized evaluations

All Sectors and Sector leads should ensure that:

- (a) The Evaluation Policy, procedures, methodologies, and guidelines are adhered to and applied;

- (b) Staff managing evaluations and externally contracted evaluation consultants are provided with necessary information to conduct evaluations and interviews in a comprehensive, objective, and impartial manner;
- (c) The evaluability of initiatives is enhanced and systematically evaluated; and
- (d) Adequate evaluation knowledge exists among their staff; and evaluation results are effectively used and appropriately shared within their Sectors and business units.

WHAT IS EVALUATION?

An evaluation is a systematic, objective, and impartial assessment of an on-going or completed project, program, or policy, its design, implementation, and results. The aim is to determine the relevance and fulfillment of objectives, their efficiency, effectiveness, impact, and sustainability. An evaluation should contribute to learning and accountability and provide credible, evidence-based information, enabling incorporating findings and recommendations into an organization's decision-making processesⁱ.

WHY ARE BEHAVIORAL INSIGHTS IMPORTANT FOR EVALUATIONS?

Interventions designed to promote or make use of Intellectual Property (IP) require stakeholders' and decision-makers commitment to succeed.

Examples of IP interventions could include a wide range of activities such as an examiner of patents participating in IP training to become more efficient in the processing of filings; or a WIPO staff member working with governments to identify strategies for investing resources in research and development to boost economic development.

These interventions require different stakeholders to change their current behaviors to certain extent. If the intervention is not successful enough in providing necessary information, incentives, or means to encourage the desired behavioral change, the intervention is unlikely to succeed fully.ⁱⁱ

More than 202 institutions¹ globally have applied BI to public policy drawing from behavioral science's good practices, approaches, tools, and methods to create the space for change and influence better decision-making.

In today's world, evaluation practitioners will benefit from factoring in BI in the evaluations by creating choice environments to help better achieve intervention and evaluation goals. "They do not define goals, but they support stakeholders in getting there." by going beyond incentives, law, and policy regulations that do not always work."ⁱⁱⁱ

Behavioral insights evaluations assist organizations by informing about interventions advantage, human behavior and decision-making and provides diagnostic evidence on specific behavioral barriers preventing interventions from succeeding.

Behavioral insights can be entirely instrumental for evaluation if integrated early on in the evaluation process. This guide factors BI concepts, tools, and approaches across the evaluation process. The specific objectives of this guide are to:

- (a) Design and apply strategies addressing behavioral challenges during the evaluation process to enhance evaluation results; and
- (b) Identify behavioral challenges affecting interventions, and design strategies aimed to reduce their

¹ OECD. "[Behavioural insights](#)". *OECD Better Policies for Better Lives*. Organisation for Economic Co-operation and Development. Retrieved 2 November 2020.

adverse effects and ultimately contribute to strengthening the impact of interventions.

EVALUATION CRITERIA

Table 1: Evaluation Criteria^{iv}

Evaluation Criteria	Definition
Relevance	The extent to which the intervention objectives and design respond to beneficiaries ^v , global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change.
Coherence	The compatibility of the intervention with other interventions in a country, sector, or institution.
Efficiency	The extent to which the intervention delivers, or is likely to produce, results in an economical and timely way.
Effectiveness	The extent to which intended expected results are being achieved or are expected to be achieved.
Impact	The lasting changes – positive and negative, intended and unintended higher-level effects – arising from the intervention.
Sustainability	The degree to which the benefits of the intervention continue or are likely to continue after completion.
Human Rights and Gender	The extent to which the intervention follows a human rights-based approach, addresses issues such as equality and non-discrimination, accountability, and social transformation.
Gender	The extent to which the intervention is based on a gender-responsive approach, focuses on equal participation, inclusion, and empowerment of men, women, marginalized and vulnerable groups.
Additional criteria that may be used in evaluations	
Design	The extent to which appropriate and participatory planning processes took place; the existence and suitability of logical frameworks and performance indicators.
Partnership and Cooperation	The extent to which effective partnerships were established and maintained; the extent of alignment and contribution to the One UN, UNDAF, and other coordination mechanisms.
Innovation	The extent to which innovative approaches were successfully used or emerged from implementation.
Behavior	The extent to which stakeholders' behavior (beliefs, attitudes, capacities, knowledge, values, etc.) affected positively or negatively the intervention

WHAT ARE THE KEY SUCCESS FACTORS FOR IMPLEMENTING A BEHAVIORAL INSIGHTS APPROACH IN EVALUATIONS?

As a result, key stakeholders identify and discuss barriers and potential solutions to improve performance. Four key factors contribute to the success of an evaluation using behavioral insights, including:

Evaluations are intended to increase commitment, learning, and accountability.



CONSULTATION AND INCLUSIVENESS

- Utilization-focus evaluation where evaluation is linked to an event or decision that needs to be taken
- Establish stakeholders' support for the evaluation
- Stakeholders engagement in the evaluation design, conduct, and formulation of recommendations
- Potential users participation throughout the evaluation process
- Management is engaged in the feedback process and commits to taking action
- Involve the client in the behavioral reduction process
- Relevant stakeholders have been invited to identify and discuss barriers



REPORTING

- Behavioral optimized report - structure, and content
- Customized versions of the receivers of the report
- Clearly define the use of the report.



FRAMING

- Reach out to relevant stakeholders and test possible solutions with them to receive feedback and commitment
- clear link between findings, conclusions, and recommendations
- Recommendations are SMART (Specific, Measurable, Achievable, Realistic, and Time-bound)
- Actionable- recommendations make sense in the context of the program/project
- Going beyond the follow up on recommendations and engaging with the client at the beginning of a changing process



MONITORING

- The existence of a logical or results framework with clear objectives, baselines, and benchmarks
- The existence of baseline information, counterfactuals, targets, etc
- The existence of systematically reported monitoring data

WHY ARE EVALUATION AND BEHAVIORAL INSIGHTS IMPORTANT?

Ultimately, the information obtained through evaluation serves the following purposes: (i) organizational learning; (ii) accountability; (iii) knowledge generation; and (iv) opportunities for dialogue. However, organizational learning, knowledge generation, and dialogue will be more effective when stakeholders are

committed to the evaluation process. Stakeholders' commitment is essential for adhering to and implementing recommendations resulting from evaluations and ultimately producing long-term, sustainable behavior change. Behavioral insights are a valuable mechanism for evaluators as they reinforce the utilization of evaluation results, organizational learning, and knowledge creation.

ORGANIZATIONAL LEARNING

Organizational learning occurs when staff creates, acquires, transfers knowledge, and modifies its behavior to reflect new knowledge and insights.

By answering difficult questions, such as whether the organization is doing the right things and doing things right, evaluation provides an opportunity to learn from past experiences and best practices.

It supports the organization in the learning process by integrating reflection opportunities for improving and incorporating new learning in daily operations.



KNOWLEDGE GENERATION

By producing substantive knowledge about the specific topics that are part of an organization's mandate and innovative practices.

This knowledge is generally found in the recommendations and lessons learned contained in evaluation reports. Such information is compiled from multiple evaluations and then synthesized and shared



OPPORTUNITIES FOR DIALOGUE

By providing a useful platform for stakeholders to discuss the evaluation and other areas of common interest.

Inclusive evaluation processes help build relationships and better understand participants' and other stakeholders' needs and interests, and opportunities for further collaboration.



ACCOUNTABILITY

By assessing compliance with established conventions, treaties, norms, policies, and plans.

Accountability is achieved through independently conducted evaluations that accurately report on performance results to WIPO, Member States and other relevant stakeholders;



Evaluations done with the commitment of stakeholders can create the space for participants to reflect on an intervention. It may inspire new ideas that lead to change,

such as new partnerships or new initiatives. Thus, evaluation as an agent of change is often an intervention in itself.

WHAT ARE THE TYPES AND DIMENSIONS OF EVALUATION?

A range of different types of evaluations exists, providing an assessment of



interventions, organizations, and policies' performance from several different perspectives. Balancing the evaluation types with the needs of the organization is an important planning consideration.

Table 2: Types of Evaluation

Type of evaluation	Definition
Strategic evaluations	Strategic evaluations provide knowledge on the strategic relevance of programs and their contribution to achieving their strategic goals. They can also look at the organization's contribution to system-wide goals, such as the UN's Sustainable Development Goals.
Thematic evaluations	Thematic evaluations analyze the organization's experience and practice in addressing crosscutting issues such as knowledge management, capacity building, or gender equality and derive conclusions and lessons for the whole organization.
Country evaluations	These evaluations assess the performance and impact of an organization's assistance to a country and aim to contribute to optimizing the portfolio of its activities.
Program evaluations	Program evaluations assess programmatic areas of sectors or departments (as defined and described in WIPO's Program and Budget documents). They also assess the relevance and sustainability of results as contributions to medium-term and longer-term goals.
Project evaluations	This involves evaluating one or more activities designed to achieve specific objectives within specified resources and periods; the project could be part of a broader program. These evaluations are undertaken throughout the implementation cycle and improve project impact and performance by contributing to learning.
Organizational assessments	These are aimed at understanding and improving performance. While four main criteria: Effectiveness, Efficiency, Financial Sustainability, and Relevance, are often used, the purpose is to go further and use the evaluation as a diagnostic tool for organizations implementing an internal change or strategic planning process, or both. Organizational assessment goes beyond measuring the results of an organization's programs, products, and services. ^{vi}

HOW ARE EVALUATION AND BEHAVIORAL SCIENCE DIFFERENT?

In addition to evaluation, other assessment practices and disciplines help ensure decision-makers and stakeholders have sufficient understanding of interventions. This section distinguishes evaluation from other disciplines.

DISCIPLINE	DEFINITION	DIFFERENCES
	<p>BEHAVIORAL SCIENCES</p> <p>Studies the way that emotions, the environment, and social factors influence decision making. It is interested in how heuristics, biases, and framing can lead people to make irrational decisions. It partly applies methodologies developed in social sciences, such as randomized trials.</p>	<p>Evaluation focuses more on processes: Are we doing the right things? and are we doing things right?</p> <p>Behavioral scientist is people center and aims to understand human actions and decision making: why people do things, not just observe whether they did something right or not.</p>
	<p>MONITORING</p> <p>Continuous, systematic collection and analysis of data based on the intervention's indicators. Helps determine progress made towards achieving results, and using allocated resources. Keeps interventions on track and ensures timely decision-making to improve implementation.</p>	<p>Monitoring data is used in evaluation processes to understand what is happening. M&E and planning are the foundation for RBM. Evaluation is conducted at specific points in time and uses multiple sources and types of data. It is usually conducted by independent evaluators or external evaluation. It responds to whether we are doing things right, doing the right things, and why.</p>
	<p>REVIEW</p> <p>Usually internal and periodic or ad hoc assessment that typically addresses performance and operational implementation issues. Examples of reviews include rapid assessments and evaluability assessments (often conducted prior to an evaluation).</p>	<p>Evaluations are usually external and more rigorous, and broader in scope than reviews.</p>
	<p>RESEARCH</p> <p>Systematic examination undertaken to develop or contribute to the knowledge of a particular topic. Examples include in-depth baseline studies and impact studies.</p>	<p>Research often feeds information into evaluations, or decision-making processes.</p>
	<p>INSPECTION</p> <p>General examination of an organizational unit, issue or practice. It determines the extent to which the unit, issue or practice adheres to prescribed standards, good practices or other criteria. Information gained is used to make recommendations for improvement or apply corrective measures.</p>	<p>While evaluations are performed as part of the project management cycle, inspections are often performed when there is a perceived risk of non-compliance.</p>

DISCIPLINE

DEFINITION

DIFFERENCES

AUDIT

Independent, objective assurance and consulting activity designed to add value and improve organization's operations. Systematic, disciplined approach to evaluate and improve the effectiveness of risk management, control, and governance processes. Focus on certifying the true and fair view of financial statements, compliance, and performance or value for money.

While both Evaluations and Audits assess performance, evaluation places more emphasis on the needs and priorities of stakeholders, results, impact, and lessons learnt.



Evaluation is a management tool that can be used for learning and accountability. The main objective is to help decision-makers improve interventions to strengthen their impact. Evaluation should not be used to overrule political decisions, address human resources issues, wrongdoing, certify financial statements, or other strictly compliance-related exercises. Evaluators are geared to contribute to a learning culture, which can only happen by building trust and commitment from stakeholders. Evaluation for purposes other than learning and accountability would be unethical and against the UNEG's ethical evaluation principles.

Institutions and units undertaking evaluation across the UN

Most UN Organizations and their specialized agencies have specific evaluation functions that provide expertise, guidance, and management of their evaluation activities. These units undertake evaluations responding to the need to account for the use of resources, demonstrate results, and the added value to the organizations' work.

In addition to the specialized evaluation functions located within each individual organization, there are other oversight bodies undertaking evaluations across the UN, namely:

FRAMEWORK



JOINT INSPECTION UNIT

Established in 1976 by the General Assembly of the UN. The JIU is an independent external oversight body of the UN System mandated to conduct evaluations, inspections and investigations system-wide. Its objective is to enhance the efficiency of the administrative and financial functioning of the United Nations System and, to this end, it may make on-the-spot inquiries and investigations.



OFFICE OF INTERNAL OVERSIGHT SERVICES

Established in 1994 by the General Assembly. The office assists the Secretary General in fulfilling his oversight responsibilities through the provision of audit, investigation, inspection, and evaluation services but limited to the UNOs. Its mandate does not extend to any of the specialized UN agencies. It promotes responsible administration of resources, accountability, transparency, and improved Program performance.



ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The conduct of evaluation worldwide is guided by the normative work and standards of OECD. The work of the OECD has been instrumental in providing the key evaluation criteria which are the foundation of evaluation. The five criteria defined by the OECD include: relevance, effectiveness, efficiency, impact and sustainability. In addition it has defined the evaluation glossary used by evaluation units worldwide.



UNITED NATIONS EVALUATION GROUP

In 2005, UNEG released its first Norms and Standards which set out clear principles to strengthen and harmonize evaluation practice for development interventions. UNEG updated its Norms and Standards in 2016. This updated version introduces four new norms - human rights and gender equality, national evaluation capacities, and professionalization of evaluation.

Table 3: List of UNEG 2016 General Norms.

Norm	Description
Internationally agreed principles, goals and targets	Upholding and promoting UN principles and values is the responsibility of evaluation decision makers and evaluators. This includes respecting, promoting and contributing to the goals and targets set out in the 2030 Agenda for Sustainable Development.
Utility	There must be clear intention to use the evaluation's analysis, conclusions, and recommendations. This includes relevant and timely contributions to Organizational learning, decision-making, and accountability.
Credibility	This requires independence, impartiality, rigorous methodology, and ethical conduct. Key elements include transparent processes, inclusive approaches involving relevant stakeholders and robust quality assurance systems.
Independence	Evaluators must have the freedom to conduct their work without influence from any party or threat to their careers. In addition, the organization's evaluation function must be positioned separately from other management functions, be responsible for setting the evaluation agenda, and have adequate resources to do its work.
Impartiality	This entails objectivity, professional integrity, and absence of bias at all stages of the evaluation process. Evaluators must not be or expect to be directly responsible for the policy setting, design or management of the evaluation subject.
Ethics	Evaluations need to be conducted with the highest standards of integrity and respect for the social and cultural environment, for human rights and gender equality, and for the "does no harm" principle.
Transparency	This is essential to establish and build trust, confidence, stakeholder ownership, and public accountability. It includes making evaluation products publicly available.
Human rights and gender equality	These universally recognized values and principles need to be integrated into all stages of an evaluation, underpinning the commitment to the principle of "no-one left behind".
National evaluation capacities	Building capacity for evaluation development activities at the country level is vital and is to be supported when requested by Member States.
Professionalism	To ensure credibility, evaluations need to be conducted with professionalism and integrity.



GENDER AND HUMAN RIGHTS

2. HUMAN RIGHTS, GENDER EQUALITY AND BEHAVIORAL INSIGHTS EVALUATIONS

The promotion and protection of human rights (HR) and gender equality (GE) guide all UN entities. There is virtually no aspect of the UN's work that does not have an HR dimension. Intellectual property rights are enshrined as human rights in the Universal Declaration of Human Rights. Article 27 of the Universal Declaration states:

- (a) Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts, and to share in scientific advancement and its benefits; and
- (b) Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary, or artistic production of which he is the author.^{vii}

Whether we are talking about peace and security, development, humanitarian action, or climate change, none of these societal issues can be addressed in isolation from HR. Moreover, in the UN context, evaluation is closely connected to the Sustainable Development Agenda and GE. Evaluations that do not follow HR, Sustainable Development and GE principles risk reinforcing or incurring in harmful patterns. This includes negative ratings of evaluation reports in the UN System-wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP)^{viii} reporting. The United Nations Evaluation Group has been instrumental in providing guidance on how evaluations can address the principles of HR and GE. It helps clarify and provide context for the key terminology used in such discussions.

Gender-responsive evaluations provide "credible and reliable evidence-based information about an intervention progress (or the lack thereof) towards intended and/or unintended results regarding gender equality and the empowerment of women"^{ix}.

Gender equality serves to the advantage of both men and women, girls and boys, and all individuals/groups marginalized and/or discriminated based on their gender. Equality cannot be achieved without the full engagement of all of them. Furthermore, men and women are subject to different, often contextually specific forms of bias and discrimination (e.g., gender identity, class, religion, caste, ethnicity, age, disability, sexual orientation, location, among others). Thus, gender-responsive evaluations should be sensitive to and assess all forms of discrimination that women and men face. Gender-responsive evaluation should determine the extent to which some of these challenges are due to behavioral aspects and suggest recommendations to address normative and behavioral difficulties.

HUMAN RIGHTS AND GENDER EQUALITY CONSIDERATIONS IN EVALUATION

The main concepts underlying human rights and gender-responsive evaluations are **inclusion, participation, non-discrimination, and fair power relations**. These concepts help improve programming by considering important social and cultural issues that can make interventions more effective and sustainable.

The United Nations Evaluation Group has developed two sets of guidance documents that are useful resources for evaluators and those who manage evaluation processes. The UNEG's 2011 publication "*Integrating Human Rights and Gender Equality in evaluation – towards UNEG guidance*" is an abridged version that provides step-by-step advice for preparing, conducting, and using HR and GE evaluations. Its 2014 publication, "*Integrating Human Rights and Gender Equality in Evaluations*", provides more in-depth theoretical and practical information, tools, and suggestions. While these guidance documents include some behavior-related questions, evaluation practitioners and policymakers need a more detailed guidance to be able to assess behavioral aspects. This document provides specific tools and approaches to

undertake a full-fledged assessment of HR and GE when conducting evaluations.

In 2012, the UN-SWAP set common performance standards for the gender-related work of all UN entities. Starting from 2013, all UN entities must self-assess and report their progress in meeting 15 Gender Equality and the Empowerment of Women performance indicators annually. The indicators track six components of gender mainstreaming: accountability, results-based management, oversight (including evaluation), human and financial resources, capacity, and knowledge exchange and networking.

The UN-SWAP indicator dedicated to evaluation processes is linked to meeting the gender-related UNEG norms & standards and demonstrating effective use of the UNEG guidance on integrating human rights and gender equality in evaluation.

The UNEG working group on GE and HR developed a technical note and scorecard for the Evaluation Performance Indicator. It aims to support more systematic and harmonized reporting using a common tool that allows for improved comparability across the UN system.

To ensure a continuous improvement in mainstreaming gender and HR perspectives into the evaluation cycle, the Evaluation Section engages in the following practices:

- (a) Factoring BI tools and approaches in evaluations facilitating the identification of behavioral barriers, friction factors and opportunities to develop solutions in order to reduce their negative effects;
- (b) Raising awareness of internal and external stakeholders about GE and HR. This includes ensuring all evaluation teams receive relevant guidance as part of their key reading material;
- (c) Ensuring, to the extent possible, that there is an equal representation of both genders in all evaluation teams;

(d) When possible, including HR and gender experts in evaluation teams; and

(e) Having members from the evaluation office actively participate in and contribute to UNEG's work on HR and GE.



EVALUATION THROUGH BEHAVIORAL INSIGHTS LENSES

3. THE EVALUATION PROCESS INTEGRATING BEHAVIORAL INSIGHTS

This part of the guide describes the key steps in conducting an evaluation integrating BI.

This Behavioral Insights Evaluation guide using the ENABLE WIPO Evaluation Approach has been divided into three

phases described below, along with the objective of each phase. Phase one - Commitment (steps 1 to 2); Phase two - Diagnostic (steps 3 and 4); and Phase three - Impact (steps 5 and 6). Each phase and stage has been designed to address specific behavioral challenges evaluators face during the evaluation process. The ENABLE WIPO Evaluation Approach which factors behavioral insights strategies has been designed to address those challenges.



PHASE 1
COMMITMENT

PHASE 2
DIAGNOSTIC

PHASE 3
IMPACT

An extensive list of additional resources and tools is provided in the appendices.

A quality checklist may be used to ensure the completeness of an evaluation is also included in the annexes. The checklist includes additional administrative procedures that are specific to Organization and/or the IOD Evaluation Section.



PHASE 1 COMMITMENT

PHASE 1: COMMITMENT

When working towards long-term goals such as the ones included in programs and evaluations, individuals often have difficulty staying motivated without guidance, support, or feedback.

It is hard to stick to some of the goals set for various internal and external reasons: willpower, self-regulation and self-control or external choice architectures that distract our attention.

Determination challenges can also create mental exhaustion, which decreases decision-making abilities leading to self-directed blame.

People not engaging and not acting according to their intentions (so-called

intention-action gap) share behavioral problems related to commitment issues. Supporting the evaluation process might not be a significant challenge, yet maintaining that support over time requires repeated mobilization of motivation in the face of issues posed by competing goals and priorities.

Factoring BIs at this early stage helps evaluation practitioners be better aware of commitment biases by focusing on designing pre-commitment devices and long-term commitment behavior strategies.

Some of the guiding questions to identify commitment biases include:



What are the **points of friction** vis a vis the desired behavior?



Is it too easy to **overlook** important information or miss key recommendations in the evaluation reports?



Are key stakeholders, who read the reports, given enough time to provide relevant **inputs to the reports**?



- How do recommendations affect (+/-) the context and key stakeholders' groups?
- Do people **commit** to recommendations resulting from evaluations privately or publicly?

STEP 1- PREPARATION: SOCIAL EXPECTATIONS



Identification of evaluation topics

The Evaluation Section establishes the annual evaluation plan considering a number of factors including relevance and feedback received from WIPO Management, Member States, and available resources.

This section of the guide addresses common behavioral issues that affect the evaluation process at its very beginning and identifies strategies to tackle those issues.

Evaluations are heavily rooted on learning; and learning is more effective if there is commitment. The evaluation ENABLE

approach with BIs complements the traditional identification of evaluation areas with the integration of social expectations aiming to enhance the commitment of stakeholders in the evaluation process as early as possible. It also focuses on addressing behavioral issues such as people being reluctant to change and sticking to the status quo. Evaluators can help stakeholders facilitate the decision-making process and design strategies to frame the evaluation prospects while considering stakeholders' motives and perspectives. Evaluators can thus clarify during the evaluation design stage how the clients and the intervention will benefit from the evaluation design.

Evaluability Criteria

Usefulness
Can the evaluation have a significant effect on the subject/program being evaluated?

Design
The extent to which the intervention logic has been correctly defined and formalized, e.g., in a specific performance framework (e.g., Logical Framework for projects).

Timeliness
Does the timing fit with the planning cycle or upcoming decision-making? Are there any impediments or priorities conflicting with the planned evaluation, e.g., significant activities or absences of key stakeholders?

Coverage
Would the evaluation contribute to a balanced work schedule? Has the area been under/over-evaluated (or audited)?

Resources
Are sufficient resources and means available to evaluate within the planned time frame? Are all relevant data available?

Behavioral Insights

Reframing the evaluation team and purpose
Creating a clear differentiation between evaluation and other oversight functions by asking the questions:

- How are evaluations different from audits? and How will the client benefit from having his/her field of work evaluated?

Commitment
Identify key stakeholders and establish their support for the evaluation and possible future recommendations.

Additional Selection criteria
Ensure that the evaluation selection criteria includes authority, budget, and motivation for the area under consideration.

STEP 2 – LAUNCH: MAKE IT EASY AND SOCIAL



Make it social and build commitment

Major stakeholders in evaluations are agencies, organizations, groups, or individuals who directly or indirectly are interested in the implementation of the intervention (program/policy) or its evaluation.^x Nevertheless, stakeholders' willpower is limited and subject to different psychological biases.

Even though managers and decision-makers adhere to evaluation recommendations, it is frequently difficult to implement them. The traditional evaluation approach assumes that stakeholders who

have agreed on the recommendation and decided to pursue specific long-term goals (recommendations) will stick to them. However, it is not easy to stick to long-term goals in a busy work environment.

Stakeholders not engaging and not acting on their intentions (the so-called intention-action gap) share behavioral problems related to commitment issues.

The BIs approach takes into consideration this determination challenge, which can decrease the decision-making abilities of managers and other decision-makers. Taking the right decision is not enough.

Managers and stakeholders can have difficulty sticking to some of their choices over time because of willpower, self-regulation, self-control, and choice architectures that guide them away from their decisions. This leads to problems with determination involving psychological discomfort of not being able to sustain their commitments or, in the case of evaluation, implement recommendations.

Behavioral insights identify potential commitment bias and design strategies to enhance stakeholders' pre-commitment devices and long-term commitment to the evaluation process. Some strategies aim to respond to questions, which should be discussed and defined as early as possible in the evaluation process and included in the evaluation terms of references (ToRs).

[Invite stakeholders to participate in the evaluation process](#)

When selecting stakeholder groups, consider their level of influence and commitment. The figure below could be used as a tool for classifying stakeholders.



Once stakeholders have been identified and categorized, evaluation managers need to establish a safe alliance with the evaluation client and key stakeholders. At this initial stage, it is of utmost importance to soften any intimidating atmosphere. This first communication could consist of an initial e-mail or if necessary, a memorandum. In this case, evaluation managers should consider the language and formality used in this first message that could create an intimidating atmosphere. This might be softened by including less formal language in

memorandum and keeping formal communication strictly to the minimum. The initial communication should be followed swiftly by a conversation with the implementing unit to inform them about the evaluation purpose, process and, more importantly, to listen to their needs, expectations, and concerns. Understanding the clients' expectations helps evaluators identify strategies to manage those expectations and reduce any uncertainties by letting the client know what is happening next.



Inclusion

How will key stakeholder groups be involved in the process?



Contribution

How are they expected to contribute to the evaluation?



Expectations





what they expect from the evaluation to enhance its value?

Strategies to build commitment

There are many different approaches to involve stakeholders in the evaluation process. Evaluators choose them depending on particularities of each evaluation. Evaluators engage stakeholders through one-to-one consultations, group consultations, workshops, world cafes, among others. All these communication channels are used to gather stakeholders' views and increase participation. Still, they might not necessarily be designed as pre-commitment devices for long-term commitments. Whatever channels

evaluators choose to engage their stakeholders, evaluation managers should be mindful of the group's needs being targeted. Depending on the target group and their needs, it might be helpful to draw to some determination strategies (listed below) to overcome the intention-action gap and enhance stakeholders' commitment.

If a decision is made and a reference group is put in place, a specific ToR for the reference group should be developed. A ToR for Reference Groups Template is included in the Annexes.

DETERMINATION STRATEGY	DEFINITION	EXAMPLE
	<p>SOCIAL NORM</p>	<p>Social norms can be instrumental in encouraging other behaviors if appropriately used. Social norms can be used to enhance the commitment of stakeholders. It can also serve as an evaluation recommendation when addressing behavioral challenges.</p>
	<p>CREATE A COMMITMENT</p>	<p>Make use of strategies that will help evaluation stakeholders follow their plan of action.</p>
	<p>PLANS AND FEEDBACK</p>	<p>Maintain the stakeholders' commitment during and after the evaluation by making use of action plans.</p>
	<p>WORK WITH FRICTION</p>	<p>Keep it simple and easier to take up. The simpler it is to perform an action, the more likely it is to work towards the goals or implement the recommendations.</p>

Source: OECD (2019). Tools and Ethics for Applied Behavioral Insights: The Basic Toolkit. OECD Publishing. Paris. <https://doi.org/10.1787/9ea76a8f-en>.

Frame the prospects

At times, the evaluation function is subsumed within other organizational functions, an oversight division or a program and management unit. In such a case, it is of utmost importance that the evaluation function clearly presents its unique added value by clarifying the difference between evaluation and other functions to the client. For instance, if the evaluation function is within



an oversight division, it might need to clarify the difference between evaluation, investigation and audit to the client to avoid confusion.

If stakeholders are not clear about the distinct role and purpose of evaluation, it could create uncertainty and a feeling of unease among certain stakeholders. Clarifying the evaluation's purpose and expectations helps frame its scope and design and is more likely to lead to a good collaboration. It is crucial to involve and consult relevant stakeholders in this process, especially to continue building on the commitment across the whole evaluation process.

The purpose of an evaluation with BIs should ideally be tied to a future event or decision or to a specific organizational need or learning opportunity. Examples of five developmental (learning based) evaluation purposes^{xi} include:

- (a) Ongoing development in adapting an intervention to new condition in complex dynamic systems and working on the debiasing stakeholders assumptions whenever possible;
- (b) Adapting effective principles and strategies to a new context as innovation and ideas are taken from elsewhere and developed within a new setting. This is useful, especially when stakeholders' disagree about the process and priorities;
- (c) Developing a rapid response in the middle of a significant change or crisis, such as the Coronavirus Disease 2019 pandemic. The purpose of such evaluation would be to explore real-time solutions and generate innovation and helpful BIs strategies to address existing challenges;
- (d) Preformative development of potentially scalable innovation helps emerge ideas and visions exploration that could be transformed into an intervention; and

- (e) Significant systems change and cross-scale developmental evaluation provide feedback on how significant system change is unfolding and being adapted.

Additionally, it is important to form a clear understanding of the scope of intervention, which entails a clarification of the relevant time and aspects to be included or excluded. Some interventions have short, defined periods; others run over extended periods but can be broken down into distinct phases, which helps determine the scope.

Approach note

Once the key evaluation design elements have been defined, it should be possible to prepare a brief behavioral optimized approach note describing the evaluation process. The behavioral optimized approach note would include: the topic for evaluation, the added value of the evaluation, differentiation between evaluation and other functions, stakeholders' expectations, their roles, the purpose, an outline of the proposed design (if known), the anticipated timing, and resources needed. The behavioral optimized approach note is not obligatory but it is a useful tool for the next phase of the evaluation. It can easily be transformed in ToR, as explained in the next section.



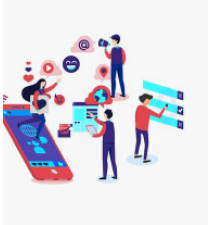


PHASE 2 DIAGNOSTIC

PHASE 2: DIAGNOSTIC

According to a standard theory of rational decision-making, individuals form their beliefs according to the rules of logic and probability. However, in reality, people rely on shortcuts (rules of thumb) and often overestimate or underestimate the outcomes and possibilities when making decisions. The so-called confirmation bias.

Individuals need to rely on a coherent worldview to make predictions and decisions. In doing so, they ignore relevant information that is against their views or in other cases, only accept information that confirms those beliefs.

The consequence can be an overestimation or underestimation of outcomes, missing relevant information, and relying too much on heuristics to make decisions, i.e., mental shortcuts or intuitive judgments, for instance, when we believe that an intervention is not successful according to our beliefs despite evidence showing the opposite or vice versa. Bias can affect the evaluation outcomes and the effectiveness of recommendations. Therefore, it is crucial to work on the implementation of debiasing strategies for each bias identify. This can be done as follows:

BIAS	DEFINITION	DEBIASING TECHNIQUES
 <p>AVAILABILITY BIAS</p>	<p>Information that appears more frequently is usually recalled by stakeholders better than less frequently occurring information (Kahneman, 2011:418)</p>	<ul style="list-style-type: none"> • Question the source of information • Ask for additional examples • Use checklist <p>(Kahneman, Lovallo & Sibony, 2011)</p>
 <p>CONFIRMATION BIAS</p>	<p>Stakeholders might search or interpret evidence in support of an opinion or pre-existing beliefs.</p>	<ul style="list-style-type: none"> • Consider the opposite • Reduce ambiguity (McKenzie, 2004:207) • Take an outside view (Koller & Lovallo, 2018). • Ask for additional examples (Kahneman, Lovallo & Sibony, 2011).
 <p>ANCHORING BIAS</p>	<p>This bias emerges when stakeholders engage in narrow thinking while problem solving by focusing too much on a starting point.</p>	<ul style="list-style-type: none"> • Question the motivation for using the anchor (Fasolo, n.d.= • Re-anchor (Kahneman, Lovallo & Sibony, 2011:56). • Consider the opposite (Larrick, 2004:323)

Source: The London School of Economics and Political Science: Decision-making, biases and nudges. London 2021

Some other guiding questions at this diagnostic stage could help reducing existing bias and are as follows:



Beliefs about the method

What are key stakeholders' pre-existing beliefs about evaluations and recommendations included in them?



Beliefs about the intervention

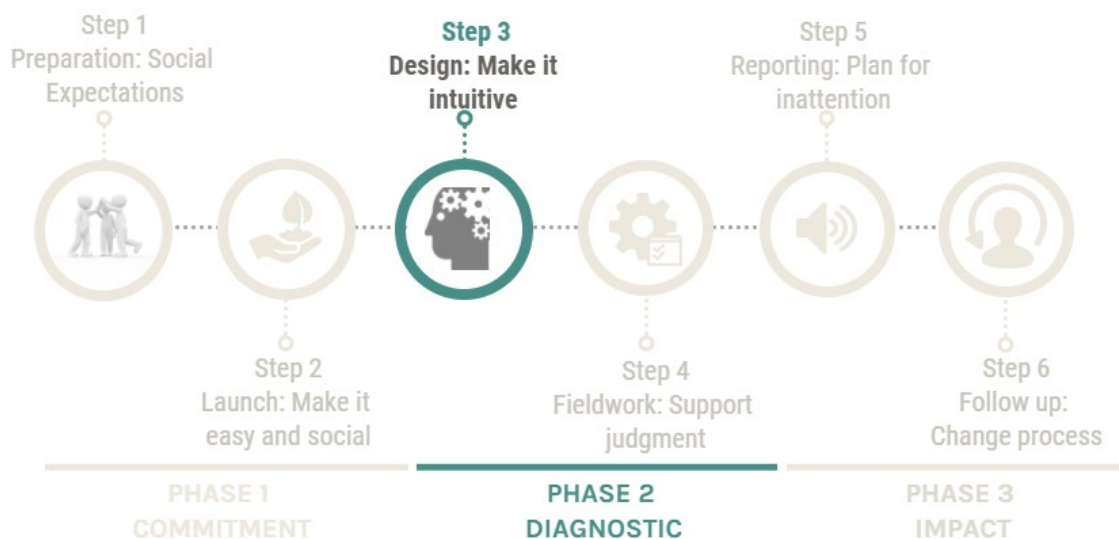
What are relevant stakeholders' attitudes towards the intervention to be evaluated?



Interaction

How do elements from the context in which relevant stakeholders are immersed interact with belief formation for the uptake of the evaluation report?

STEP 3 – DESIGN – MAKE IT INTUITIVE



Theory of change and mental models

Early in the planning process, evaluation managers need to have an in-depth understanding of the intervention (project, program, policy, or activity) to be evaluated. This will help to overcome difficulties with abstractions in decision making during the development of a Theory of Change.

The evaluation managers, in collaboration with the relevant stakeholders, will work on the following:

- (a) Summarizing the intervention logic or recreating a Theory of Change^{xii}, when possible;
- (b) Reviewing activities and outputs and comparing them to the theory of

change/logic model ensures that the model is a good guide for the evaluation;

(c) Making a behavioral reduction with the client facilitates collaboration. It develops a common understanding of the issues to be evaluated;

(d) Setting up a plan to do a behavioral mapping of the area. Use flowcharts to learn how things are working at the current moment.

(e) Identifying relevant stakeholders to interview or include in the process.

These activities help create a good understanding of the intervention. Moreover, they guide the planning, design, and diagnostic process.

Behavioral reduction and mapping

The earlier BIs are integrated into the evaluation process, the more effective the evaluation report will be. Behavioral reduction is a tool to identify behaviors affecting the intervention. Not all issues in an intervention will be of behavioral nature, so it is important that evaluation managers distinguish between behavioral issues and process related issues.

A behavioral reduction can help bridge the gap between the overarching intervention goal and the operational level activities. The main objective of this stage is to identify in collaboration with the client the behavioral challenges that could materialize in the future and link them with the respective behavioral informed solutions. This process also facilitates collaboration with the client and creates a shared understanding of the issues to be evaluated. The identified potential behavioral components will be triangulated during the evaluation process.

Behavioral mapping is a tool used to observe and record behaviors in a specific context.^{xiii} It reflects how people actually behave rather than how they should behave.^{xiv} At this stage, evaluation managers will consult relevant stakeholders about the

current behavior. Nevertheless, direct observations will also be helpful, especially considering that stakeholders might not feel comfortable providing their honest feedback.

The behavioral evaluation questions' matrix

The matrix shows the relationship between evaluation questions, sub-questions, indicators, techniques used in the evaluation, and critical sources of information (each of which is described below and in the next Chapter). The matrix allows the evaluators to ensure that the data collection tools and methods will meet the evaluation information needs and provide reliable and valid information for triangulation purposes.

The traditional evaluation process very seldom considers the identification of behavioral issues. Once a list of behaviors affecting the intervention has been identified, key behavioral questions can be included in the evaluation question matrix. This could be achieved by making stakeholders a part of the identification process and strengthening the evaluation findings. The questions below are a sample of questions that can be added to the evaluation question matrix (EQM):

Sample questions	Stakeholders' views from very likely to very unlikely
<p>Relevance Is behavior change an organizational priority?</p>	
<p>Effectiveness Do the specific behaviors observed affect the achievement of results?</p>	
<p>Efficiency Does the specific behavior affect the efficient delivery of the intervention?</p>	
<p>Impact Will changing the specific behavior enhance the intervention's longer-term impact?</p>	
<p>Feasibility Is it feasible to address this behavior? Are there any potential risks or unintended consequences when pursuing the desired behavior?</p>	

A Behavioral Evaluation Questions' Matrix template is included in the Annexes.

Formulating questions and sub-questions

An evaluation aims to respond to questions identified in the EQM. The matrix should explore the significant issues associated with the evaluation purpose by responding to the evaluation criteria defined in the introduction (relevance, efficiency, effectiveness, impact, and sustainability). They should be limited to the few most critical ones. It is recommended to have a maximum of five, i.e., one or two per relevant criterion. Keeping in mind that stakeholders' attention is limited, questions will be designed to identify lessons learned, good practices, and challenges that could be both of procedural and behavioral nature.

Evaluation and behavioral question indicators and measures

Question indicators help determine the data that needs to be collected to assess the intervention's progress and whether it is on track to achieve its goals and objectives.

Question indicators should provide a measurement to answer the question: What would we expect to see as verifiable evidence of the main accomplishment under each criterion?

When designing indicators, a few general concepts should be observed:

- (a) Limit the number of indicators to one or two per statement to avoid data overload;
- (b) Include at least one qualitative indicator per statement to provide greater depth to the analysis;
- (c) Ensure that indicators are gender-transformative, wherever possible. A gender-transformative indicator is a measure of change over time promoting gender equality and positive development outcomes. It indicates whether the actions taken are transforming unequal gender relations to promote shared power, control of resources, decision-making, and support for women's empowerment. For example: The

extent to which women and men's perspectives are considered in decision-making processes; and

- (d) Indicators should be SMART: Specific, Measurable, Attainable/Achievable, Relevant, and Time-bound.

Further resources on indicators are included in the Annexes.

Budgeting and resource planning

All the administrative procedures should be appropriately managed. This includes the overall cost frame for the evaluation, resource planning, and securing of funds.

Adequate planning requires that administrative procedures be adequately managed. As early as possible in the process, the evaluation commissioner should be considering:

- (a) The overall cost frame for the evaluation;
- (b) Resource planning; and
- (c) Securing funding.

Cost frame

The common cost categories for creating a program evaluation budget are:

- (a) Staffing – existing staff, experts or consultants, and external evaluators
- (b) Materials and supplies – general office supplies, including presentation materials, meeting rooms, telecommunication, etc.
- (c) Travel costs – particularly important when missions to countries will form a significant component of the evaluation (transportation and daily subsistence allowance).

Resource planning

Resource planning should be conducted as early as possible.

Securing funding

An evaluation should not proceed until the funding has been secured or "obligated" from the organization's budget. Every organization will have instructions on how to secure financing. An example of WIPO's Organization Procurement Guide has been provided. Organization procurement procedures must be followed for the engagement of any consultants or experts.

Additional resources also include what other programs are dedicating to evaluations. Programs need to reserve an appropriate amount in their respective budgets for monitoring and evaluation.

Terms of reference of the evaluation

The ToR document defines all aspects of the evaluation and is an essential step for managing a quality evaluation. An accurate and well-defined ToR sets the parameters against which the success of the assignment can be measured. At a minimum, the ToR should:

- (a) Briefly explain why the evaluation is taking place, what it will focus on and who will participate;
- (b) Describe the evaluation purpose, use, scope, and focus;
- (c) Identify the key questions that will be answered and any cross-cutting themes;
- (d) Explain the design and methodology selected;
- (e) Define the timing of the evaluation;
- (f) Provide background information on the intervention to be evaluated, along with a theory of change;
- (g) Outline the roles and responsibilities of the main participants in the evaluation; and

- (h) Make reference to any relevant norms, standards, or conventions (UNEG Norms and Standards)

ToRs can state the expertise required from internal and external evaluators. It does not usually contain the evaluation matrix or a detailed budget. However, the total budget for the evaluation may be included.

An example of ToR is included in the Annexes.

Using external specialists

When should external evaluators and behavioral experts be engaged?

Evaluation, behavioral or technical experts are generally engaged when there is a need for specific technical expertise. They may be hired to conduct the complete evaluation or a particular part of the evaluation or provide input and feedback on specific topics and contexts.

Their selection is critical to the success of the evaluation, and sufficient time should be allowed to ensure that qualified and experienced experts are identified. Irrespective of the scope of the role, the skill set required (for both individuals and teams) includes:

- (a) Strong management skills – to efficiently manage the evaluation;
- (b) Stakeholder engagement experience;
- (c) Methodological and conceptual expertise – to effectively undertake the necessary research, analysis, and synthesis of recommendations; and
- (d) Presentation and writing expertise – to present findings and recommendations effectively.

External Technical/Local Experts

External evaluators may be engaged to provide specific expertise, including:

(a) Subject matter expertise – examples include specific knowledge of trademark law or copyright and expertise in particular types of evaluations (e.g., impact or communications evaluations) or data collection/research techniques; and

(b) Local knowledge – understanding the politics, economy, culture, and traditions of a particular region may be necessary to collect the information required for the evaluation. Language expertise and access to local authorities' networks are always critical during in-country missions.

Recruitment/Contracting

Recruitment of consultants should be completed according to an organization's recruitment policies, procedures, and the Organization Procurement Guide.

Specific *ToRs for External Consultants* help define the experience and qualifications required and vet expressions of interest (Eoi) or tender documents. All proposals received should be checked for their compliance to the requirements defined in the specific ToR based on the following information:

- (a) A generic description of how the candidate(s) correspond(s) to the profile required for the type of mandate;
- (b) Availability of the experience and expertise relevant to the specific evaluation (individual and collective in case of teams);
- (c) Curriculum Vitae of all team members;
- (d) Understanding of the mandate and the proposed approach;
- (e) Estimation of the number of working days for the accomplishment of the tasks and delivery of the required outputs; and

(f) Indication of daily rates and incidental costs.

The final ToR of the evaluation should include the roles, responsibilities, and the agreed deliverables for all external experts and their specific ToR.

Inception Reports












An inception report is sometimes prepared at the end of the Planning and Design phase to clarify the team's understanding of the intervention. While the ToR will often suffice, it is recommended that an Inception Report be the first deliverable when engaging external experts. It is a standard practice for the inception report to be formally accepted before moving on to the fieldwork stage, as any lack of understanding can significantly affect the conduct or outcomes of the evaluation.

An inception report provides consultants with a more detailed understanding of the intervention. This helps them understand the intervention before drafting an EOI. The inception report should be reasonably brief, describing the areas not covered in ToR, including:

- (a) Remaining evaluability challenges;
- (b) Interpretation of evaluation and behavioral questions;
- (c) Methodology for data collection and analysis; and
- (d) The detailed evaluation work plan.^{xv}

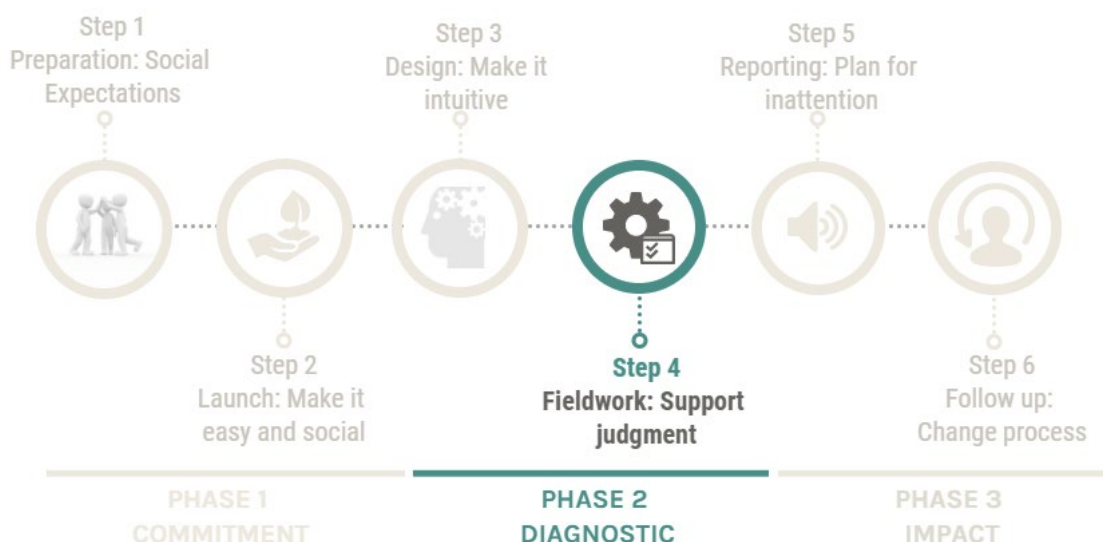
[Wrapping-up: The Planning Checklist](#)

Before moving onto the fieldwork phase of the evaluation, it is useful to reflect on the evaluation design. Using the following checklist is a helpful way to ensure that the design of the evaluation will deliver the necessary information, which will serve as a basis for useful and valuable conclusions and recommendations.

-  Have the purpose of the evaluation with behavioural insights been clearly spelled out and the use of the evaluation determined?
-  Have several key questions been identified as the focus of the evaluation with behavioural insights?
-  Have the various design options been considered and an appropriate design selected?
-  Has a stakeholder analysis been carried out and an outreach strategy prepared?
-  Have stakeholders (women and men) been identified to participate in the evaluation and their roles and responsibilities described?
-  Have key data sources been identified?
-  Have various methods been considered and a combination of the most appropriate methods selected in light of the evaluation purpose and context?
-  Have appropriate, equity-sensitive indicators been selected for each outcome and impact?
-  Do the design and methods selected take into account the key evaluation questions?
-  Do the design and methods allow for a thorough analysis of cross-cutting issues, including human rights and gender equality?
-  Have solutions been explored for the challenges anticipated?

Source: Adapted from UNEG Guide for Conducting Evaluations of Normative Work in the UN System, November 2013

STEP 4 – FIELDWORK – SUPPORT JUDGMENT



Data collection

It is critical to focus on quality data collection to conduct a credible analysis of the data and produce a valuable report. Data should be collected using a combination of social and behavioral science research methods.

A combination of the different methods described below can include desk research of the intervention's documentation and context. The various methods used will depend on the evaluation objectives, design, questions, resources allocated to the evaluation, data availability, and access to stakeholders. Evaluators and behavioral experts apply

two fundamental principles when gathering and analyzing information:

- (a) A measure must be valid: It must measure what it claims to measure.^{xvi} The methods used need to capture all outcomes – expected and unexpected, positive and negative; and
- (b) A measure must be reliable: The measure ought to produce about the same reading each time it is used.^{xvii} There must be a sufficient variety of reliable information from various sources to allow triangulation.

Description	Advantages	Challenges
Desk and Document Reviews		
Systematic analysis of existing documentation, including quantitative and descriptive information about the initiative, outputs, and outcomes, such as documentation from capacity development activities, donor reports, and other evidence.	Cost-efficient.	Documentary evidence can be complex to code and analyze. Difficult to verify the reliability and validity of data.
Interviews		
Solicit responses to questions designed to obtain in-depth information about facts and a person's experiences and views. It can be fully structured, semi, or unstructured.	Facilitates fuller coverage, range, and depth of information on a topic.	It can be time-consuming, difficult to analyze, and costly.
Direct Observation		
Notes from direct observations, pictures, recordings, etc.	Can see operations of a program as they are occurring.	It can be difficult to categorize or interpret, expensive.
Focus Group Interviews		
A small group of stakeholders (6 to 8 people) is interviewed to explore their opinions or judgments about a development initiative or policy and gather information about their behaviors, understanding, and perceptions.	A quick, reliable way to obtain common impressions from diverse stakeholders. An efficient way to get a high degree of range and depth of information in a short time.	It can be hard to analyze responses. Requires trained facilitator. It may be difficult to schedule.
Surveys (samples of respondents, including project/ program and control observations)		
A sample of the intervention population with sex disaggregation (and possibly of a control group) is extracted. Questionnaires are usually administered face-to-face by enumerators based on a prewritten and pre-coded questionnaire. Entries are recorded electronically and analyzed using quantitative and qualitative data analysis software that help produce clearly articulated findings using rigorous econometric techniques and evidence.	The sampling procedure should aim to select a statistically representative subset of the population. Large samples allow for analysis that is more refined and represent more subcategories of the population (sub-region, province, etc.)	Trained specialists are required for survey design planning and data analysis. Surveys that are more extensive can be costly and time-consuming to implement.

Sources (adapted from): (International Labour Organization, 2013) (OIOS, 2014)

Data analysis and triangulation

Once data has been obtained using one or more of the methods described, it needs to be analyzed. The purpose of data analysis is to use the data collected to develop answers to the evaluation questions. It is the process of taking the raw data and turning it into useful information from which findings, conclusions, and recommendations can be made. In analyzing the data, an evaluator is looking for themes, patterns, or recurrence of information that will help develop an understanding of the data.

The process will involve:

- (a) Assessing the validity and reliability (completeness and accuracy) of the data;
- (b) Compiling and organizing the data;
- (c) Analyzing the different data sets and making sense of the information;
- (d) Aggregating and analyzing the overall data; and
- (e) Triangulation of data to verify results and findings.

Developing findings

The findings are the results that can be seen in the analyzed data. Engaging key

stakeholders in the review of the findings serves the following purposes:

- (a) Provides a broader range of perspectives on what the findings mean and what is relevant;
- (b) Update behavioral flowcharts following new knowledge;
- (c) Provides the stakeholders with the chance to view, discuss and debate the findings, increasing their understanding and acceptance of the conclusions and recommendations; and
- (d) Strengthens the commitment of relevant stakeholders during the evaluation process.

Drawing conclusions

Conclusions provide summary judgments about the strengths and weaknesses of the evaluated intervention. They should be fair, impartial, and backed by evidence.^{xviii} Conclusions should also investigate the underlying reasons behind each issue, including behavior that is observed/documentated. Ask “WHY” and not “HOW”. The table below includes four aspects that can be used to identify the cause of the problem.

Conclusions	Behavioral problems			
	Attention	Belief Formation	Choice	Determination
Issues, challenges, including behavioral problems, will be assessed by elimination process using four behavioral insights aspects.	Is the issue due to forgetting, overlooking, multitasking, distraction, among other attention issues?	Is the issue related to pre-existing beliefs, over-or under confidence, relying excessively on rules of thumb, among other beliefs issues?	Is the issue related to a social motive, sensitivity to framing, sticking to the status quo, and disappointment?	Excessive self-directed blame, inertia, mental exhaustion, cognitive dissonance, among others?

Conclusions should be clearly linked to the findings. They must also clearly support the recommendations. The evaluators' professional views on the assessed criteria (relevance, efficiency, effectiveness, impact, sustainability, coherence, coverage, and

sustainability) and significant issues identified by the evaluation are the conclusions. The conclusions build the bridge between the past and present (results/findings) with the future (recommendations).



Findings

What and How happened
The Facts



Conclusions

Why something occurs?
A Judgement



Recommendations

Solution
strategies/recommendations

Developing recommendations

Recommendations in the evaluation report must be clearly supported by the evaluation findings, which are clearly linked to the data collected. Evaluators need to connect the underlying findings, conclusions, and reasons to the relevant solution, strategies, or recommendations.

The behavioral insights approach should be applied during the design process of the evaluation. Evaluators should reach out to relevant stakeholders and test possible solutions with them to receive feedback and commitment. During this process, they should leverage social proof, a viable strategy combined with the knowledge gained through the behavioral reduction tool. Unsupported recommendations, no

matter how brilliantly worded, risk lacking credibility, and are less likely to be adopted.

Many people prefer to choose the options that require less work, even when that means keeping the status quo and not taking an active decision. Evaluators need to reduce the resistance and facilitate the decision-making process by designing recommendations that enable the decision-making process.

Recommendations can be developed during or following the fieldwork stage. While several aspects could be considered when drafting strategies or suggestions with a BIs perspective, this guide will prepare recommendations applying the behavioral framework when designing recommendations to address a behavioral problem. The table below can be used as a tool for identifying a solution to the behavioral issue.

Recommendations	Behavioral Insights Recommendations' Strategies			
	Attention	Belief Formation	Choice	Determination
Identify the right solution to the issues	<p>If there is an attention issue, then your strategy should focus on:</p> <ul style="list-style-type: none"> • Making it relevant: state of mind, timing, placement • Seizing attention: silence, reminders, prompts, social attention • Planning for inattention: Change default, safety mechanisms 	<p>If the issue is related to a pre-existing belief, then your strategy should focus on:</p> <ul style="list-style-type: none"> • Supporting judgment; social proof, adapting or utilizing heuristics • Making it intuitive: mental models, intuitive coding • Guide searching. Use question trees, searching by aspects 	<p>If the issue is due to choice, then your strategy should focus on:</p> <ul style="list-style-type: none"> • Making it social: create a sense of community, connect with social identities • Framing interventions: framing prospects, arrange choices • Making it attractive: trigger emotions, consider motives and perspectives 	<p>If the issue is a determination, then your strategy should focus on:</p> <ul style="list-style-type: none"> • Making it easy: work with friction, default effectively cognitive avoidance • Providing plans and feedback: intentions, providing feedback • Creating social expectations. Creating commitments, leveraging social norms

Regardless of whether recommendations target a behavioral or a structural issue, recommendations should respond to the evaluation purpose and the questions that the evaluation sought to answer. Some other basic principles to consider when drafting recommendations are as follows:

- (a) Recommendations should be directed to people who will be responsible for implementing them;
- (b) Recommendations need to be practical solutions to the problem identified;
- (c) Involve stakeholders and decision-makers to get feedback and input into recommendations;

(d) The purpose of an evaluation is to contribute to improvements in the relevance, efficiency, effectiveness, impact, and/or sustainability of activities – reflecting back on these criteria helps in drafting valuable recommendations; and

(e) Ensure that recommendations are addressed explicitly to the intervention being evaluated, but consider broader applications where relevant.

Recommendations, like indicators, need to be SMART (Specific, Measurable, Achievable, Realistic, and Timely/Time-bound).



PHASE 3 IMPACT

PHASE 3: IMPACT

During this stage, evaluators will report on their findings, conclusions, and recommendations. Once the report is accepted and decision-makers have agreed to the recommendations, the follow-up process of recommendations will start. While this process looks pretty straightforward and linear, reality shows different landscape. Limited evaluation reports are used for decision-making purposes or to influence change, and a limited number of recommendations are converted into actions that could potentially have a positive impact.

The limited use of evaluation reports, could be justified for a variety of reasons:

- (a) Evaluations are used as a compliance tool, for instance, when evaluations are conducted to be accountable to a specific donor (the resources have been utilized as planned);
- (b) Evaluations are conducted for institutional use, for instance, a decision about an intervention has already been made, and the evaluation is just a confirmation of the decision adopted; and
- (c) Other reasons for limited use might be of behavioral nature, such

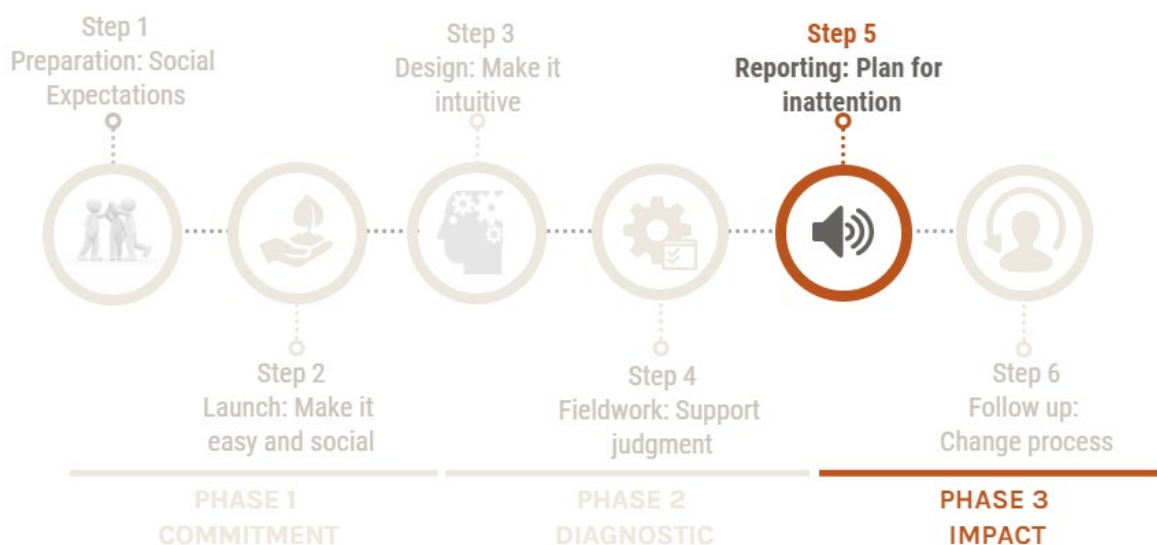
as the commitment of stakeholders, time of the evaluation, attention, and beliefs, among other behavioral issues.

This guide will focus on evaluation reports which have not addressed behavioral challenges when preparing for the last phase of the evaluation (impact). As with any interventions, evaluators and commissioners need to plan for the impact of the evaluation. Many evaluation reports do not get the expected attention, as this is often scarce, easily distracted, and quickly overwhelmed.

Decision-makers have a hard time making choices based on relevant information and aligning their decisions with their intrinsic preferences. Common examples of inattention are as follows: (i) forgetting the implementation of recommendations; (ii) multitasking; and (iii) overlooking non-obvious information. Evaluators and commissioners need to design evaluations assuming that people can not focus on everything. Behavioral problems in the evaluation process are often partially caused by attention issues in the broader definition of the behavioral principle. To strategically leverage this domain, relevant evaluation stakeholders should focus on the most critical aspects of the evaluation assignment in light of defined priorities for the office. The following questions could guide commissioners and evaluators during this phase:



STEP 5 – REPORTING – PLAN FOR INATTENTION



One of the most critical deliverables in the evaluation process is the evaluation report. A written report is a standard evaluation procedure, and it may be supplemented with presentations or other communications tools.

The objective of the evaluation report is to present the findings, conclusions, and any SMART recommendations^{xix}, to address issues and gaps and improve accountability and learning in the organization.

Draft report

It is essential to prepare a draft report, to which decision-makers and stakeholders will provide their feedback regarding the conclusions and recommendations in there. This feedback serves as a significant contribution to the final report (see next step). Failure to engage stakeholders at this stage will have considerable impacts on the acceptance and adoption of recommendations.

A checklist for preparing draft reports is as follows:

- (a) Consider customized versions for different report audiences;
- (b) Prepare a behavioral optimized report – structure and content;

- (c) Write in plain English style;
- (d) Summarize the key findings, conclusions, and recommendations up-front (e.g., in an Executive Summary and/or findings, conclusions and recommendations Matrix);
- (e) Ensure that the report conforms with the purpose stated ToR;
- (f) Ensure that the findings, recommendations, and conclusions address the key evaluation questions;
- (g) Target the report to the intended audience and, if possible, to those who will be responsible for implementing any recommendations;
- (h) Remember to outline the objectives, scope, and methodology of the evaluation and identify any limitations; and
- (i) Keep the report concise. Detailed findings and other relevant information should be included in annexes.

Evaluation reports should be drafted to seize the attention of decision-makers. The following principles should be applied for evaluation reports to be effective:

- (a) Making it relevant is a prerequisite for working effectively to create a behavioral effect to increase impact from evaluations;
- (b) Strategic placement is an overlooked dimension of making behavioral interventions relevant. Hence, failure to get people's attention to the optimally calibrated actions is a standard issue in many behavioral problems;
- (c) Seizing attention in a broad sense means that the fundamental problem of inattention is that people usually fail to attend to essential parts in a given document or a set context. Whether people forget or overlook something, and whether this is due to relegating, multitasking, or being distracted, focusing on one thing implies by definition that one is not paying attention to something else;
- (d) Leveraging saliency denotes a feature of choice architecture that draws our attention to relevant information, events, or options, at the expense of other features. The key is to understand that our attention is drawn to what is novel and seems relevant to us; and
- (e) Leveraging social attention considers BIs in relation to social aspects of attention. A concrete mechanism to be considered is the spotlight effect, which refers to how people focus on the same contextual features and options.

In addition, evaluators need to make appropriate use of visuals and illustrations whenever relevant through graphics, tables, boxes, pictures, etc.

An Example draft report is included in the Annexes. The same template applies to the final report; and

To further assist in delivering quality evaluation reports, IOD has produced a Checklist for assuring the Quality of Evaluation Reports included in the Annexes.

Stakeholders' comments

The draft report may be shared in its entirety or in a summary form. It may be preferable to share selected information with some stakeholders to focus stakeholders' attention on relevant information and/or maintain momentum. Stakeholders should be given sufficient time to provide their comments.

Sharing the draft report with stakeholders serves to:

- (a) Ensure factual accuracy and completeness of the report;
- (b) Provide a final opportunity to verify the findings, conclusions, and recommendations and ensure their accuracy and appropriateness before the publication of the final report;
- (c) Continue the process of engagement with key stakeholders, prompting ownership of the findings, conclusions, and recommendations; and
- (d) Ultimately lead to better acceptance and utilization of the report

Obtaining stakeholders' feedback may be conducted in face-to-face meetings where stakeholders have the opportunity to discuss and/or debate the findings, conclusions, and recommendations or by email, phone, video conferences or other methods.

Management response & matrix

In addition to responding in general to the draft report, management should provide a formal response to the recommendations in a Management Response Matrix. The purpose of the Management Response Matrix is to ensure that managers fully understand the recommendations and provide an action plan to implement the recommendations.

Recommendations from the draft report are entered into the Management Response Matrix by the evaluator and then shared with the Program Management. For each

recommendation, management must provide the following response:

- (a) Comments – general comments on the recommendation, including a statement of agreement or disagreement;
- (b) Actions –action(s) that will be taken for the recommendation;
- (c) Person responsible – identify the person responsible for following up on the implementation of the recommendation and actions;
- (d) Deadline – a date by which the actions will be completed; and
- (e) Closing criteria – the criteria that must be demonstrated for the recommendation to be considered implemented or closed.

Program decision-makers should be provided with a minimum of 10 working days to complete the formal response.

A Management Response Matrix Template is included in the Annexes.

Incorporating management comments

It is important to remember that this response process is intended to facilitate the refinement of the draft report. The evaluation manager will need to exercise professional judgment in incorporating comments from stakeholders in the final report to maintain a level of independence and credibility.

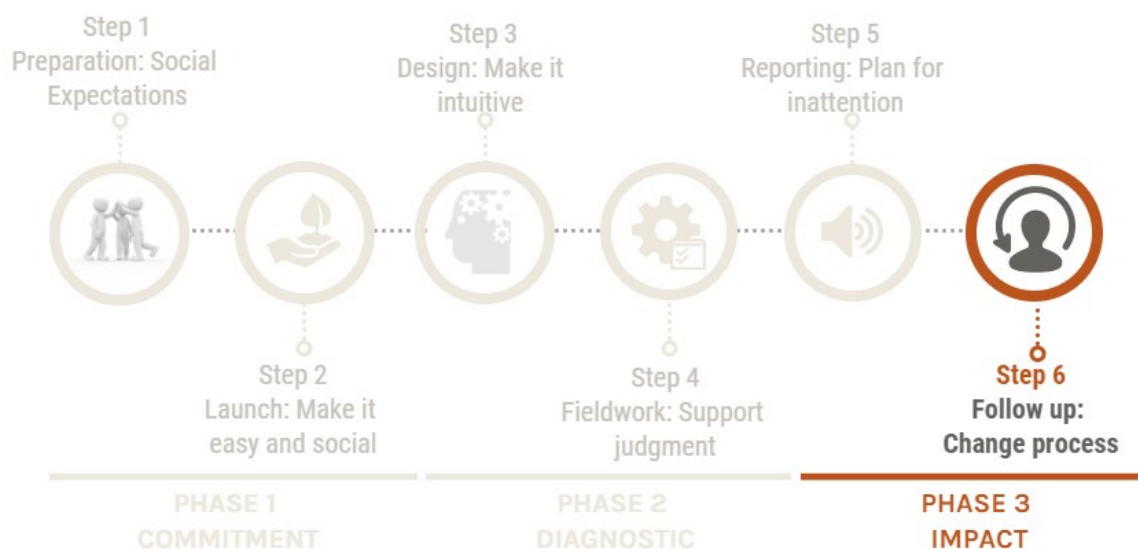
It is a standard organizational practice to include the Management Response and Action Plan in a Recommendation Table Matrix at the end of the final report. Other comments may or may not be included, depending on whether they add value to the outcome.

Outreach and communicating evaluation results

At this stage, evaluators should prepare a plan for communicating the evaluation results. This should undoubtedly become a dynamic process beyond sharing the report via email or publishing it on a website. Communicating evaluation results to different audiences should complement and reinforce the evaluation process. Evaluators and commissioners should consider the aspects presented in the figure below^{xx}:



STEP 6 – FOLLOW-UP – CHANGE PROCESS



Follow-up: Process Change

In addition to tracking the implementation of recommendations through an automatized follow-up mechanism or through formal follow-up by the evaluators, the evaluation function should reframe the messages. The evaluators and the client are at the beginning of a change process.

Evaluators should (i) highlight the benefits of the suggested recommendations or the consequences of keeping the status quo; (ii) emphasize what the organization could lose if there is no action; (iii) emphasize what the organization will gain by acting.^{xxi} At this stage, the evaluation managers should invite relevant stakeholders to identify and discuss barriers: What should be done for this to happen? (incentive structure)

As part of the change, the process includes the following review meetings with the program decision-maker after the issue of the final report:

- (a) First, a meeting to discuss the implementation of the recommendations and obtain feedback on the evaluation process;
- (b) One year after to review the implementation of the recommendations and assess its

impact, to compare before and after the implementation of the recommendation and gather the different stories; and

- (c) Further, follow up as necessary until the implementation of recommendations is completed.

While the primary purpose of these review meetings is to ensure that recommendations are implemented, an equally important reason for meeting with program decision-makers is to obtain feedback from program decision-makers that can lead to the improvement of future evaluations and recommendations.

Using evaluation results

Evaluations are undertaken to improve commitment, organizational learning, and knowledge generation. However, evaluations can be effective in advancing the organization's work only if they are used. This chapter begins by considering how the evaluation is critical for managing results and improving interventions, and how results are integrated into the program lifecycle. It also looks at how evaluations are shared at the organization and concludes with good practices for communicating results.

Much focus of evaluation processes tends to be on getting the evaluations done and doing so in time to meet deadlines. There is often less emphasis placed on using the information and evidence gained from evaluations beyond the formalities of distributing the report to management and other stakeholders.

The use of information and evidence is emphasized in the UNEG's definition of evaluation. The United Nations Evaluation Group's 2016 norms and standards for evaluation state that evaluations "should provide credible, useful evidence-based information that enables the timely incorporation of its findings, recommendations, and lessons into the decision-making processes of the organizations and stakeholders."

After all the work that goes into producing an evaluation, evaluators and decision makers need to make it meaningful and make it count. Anticipating who the different audiences may be and how they might use the results of an evaluation will help determine the best ways to organize and present results.

Timing and timeliness are also essential factors. The impact of an evaluation can depend on the decision makers' behavior and how well and quickly its findings can be available to inform strategic priorities and be communicated in appropriate ways to stakeholders when they need it. Evaluators should be aware of any behavioral and time-sensitive issues and be prepared to share preliminary key findings as they emerge.

Evaluations are more likely to be viewed more positively and acted upon when decision-makers embrace these exercises to manage results and improve interventions. An evaluation culture and enabling environment exist where it is understood that learning comes from both successes and failures.

Finally, the adoption of evaluation results is enabled when the information is customized according to the audience's needs and included in an optimized behavioral report. Good practices for effective communication of results in reports and other evaluation products are described in more detail in chapter five, including, for instance, the following approaches:

- (a) Telling a story: A report can be brought to life by using stories;
- (b) Keeping things simple: Write in an accessible style using plain language;
- (c) Being purpose-driven: Focus on what aspects are most relevant to the audience;
- (d) Segmenting the audience: Present messages in simple and easily understandable formats tailored to the specific needs of different users; and
- (e) Visualizing results: Make frequent use of visual aids such as diagrams, pictures, charts, graphs, and maps.

ACKNOWLEDGMENT

IOD wishes to thank all relevant staff members for their assistance, cooperation, and interest during this assignment.

Prepared by: Julia Engelhardt, Senior Evaluation Officer, in collaboration with Mr. Sebastian Fonzo, Managing Director, International Operations, iNudgeyou

Reviewed by: Adan Ruiz Villalba, Head of IOD Evaluation Section

Approved by: Rajesh Singh, Director IOD



ANNEXES

ANNEXES

Annex I.	Evaluation Template
Annex II.	Decentralized Evaluations Template
Annex III.	Evaluation Report Template
Annex IV.	Resources and Tools
Annex V.	Endnotes

Annex I: Evaluation Template

Evaluation Announcement Template

If a memorandum is required, then the following should be used as guidance:

- (a) Include, if possible, all information in the key information in the first two paragraphs
- (b) Make clear the actions stakeholders need to follow and what happens next

To/À:

Date:

From/De:

Cc :

**Subject/
Object:**

IOD would like to announce the upcoming *Topic of Evaluation* as per the Oversight Plan (*Year*) and invites you to collaborate with us in this evaluation process, which will take place between Month Day year and Month Day, year.

Insert a brief description of the type of evaluation to be conducted, the benefits for the stakeholders, and the expected contribution from decision-makers.

Briefly explain why the subject was selected.

Evaluations have proved to be most useful when having a designated focal person from your business unit collaborating with the evaluators during this undertaking. Ideally, the focal point should be a member of professional staff expertise in the areas of *Insert expertise required*.

Please let us know at your earliest convenience if you wish to nominate a colleague to work closely with us during this process.

We are delighted to invite you and your team for our first evaluation meeting scheduled for *Date, Time, and Location* to listen to your needs and expectations,

Do not hesitate to contact our Evaluation Section colleagues, Julia and Adan, at any time if you would like to know more about this evaluation process.

We very much look forward to our collaboration.

Your IOD Evaluation Section

Evaluability Assessment Template

Assessment Criteria

Criteria	Weighting
Decision-makers commitment	
Program Design	
Internal Management and Governance	
Availability of Information	
Context of the Program	

Performance Scores

Raw Score	Performance level	Performance requirements
4	Very good	The criteria are fully meet (evidence outmatch requirements)
3	Good	The criteria are met
2	Relatively good	The criteria are partially met
1	Poor	Insufficient identification of information related to the criteria
0	Not identified	No identification of information related to the criteria

Evaluability Scores

Evaluability	Score
Fully evaluable	≥ 3.5
Mostly Evaluable	2. 5 to 3.5
Restricted Evaluability	1. 5 to 2.5
Not Evaluable	<1.5

Evaluability Assessment Criteria	Score	Source(s) of information	Additional information
Criteria 1: Intervention Design			
<p>The definition and quality of the Program Theory (the theory of change along with the intervention theory), including its adequacy, clarity, coherence, feasibility, and relevance.</p>			
<p>The program clearly defines the problem (and thus long-term impact and outcomes) that it aims to change.</p>			
<p>The proposed steps towards achieving outcomes (logic model) are clearly and coherently defined through a continuous causal chain, and they are measurable.</p> <p>It is possible to identify which linkages in the causal chain will be most critical to the intervention's success that could thus be the focus of evaluation questions.</p>			
<p>The intended beneficiary groups are clearly identified.</p>			
<p>The intervention is relevant to the needs of the target groups as identified by any form of situation analysis, baseline study, or other evidence and argument.</p>			
Criterion Score			
Criteria 2: Internal Management and Governance			
<p>The effective management and result-oriented approach of the program, including the clear understanding and agreement of the objectives by the Division, and the assessment of risks and assumptions.</p>			

Evaluability Assessment Criteria	Score	Source(s) of information	Additional information
The implementation of the program is related to plans. Effectiveness can be assessed.			
Boundaries (if any) to achieving outcomes are identified.			
The risks associated to the intervention processes have been identified and actions to tackle them defined.			
There is ownership of staff on the Program purpose, design and implementation.			
Different stakeholders hold a common view about the intervention objectives and how they will be achieved.			
Criterion Score			
Criteria 3: Availability of Information			
The existence of adequate contents and systems for making information available.			
The program is capable of providing a complete and relevant set of documents for the evaluation.			
If reviews or evaluations have been carried out, there are reports available.			
There is a monitoring system to systematize information with defined responsibilities, sources and periodicity.			
If data is not available, staff and systems have the capacity to report evidence-based information.			
Data is being collected for all indicators.			
Baseline measures exist and provide relevant information			

Evaluability Assessment Criteria	Score	Source(s) of information	Additional information
for a future evaluation.			
Criterion Score			
Criteria 4: Context of the Program			
The conduciveness of contextual factors (including perceptions and availability of resources by the Division) to conduct the evaluation.			
The Division is capable of delivering what is needed.			
Stakeholders' expectations of the evaluation objectives and uses are conducive to an efficient evaluation process (i.e. there is alignment of stakeholders' information needs and the availability of information and resources).			
The timing of the evaluation fits into the Program cycle (i.e. there is opportunity for the evaluation to have an influence).			
There are no other factors that can hamper the evaluation (e.g. program's capacity to manage negative findings)			
Criterion Score			
Final Score (Weighted Score)			

Reference Group Terms of Reference Template

EVALUATION TOPIC



Terms of Reference for the Reference Group (RG)

Objective: Members of the RG are expected to provide technical inputs and to ensure that information is exchanged on their area of work throughout the evaluation process and to provide feedback to evaluation products (ToR, inception, and draft reports)

Composition and Scope:

The Reference Group (RG) is composed of key stakeholders who have contributed to the work in the area being evaluated. It is established for interested people to increase ownership, transparency, and learning during the evaluation process.

Role and function of RG Members:

The RG members participate in meetings convened by the evaluation decision-maker at critical junctures during the evaluation and provide feedback and advice during the evaluation process throughout its cycle.

Specifically, the RG reviews and comments

- The inception report
- The draft evaluation reports,
- The main recommendations and lessons identified and related follow-up.

Members of the RG are usually nominated by Organization Sector leads. The Evaluation Section can propose names of potential members to the Sector leads

Annex II: Decentralized Evaluations Template

Checklist for ensuring the quality of evaluation reports

Evaluation Title:

Commissioning Office/Organization:

Type of evaluation:

1. The Report Structure

1.0 The report is well structured, logical, clear and complete.

1.1 Report is logically structured with clarity and coherence (e.g., background and objectives are presented before findings, and findings are presented before conclusions and recommendations).

1.2 The title page and/or opening pages provide key basic information

1. Name of the evaluation object
2. Timeframe of the evaluation and date of the report
3. Locations (country, region, etc.) of the evaluation object
4. Names and/or Organizations of evaluators
5. Name of the Organization commissioning the evaluation
6. Table of contents which also lists Tables, Graphs, Figures, and Annexes
7. List of acronyms.

1.3 The Executive Summary is a stand-alone section of 2-3 pages that includes1:

1. Overview of the evaluation object
2. Evaluation objectives and intended audience
3. Evaluation methodology
4. Most essential findings and conclusions
5. Main recommendations

1.4 Annexes increase the credibility of the evaluation report. They may include:

1. TORs
2. List of persons interviewed and sites visited.
3. List of documents consulted
4. More details on the methodology, such as data collection instruments,
including details of their reliability and validity
5. Evaluators biodata and/or justification of team composition
6. Evaluation matrix
7. Results framework

2. Object of Evaluation

2.0 The report presents a clear and complete description of the 'object' of the evaluation

2.1 The logic model and/or the expected results chain (inputs, outputs, and outcomes) of the object is clearly described

2.2 The context of key social, political, economic,

Evaluation Title:

demographic, and institutional factors that have a direct bearing on the object is described. For example, the partner government's strategies and priorities, international, regional or country development goals, strategies and frameworks, the concerned agency's corporate goals and priorities, as appropriate

2.3 The scale and complexity of the object of the evaluation are clearly described, for example:

- The number of components, if more than one, and the size of the population each component is intended to serve, either directly and indirectly.
- The geographic context and boundaries (such as the region, country, and/or landscape and challenges where relevant
- The purpose and goal, and Organization/management of the object
- The total resources from all sources, including human resources and budget(s) (e.g. concerned agency, partner government and other donor contributions and actual expenditures
- The duration

2.4 The key stakeholders involved in the object implementation, including the implementing agency(s) and partners, other key stakeholders and their roles

2.5 The report identifies the implementation status of the object, including its phase of implementation and any significant changes (e.g. plans, strategies, logical frameworks) that have occurred over time and explains the implications of those changes for the evaluation

3. Evaluation Purpose, Objective(s) and Scope

3.0 The evaluation's purpose, objectives and scope are fully explained

3.1 The purpose of the evaluation is clearly defined, including why the evaluation was needed at that point in time, who needed the information, what information is needed, how the information will be used

3.2 The report should provide a clear explanation of the evaluation objectives and scope including main evaluation questions and describes and justifies what the evaluation did and did not cover

3.3 The report describes and provides an explanation of the chosen evaluation criteria, performance standards, or other criteria used by the evaluators

3.4 As appropriate, evaluation objectives and scope include questions that address issues of gender and human rights

4. Evaluation Methodology

Evaluation Title:

4.0 The report presents transparent description of the methodology applied to the evaluation that clearly explains how the evaluation was specifically designed to address the evaluation criteria, yield answers to the evaluation questions and achieve evaluation purposes

4.1 The report describes the data collection methods and analysis, the rationale for selecting them, and their limitations. Reference indicators and benchmarks are included where relevant

4.2 The report describes the data sources, the rationale for their selection, and their limitations. The report includes discussion of how the mix of data sources was used to obtain a diversity of perspectives, ensure data accuracy and overcome data limits

4.3 The report describes the sampling frame – area and population to be represented, rationale for selection, mechanics of selection, numbers selected out of potential subjects, and limitations of the sample

4.4 The evaluation report gives a complete description of stakeholder's consultation process in the evaluation, including the rationale for selecting the particular level and activities for consultation

4.5 The methods employed are appropriate for the evaluation and to answer its questions

4.6 The methods employed are appropriate for analysing gender and rights issues identified in the evaluation scope

4.7 The report presents evidence that adequate measures were taken to ensure data quality, including evidence supporting the reliability and validity of data collection tools (e.g. interview protocols, observation tools, etc.)

5. Findings

5.0 Findings respond directly to the evaluation criteria and questions detailed in the scope and objectives section of the report and are based on evidence derived from data collection and analysis methods described in the methodology section of the report.

5.1 Reported findings reflect systematic and appropriate analysis and interpretation of the data

5.2 Reported findings address the evaluation criteria (such as efficiency, effectiveness, sustainability, impact and relevance) and questions defined in the evaluation scope

5.3 Findings are objectively reported based on the evidence

5.4 Gaps and limitations in the data and/or unanticipated findings are reported and discussed

Evaluation Title:

5.5 Reasons for accomplishments and failures, especially continuing constraints, were identified as much as possible

5.6 Overall findings are presented with clarity, logic, and coherence

6. Conclusions

6.0 Conclusions present reasonable judgments based on findings and substantiated by evidence, and provide insights pertinent to the object and purpose of the evaluation

6.1 The conclusions reflect reasonable evaluative judgments relating to key evaluation questions

6.2 Conclusions are well substantiated by the evidence presented and are logically connected to evaluation findings

6.3 Stated conclusions provide insights into the identification and/or solutions of important problems or issues pertinent to the prospective decisions and actions of evaluation users

6.4 Conclusions present strengths and weaknesses of the object (policy, Programs, project's or other intervention) being evaluated, based on the evidence presented and taking due account of the views of a diverse cross-section of stakeholders

7. Recommendations

7.0 Recommendations are relevant to the object and purposes of the evaluation, are supported by evidence and conclusions, and were developed with the involvement of relevant stakeholders

7.1 The report describes the process followed in developing the recommendations including consultation with stakeholders

7.2 Recommendations are firmly based on evidence and conclusions

7.3 Recommendations are relevant to the object and purposes of the evaluation

7.4 Recommendations clearly identify the target group for each recommendation

7.5 Recommendations are clearly stated with priorities for action made clear

7.6 Recommendations are actionable and reflect an understanding of the commissioning Organization and potential constraints to follow-up

8. Gender and Human Rights

8.0 The report illustrates the extent to which the design and implementation of the object, the assessment of results and the evaluation process incorporate a gender equality perspective and human rights based approach

8.1 The report uses gender sensitive and human

Evaluation Title:

rights-based language throughout, including data disaggregated by sex, age, disability, etc.

8.2 The evaluation approach and data collection and analysis methods are gender equality and human rights responsive and appropriate for analyzing the gender equality and human rights issues identified in the scope.

8.3 The report assesses if the design of the object was based on a sound gender analysis and human rights analysis and implementation for results was monitored through gender and human rights frameworks, as well as the actual results on gender equality and human rights

8.4 Reported findings, conclusions, recommendations and lessons provide adequate information on gender equality and human rights aspects

A four-point rating scale: 1=poor, 4 = excellent is applied

Terms of Reference Template



Internal Oversight Division

Reference:

Terms of Reference

EVALUATION TOPIC

Date:

TABLE OF CONTENTS

LIST OF ACRONYMS.....	3
1. BACKGROUND.....	4
2. OBJECTIVES AND SCOPE OF THE EVALUATION.....	4
(A) OBJECTIVES and Use OF THE EVALUATION.....	4
(B) SCOPE.....	4
3. EVALUATION QUESTIONS (BY EVALUATION CRITERIA).....	4
(A) Questions on Relevance.....	4
(B) Questions on Effectiveness.....	4
(C) Questions on Efficiency.....	4
(D) Questions on Sustainability.....	4
(E) QUESTIONS ON LIKELIHOOD OF IMPACT.....	5
4. METHODOLOGICAL APPROACH.....	5
(A) EVALUATION DESIGN AND DESK REVIEW PHASE.....	5
(B) FIELD PHASE.....	5
(C) REPORTING PHASE.....	5
(D) KEY ISSUES THAT COULD LIMIT THE EVALUATION PROCESS.....	5
5. DELIVERABLES.....	5
6. TIME TABLE.....	6
7. evaluation management and process.....	6
ANNEXES.....	7

ACRONYMS

ASPAC	Asia and the Pacific
IAOC	Independent Advisory Oversight Committee
IOD	Internal Oversight Division
IP	Intellectual Property
LAC	Latin America and the Caribbean
LDCs	Least Developed Countries
RBMF	Results-Based Management Framework
RG	Reference Group
SG	Strategic Goal
SMT	Senior Management Team
ToR	Terms of Reference
UNEG	United Nations Evaluation Group
Organization	World Intellectual Property Organization

BACKGROUND

1. These Terms of Reference (ToR) have been developed according to United Nations Evaluation Group (UNEG) Standards and in consultation with the Development Sector, in particular the Division for Least-Developed Countries (LDCs).
2. Description of the evaluation topic.
3. Reasons for the evaluation including timing, relationship to decision-making events.
4. Other relevant background information.

OBJECTIVES AND SCOPE OF THE EVALUATION

(A) OBJECTIVES AND USE OF THE EVALUATION

5. Describe primary and any secondary objectives for the evaluation.
6. Describe how the results are expected to be used in terms of performance improvement, learning and accountability.

(B) SCOPE

7. Describe the broad scope of the evaluation: what will be included and what will be excluded.
8. Outline any sampling anticipated (optional at this stage)

EVALUATION QUESTIONS (BY EVALUATION CRITERIA)

9. Outline the key criteria that the evaluation will seek to respond to (Relevance, Efficiency, Effectiveness, Impact, Sustainability plus Coverage, Coherence, Coordination if being evaluated).

(A) QUESTIONS ON RELEVANCE

- (a) List key questions on Relevance that the evaluation will seek answers to.

(B) QUESTIONS ON EFFECTIVENESS

- (a) List key questions on Relevance that the evaluation will seek answers to.

(C) QUESTIONS ON EFFICIENCY

- (a) List key questions on Relevance that the evaluation will seek answers to.

(D) QUESTIONS ON SUSTAINABILITY

- (a) List key questions on Relevance that the evaluation will seek answers to.

(E) QUESTIONS ON IMPACT

- (a) List key questions on Relevance that the evaluation will seek answers to.

METHODOLOGICAL APPROACH

10. Identify the methodologies to be employed.
11. Outline the key data collection techniques anticipated.
12. Outline anticipated methods for data analysis.
13. Outline the key phases of the evaluation and what each phase will involve and achieve.

(A) EVALUATION DESIGN AND DESK REVIEW PHASE

14. Describe the key elements of the phase.

(B) FIELD PHASE

15. Describe the key elements of the phase.

(C) REPORTING PHASE

16. Describe the key elements of the phase.

(D) KEY ISSUES THAT COULD LIMIT THE EVALUATION PROCESS

17. Identify any issues that could impact on the evaluation in terms of achievement of objectives, scope, timeframe for the evaluation etc.

DELIVERABLES

18. Detail the anticipated deliverables. At a minimum these should include:

- (a) Final Terms of Reference
- (b) Draft Evaluation Report with Findings and Recommendations
- (c) Final Evaluation Report for publication.

19. Specify in which language(s) the deliverables will be produced.

TIME TABLE

20. Provide a draft timetable for key activities and deliverables.

Timeframe	Main phases of the evaluation
-----------	-------------------------------

(a) (b)

(c) (d)

(e) (f)

(g) (h)

(i)

EVALUATION MANAGEMENT AND PROCESS

21. Identify who will be responsible for conducting the evaluation.
22. Outline the roles and responsibilities of the various team members including any external consultants/contractors.

OVERALL EVALUATION BUDGET

23. Provide an overall evaluation budget (Optional) (Note: For External Evaluations, the Final Terms Of Reference must contain the agreed budget for the evaluation as the external contractor must be accountable for the completion of the evaluation within the budget.)

ANNEXES

(OPTIONAL – ADD/DELETE AS NECESSARY)

Annex VI.	INITIAL ANALYSIS OF THE EVALUATION TOPIC
Annex VII.	TENTATIVE LIST OF STAKEHOLDERS
Annex VIII.	DETAILED BUDGET

EXAMPLE OF ANNEX

Example Template for Categorizing Stakeholders

Beneficiaries of the intervention	Users (clients) of the evaluation
<p>Primary (direct/indirect) Government counterparts (direct) Groups representing minorities (direct) Minorities (indirect)</p>	<p>Primary Intervention team Local decision makers High level decision makers (headquarters and region) Government counterparts Member states Specific committees</p>
<p>Secondary Senior government officials (direct) Minority cultural specialists, linguistic scholars and religious leaders (direct and indirect)</p>	<p>Secondary Collaborating donors Collaborating partners, NGOs etc. Other UN Organizations Groups representing minorities Academics and evaluation professionals</p>
<p>Others Other minority groups or government officials not reached by the intervention</p>	<p>Beneficiaries</p>

Evaluation Questions Matrix

	Overarching Question	Sub-questions	Indicators/Measures	Key Sources of Information	Main Methods
Relevance	1.		•	•	•
	2.		•	•	•
Effectiveness	3.		•	•	•
	4.		•	•	•
Efficiency	5.		•	•	•
	6.		•	•	•
Impact	7.		•	•	•
	8.		•	•	•
Sustainability	9.		•	•	•
	10.		•	•	•
Coherence	11.		•	•	•
	12.		•	•	•
Coordination	13.		•	•	•
	14.		•	•	•
Coverage	15.		•	•	•
	16.		•	•	•

Overview of Findings, Conclusions and Recommendations

FINDINGS	CONCLUSIONS	RECOMMENDATIONS
Criterion A		
Finding A.1		
Finding A.2	Conclusion A	
Finding A.3		Recommendation #
Finding A.4	Conclusion B	
Criterion B		
Finding B.1		
Finding B.2	Conclusion C	
Finding B.3		Recommendation #
Finding B.4	Conclusion D	
Criterion C		
Finding C.1		
Finding C.2	Conclusion E	
Finding C.3		Recommendation #
Finding C.4	Conclusion F	

Management Response Matrix

Recommendation #	Accepted / Rejected (indicate reason for rejecting)	Person(s) Responsible	Deadline	Management Comments and Action Plan
Recommendation #				
Recommendation #				
Recommendation #				
Recommendation #				

Annex III: Evaluation Report Template



Internal Oversight Division

Reference:

EVALUATION REPORT

Report details

Date

TABLE OF CONTENTS

<u>LIST OF ACRONYMS</u>	3
<u>EXECUTIVE SUMMARY</u>	4
<u>1. INTRODUCTION</u>	5
<u>2. DESCRIPTION OF THE INTERVENTION</u>	5
<u>3. EVALUATION SCOPE AND OBJECTIVES</u>	5
<u>4. EVALUATION APPROACH AND METHODS</u>	6
<u>5. DATA ANALYSIS</u>	6
<u>6. FINDINGS AND CONCLUSIONS</u>	7
<u>7. RECOMMENDATIONS</u>	7
<u>8. LESSONS</u>	7
<u>9. REPORT ANNEXES</u>	7
<u>10. ACKNOWLEDGMENT</u>	8
<u>TABLE OF RECOMMENDATIONS</u>	9
<u>ANNEXES</u>	

LIST OF ACRONYMS

Organization	World Intellectual Property Organization
IOD	Internal Oversight Division

EXECUTIVE SUMMARY

A stand-alone section of two to three pages that should:

- Briefly describe the intervention (the project(s), program(s), policies or other interventions) that was evaluated.
- Explain the purpose and objectives of the evaluation, including the audience for the evaluation and the intended uses.
- Describe key aspect of the evaluation approach and methods.
- Summarize principle findings, conclusions, and recommendations.

INTRODUCTION

Should:

- Explain why the evaluation was conducted (the purpose), why the intervention is being evaluated at this point in time, and why it addressed the questions it did.
- Identify the primary audience or users of the evaluation, what they wanted to learn from the evaluation and why and how they are expected to use the evaluation results.
- Identify the intervention (the project(s) program(s), policies or other interventions) that was evaluated—see upcoming section on intervention.
- Acquaint the reader with the structure and contents of the report and how the information contained in the report will meet the purposes of the evaluation and satisfy the information needs of the report's intended users.

DESCRIPTION OF THE INTERVENTION

Provides the basis for report users to understand the logic and assess the merits of the evaluation methodology and understand the applicability of the evaluation results. The description needs to provide sufficient detail for the report user to derive meaning from the evaluation. The description should:

- Describe **what is being evaluated, who seeks to benefit**, and the **problem or issue** it seeks to address.
- Explain the **expected results map or results framework, implementation strategies**, and the key **assumptions** underlying the strategy.
- Link the intervention to **national priorities**, UNDAF priorities, corporate multiyear funding frameworks or strategic plan goals, or other **program or country specific plans and goals**.
- Identify the **phase** in the implementation of the intervention and any **significant changes** (e.g., plans, strategies, logical frameworks) that have occurred over time, and explain the implications of those changes for the evaluation.
- Identify and describe the **key partners** involved in the implementation and their roles.
- Describe the **scale of the intervention**, such as the number of components (e.g., phases of a project) and the size of the target population for each component.
- Indicate the **total resources**, including human resources and budgets.
- Describe the context of the **social, political, economic and institutional factors**, and the **geographical landscape** within which the intervention operates and explain the effects (challenges and opportunities) those factors present for its implementation and outcomes.
- Point out **design weaknesses** (e.g., intervention logic) or other **implementation constraints** (e.g., resource limitations).

EVALUATION SCOPE AND OBJECTIVES

The report should provide a clear explanation of the evaluation's scope, primary objectives and main questions.

- Evaluation scope - The report should define the parameters of the evaluation, for example, the time period, the segments of the target population included, the geographic area included, and which components, outputs or outcomes were and were not assessed.
- Evaluation objectives - The report should spell out the types of decisions evaluation users will make, the issues they will need to consider in making those decisions, and what the evaluation will need to achieve to contribute to those decisions.

- Evaluation criteria - The report should define the evaluation criteria or performance standards used. The report should explain the rationale for selecting the particular criteria used in the evaluation.
- Evaluation questions - Evaluation questions define the information that the evaluation will generate. The report should detail the main evaluation questions addressed by the evaluation and explain how the answers to these questions address the information needs of users.

EVALUATION APPROACH AND METHODS

The evaluation report should describe in detail the selected methodological approaches, methods and analysis; the rationale for their selection; and how, within the constraints of time and money, the approaches and methods employed yielded data that helped answer the evaluation questions and achieved the evaluation purposes. The description should help the report users judge the merits of the methods used in the evaluation and the credibility of the findings, conclusions and recommendations. The description on methodology should include discussion of each of the following:

- Data sources - The sources of information (documents reviewed and stakeholders), the rationale for their selection and how the information obtained addressed the evaluation questions.
- Sample and sampling frame - If a sample was used: the sample size and characteristics; the sample selection criteria (e.g., single women, under 45); the process for selecting the sample (e.g., random, purposive); if applicable, how comparison and treatment groups were assigned; and the extent to which the sample is representative of the entire target population, including discussion of the limitations of the sample for generalizing results.
- Data collection procedures and instruments - Methods or procedures used to collect data, including discussion of data collection instruments (e.g., interview protocols), their appropriateness for the data source and evidence of their reliability and validity.
- Performance standards - The standard or measure that will be used to evaluate performance relative to the evaluation questions (e.g., national or regional indicators, rating scales).
- Stakeholder engagement - Stakeholders' engagement in the evaluation and how the level of involvement contributed to the credibility of the evaluation and the results.
- Ethical considerations - The measures taken to protect the rights and confidentiality of informants (see UNEG' Ethical Guide for Evaluators' for more information).
- Background information on evaluators - The composition of the evaluation team, the background and skills of team members and the appropriateness of the technical skill mix, gender balance and geographical representation for the evaluation.
- Major limitations of the methodology - Major limitations of the methodology should be identified and openly discussed as to their implications for evaluation, as well as steps taken to mitigate those limitations.

DATA ANALYSIS

The report should describe the procedures used to analyze the data collected to answer the evaluation questions. It should detail the various steps and stages of analysis that were carried out, including the steps to confirm the accuracy of data and the results. The report also should discuss the appropriateness of the analysis to the evaluation questions. Potential weaknesses in the data analysis and gaps or limitations of the data should be discussed, including their possible influence on the way findings may be interpreted and conclusions drawn.

FINDINGS AND CONCLUSIONS

The report should present the evaluation findings based on the analysis and conclusions drawn from the findings.

- Findings - Should be presented as statements of fact that are based on analysis of the data. They should be structured around the evaluation criteria and questions so that report users can readily make the connection between what was asked and what was found. Variances between planned and actual results should be explained, as well as factors affecting the achievement of intended results. Assumptions or risks in the intervention or program design that subsequently affected implementation should be discussed.
- Conclusions - Should be comprehensive and balanced, and highlight the strengths, weaknesses and outcomes of the intervention. They should be well substantiated by the evidence and logically connected to evaluation findings. They should respond to key evaluation questions and provide insights into the identification of and/or solutions to important problems or issues pertinent to the decision-making of intended users.

RECOMMENDATIONS

The report should provide practical, feasible recommendations directed to the intended users of the report about what actions to take or decisions to make. The recommendations should be specifically supported by the evidence and linked to the findings and conclusions around key questions addressed by the evaluation. They should address sustainability of the initiative and comment on the adequacy of the intervention's exit strategy, if applicable.

LESSONS

As appropriate, the report should include discussion of lessons from the evaluation, that is, new knowledge gained from the particular circumstance (intervention, context outcomes, even about evaluation methods) that are applicable to a similar context. Lessons should be concise and based on specific evidence presented in the report.

REPORT ANNEXES

Suggested annexes should include the following to provide the report user with supplemental background and methodological details that enhance the credibility of the report:

- ToR for the evaluation
- Additional methodology-related documentation, such as the evaluation matrix and data collection instruments (questionnaires, interview guides, observation protocols, etc.) as appropriate
- List of individuals or groups interviewed or consulted and sites visited
- List of supporting documents reviewed
- Intervention results map or results framework
- Summary tables of findings, such as tables displaying progress towards outputs, targets, and goals relative to established indicators
- Short biographies of the evaluators and justification of team composition
- Code of conduct signed by evaluators

ACKNOWLEDGMENT *IOD wishes to thank all relevant members of staff for their assistance, cooperation and interest during this assignment.*

Prepared by:

Reviewed by:

Approved by:

TABLE OF RECOMMENDATIONS

Recommendation #	Accepted / Rejected (indicate reason for rejecting)	Person(s) Responsible	Deadline	Management Comments and Action Plan
Recommendation #				
Recommendation #				
Recommendation #				

Annex IV: Resources and Tools

Organization Links and Tools

Organization Evaluation Policy, 2016

Organization Evaluation Strategy, 2016

[Organization Evaluation Code of Conduct](#)

[Organization IOD Charter, 2014](#)^{*1}

Internal Oversight Division, 2016 Oversight Plan

[Knowledge Management in Organization.](#)

TeamCentral User Guide

Checklist for assessing the quality of evaluation reports of the Organization IOD Evaluation Function

Organization recruitment policies and procedures

Decentralized Evaluations Template

[Organization Procurement Guide](#)

External Links and tools

UNEG

United Nations Evaluation Group, 2005: [Norms for Evaluation in the UN System](#), New York: UNEG.

United Nations Evaluation Group, 2008: [Code of Conduct](#) for Evaluation on the UN System, New York: UNEG.

United Nations Evaluation Group, 2008: [Ethical Guide](#). New York: UNEG

United Nations Evaluation Group, 2010: [Good Practice](#) Guide for Follow up to Evaluations, New York: UNEG.

UNEG [Quality Checklist](#) for Evaluation Terms of Reference and Inception Reports.

[UNEG Integrating Human Rights and Gender Equality in Evaluations](#)

[UNEG Quality Checklist for Evaluation Reports.](#)

OECD/DAC

[OECD/DAC, 2010b](#). Quality Standards for Development Evaluation, Paris: OECD.

¹ Last amended on October 2, 2018

[OECD/DAC, 2010c](#). Managing Joint Evaluations, Paris: OECD.

[OECD/DAC, 2002](#). Glossary of Key Terms in Evaluation and Results Based Management, Paris: OECD.

OIOS-IED

[OIOS-IED Inspection and Evaluation Guide, New York, 2014: OIOS.](#)

[OIOS-IED Reference Documents](#)

Very extensive list of evaluation reference documents including a large catalogue of checklists for all stages of evaluations

Other

[IFAD, 2009. Evaluation Guide](#) - Methodology and Processes, Rome: Office of Evaluation.

United Nations Development Program, 2009: [Guide](#) on Planning, Monitoring and Evaluating for Development Results, New York: UNDP.

[ILO policy](#) guide for results-based evaluation: Principles, rationale, planning and managing for evaluations, 2nd ed. (July 2013)

Active Learning Network for Accountability and Performance in Humanitarian Action ([ALNAP](#))

[Australian Bureau of Statistics](#) – Sampling and data Organization techniques

[Better Evaluation](#)

[Gender and Evaluation](#)

[Integrated Monitoring and Documentation Information System \(IMDIS\) data](#)

[United Nations Joint Inspection Unit \(JIU\)](#)

[My M&E](#)

[Random number generators](#)

[Research Rundowns](#) – [significance testing](#)

[Sample size calculators](#)

Social Research Methods

<http://www.socialresearchmethods.net/kb/sampon.php>

<http://www.socialresearchmethods.net/kb/relandval.php>

[Sophia Learning](#)

[Stat Trek](#)

Topic Specific Resources

Gender Equity
[Gender Equality Glossary](#), UNWOMEN

Data Collection
[OIOS-IED Inspection and Evaluation Guide](#), 2014 contains some of the best, detailed information on data collection, particularly well suited to the Organization context.

Annex V: Endnotes

- ⁱ Behavioural Sciences – UN Innovation Network. Modified by Organization IOD Evaluation Section
- ⁱⁱ Behavioural Sciences – UN Innovation Network. Modified by Organization IOD Evaluation Section
- ⁱⁱⁱ Referred to as “choice architecture” in Richard Thaler and Cass Sunstein’s, *Nudge – Wie Man Kluge Entscheidungen anstoesst* (2011)
- ^{iv} The definitions have been slightly rephrased from the original reference.
<http://www.oecd.org/dac/evaluation/dacriteriaforevaluatingdevelopmentassistance.htm>
- ^v Beneficiaries is defined as, “the individuals, groups, or organisations, whether targeted or not, that benefit directly or indirectly, from the development intervention.” Other terms, such as rights holders or affected people, may also be used.
- ^{vi} Lusthaus C., Adrien M., Anderson C. and Carden F. 1999
- ^{vii} United Nations: Universal Declaration of Human Rights, Article 27 (1) and (2)
- ^{viii} <http://www.unwomen.org/en/how-we-work/un-system-coordination/promoting-un-accountability>
- ^{ix} UN Women Independent Evaluation Office (2015). How to Manage Gender-Responsive Evaluation: Evaluation Guide.
- ^x Glossary of Key Terms in Evaluation and Results Based Management. OECD Development Cooperation Directorate. France
- ^{xi} Patto, Michael Quinn. Developmental Evaluation: applying complexity concepts to enhance innovation and use. The Guildford Press. New York, 2011
- ^{xii} An analysis of the planned activities, expected outputs, outcomes and impacts and relevant indicators and assumptions which will be used to guide the direction of the evaluation
- ^{xiii} Cheuk Fan Ng (2016). Behavioral Mapping and Tracking – Research Methods for Environmental Psychology, Volume 29
- ^{xiv} OECD (2019), Tools and Ethics for Applied Behavioural Insights: The BASIC Toolkit, OECD Publishing, Paris, <https://doi.org/10.1787/9ea76a8f-en>.
- ^{xv} Looking Back, Moving Forward: Sida Evaluation Guide
- ^{xvi} Kelly, T. L. (1927). Interpretation of Educational Measurements. Macmillan Press. New York.
- ^{xvii} Nunnally, J. (1978). Psychometric Theory. McGraw Hill. New York.
- ^{xviii} ILO Policy Guide for Results Based Evaluations
- ^{xix} Specific, measurable, achievable, relevant and time bound
- ^{xx} London School of Economics : Playbook – Behavioural Change. Response – A behavioural checklist for designing effective communications. Practitioners’ Playbook
- ^{xxi} London School of Economics : Playbook – Behavioural Change. Response – A behavioural checklist for designing effective communications. Practitioners’ Playbook

[End of Annexes and of Document]