



UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION



eio

independent
evaluation

Evaluation Manual

Revision – July 2023

Office of Evaluation and Internal Oversight
Independent Evaluation Unit

Evaluation Manual

Revision – July 2023



UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

Vienna, 2023

FOREWORD

In a world where evidence and data-driven decision-making, learning and accountability are increasingly valued, evaluations play a pivotal role in objectively assessing the effectiveness, efficiency, and impact of programs, policies, and initiatives. They provide crucial insights that guide organizations, governments, and individuals in making informed choices, driving innovation, and improving outcomes, impact, and sustainable transformational changes.

Evaluation is a dynamic and multifaceted discipline. It encompasses a wide range of methodologies, tools, and approaches, all aimed at answering critical questions, uncovering valuable lessons, and informing future actions.

This Evaluation Manual has been crafted with the intention of demystifying the evaluation process in UNIDO, and empowering evaluators, project managers, and stakeholders alike. It serves as a comprehensive guide, offering a rich array of knowledge, practical guidance, and best practices to help you navigate the evaluation journey successfully. Whether you are a seasoned evaluator, a programme manager responsible for commissioning evaluations, or a policymaker interested in understanding how evaluations can inform your decisions, this manual has something valuable to offer.

Drawing on the collective wisdom of evaluation experts, this manual provides a step-by-step approach to evaluation planning, design, data collection and analysis, and reporting. It delves into emerging trends and innovative techniques that are reshaping the evaluation landscape. Furthermore, it emphasizes the importance of collaboration, stakeholder engagement, and utilization of evaluation findings to maximize the impact of evaluations and drive meaningful change.

While this manual offers practical tools and frameworks, it is essential to remember that evaluation is not a one-size-fits-all endeavour. Context matters, and every evaluation should be tailored to the unique circumstances, needs and goals of the initiative being evaluated. The manual encourages critical thinking, flexibility, and adaptability, acknowledging that evaluation is an iterative process that evolves with the changing dynamics of programmes and policies.

It is our sincere hope that this Evaluation Manual will serve as a trusted companion and valuable resource on our evaluation journey. May it inspire you to approach evaluations with curiosity, rigour, and a commitment to learning. May it equip you with the knowledge and skills to navigate the complexities of evaluation and contribute to evidence-informed decision-making. And, above all, may it foster a culture of evaluation that embraces transparency, accountability and continuous improvement.

We extend our gratitude to all the authors, contributors and reviewers who have generously shared their expertise, insights and experiences in shaping this manual. Their dedication to advancing the field of evaluation is commendable, and their contributions will undoubtedly make a lasting impact on UNIDO and beyond.

Remember, evaluation is not merely an exercise in assessing projects; it is a powerful tool for positive change. By harnessing the potential of evaluation, we can create a world where decisions are grounded in evidence, resources are optimized, and the well-being of individuals and communities is enhanced.

Let us embark on this evaluation journey together, equipped with knowledge, guided by curiosity, and driven by a shared commitment to creating a better future.

UNIDO Evaluation Team
June 2023

FIGURES	7
ACRONYMS AND ABBREVIATIONS	8
1. BACKGROUND	9
1.1 Who is the manual for?	10
1.2 Why are evaluations done?	11
1.3 When should evaluations not be undertaken?	12
2. WHAT IS EVALUATION?	13
2.1 Definition	14
2.2 Evaluation and related concepts: Role within results-based management	14
2.2.1 Results-based management	15
2.2.2 UNIDO Three Lines Model	16
2.2.3 Monitoring	16
2.3 UNIDO evaluation types and dimensions	17
2.3.1 Self-evaluations	18
2.3.1 Independent evaluations	18
2.4 Key factors behind successful evaluations	19
2.5 Quality standards	19
2.6 Evaluation principles	20
3. METHODOLOGY	22
3.1 Evaluation criteria	23
3.2 Formulating evaluation questions	25
3.3 Rating systems	31
3.3.1 Project design	32
3.3.2 Project performance	33
3.3.3 Performance of partners	33
3.3.4 Human rights and gender mainstreaming	34
3.3.5 Environmental and social safeguards (ESS)	36
3.3.6 Other evaluation approaches	40
3.3.7 Logframe	41
3.3.8 Theory of change	41
3.4 Data collection methods and tools	46
3.5 Data analysis	49
3.6 The “why” and “how” questions: Understanding change causes and processes	51
3.7 Evaluating impact: progress to impact at completion point	51
3.8 The challenge of evaluating impact	54
3.8.1 Impact evaluation techniques	55
3.9 Benchmarking	57
3.10 Formulating recommendations and management action plans (MAPs)	57

3.11 Participation	59
3.12 Using the evaluation	59
3.12.1 Evaluation users	60
3.12.2 Dissemination of the evaluation report	60
3.12.3 Lessons learned and good practices	60
4. THE PROJECT EVALUATION PROCESS: STEP-BY-STEP	62
4.1 Background	63
4.1.1 Planning an evaluation at the design stage	64
4.2 Evaluation planning and budget	64
4.3 Preparation of evaluation task	65
4.3.1 Developing terms of reference (ToR)	65
4.3.2 Intervention rationale	67
4.3.3 Assessing evaluability	67
4.3.4 Evaluation team composition and recruitment	68
4.4 Evaluation conduct and reporting	71
4.4.1 Inception phase	71
4.4.2 Quality assurance of inception report	72
4.4.3 Field work	72
4.4.4 Presentation of preliminary findings	73
4.4.5 Reporting	73
4.5 Follow-up, learning, and dissemination	75
4.5.1 Management response to evaluations and internal follow-up	75
4.5.2 Learning and dissemination	76
5. THE PROCESS FOR STRATEGIC EVALUATIONS: COUNTRY, THEMATIC, CORPORATE-LEVEL AND IMPACT EVALUATIONS	78
5.1 Background	79
5.1.1 Selection and timing	79
5.2 Country-level evaluations	81
5.2.1 Purpose	81
5.2.2 Scope	81
5.2.3 Process and methodology	82
5.3 Thematic evaluations	83
5.3.1 Purpose	83
5.3.2 Scope	83
5.3.3 Process and methodology	83
5.4 Impact evaluations	85
5.4.1 Objectives	85
5.4.2 Purpose	85
5.4.3 Scope	85

5.4.4 Challenges of impact evaluations	85
5.4.5 Process and methodology	86
ANNEX 1: References to good evaluation practices	
KEY EVALUATION QUESTIONS	88
ANNEX 2: Definition of evaluation criteria including key evaluation questions	89
FIGURES	
Figure 1: The connecting role of UNIDO’s Evaluation Manual	10
Figure 2: The purposes of evaluation in UNIDO	11
Figure 3: Categories and examples of misusing evaluation	12
Figure 4: UNIDO Three Lines Model (3LM)	16
Figure 5: The concepts of monitoring and evaluation in UNIDO	17
Figure 6: Independent evaluation types in UNIDO	18
Figure 7: Definition of evaluation criteria	23
Figure 8: Good practices for the formulation of evaluation questions	26
Figure 9: UNIDO evaluation rating scale	31
Figure 10: Formula transforming UNIDO ratings into GEF ratings	32
Figure 11: Presentation of rating a theory of change with colour coding	44
Figure 12: Rating a theory of change in tabular form	45
Figure 13: Main data collection methods used in UNIDO evaluations	47
Figure 14: Examples of data analysis in UNIDO evaluations	50
Figure 15: Progress towards impact	52
Figure 16: Impact dimensions	53
Figure 17: CHANGE framework for evaluating impact towards policy and institutional change	56
Figure 18: Flowchart for Independent project evaluations	63
Figure 19: UNEG Evaluation Competence Framework	69
Figure 20: Planning and selection of strategic evaluations	79
Figure 21: Flowchart for strategic evaluations	80
Figure 22: Common challenges in UNIDO country evaluations and possible solutions	83
Figure 23: Common challenges in UNIDO thematic evaluations and possible solutions	84

ACRONYMS AND ABBREVIATIONS

COP	Conference of the Parties
CP	Cleaner Production
DAC	Development Assistance Committee
DFAT	Australia's Department of Foreign Affairs and Trade
EO	Evaluation Officer
EU	European Union
GEF	Global Environment Facility
IDB	Industrial Development Board
IEU	UNIDO Independent Evaluation Unit
IFAD	International Fund for Agricultural Development
IP	Integrated Programme
ISID	Inclusive and Sustainable Industrial Development
LF/Logframe	Logical Framework
MP	Montreal Protocol
MRS	Management Response Sheet
MTPF	Medium-Term Programme Framework
M&E	Monitoring and Evaluation
M&R	Monitoring and Reporting
NGO	Non-Governmental Organization
ODE	Office of Development Effectiveness (Australia)
OECD	Organization for Economic Development and Cooperation
OIOS	Office for Internal Oversight Services
p.	Page
PM	Project Manager
PTC	Programme Development and Technical Cooperation
QA	Quality Assurance
QC	Quality Checklist
RBM	Results-based Management
Sida	Swedish International Development Cooperation Agency
SDG	Sustainable Development Goals
TC	Technical Cooperation
TL	Team Leader
TOC	Theory of Change
ToR	Terms of Reference
UN	United Nations
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNIDO	United Nations Industrial Development Organization
USAID	United States Agency for International Development
USD	United States Dollar
WIPO	World Intellectual Property Organization

1. Background

This section explains how the concept of evaluation at UNIDO relates to the broader sustainable development agenda. It highlights the important role of evaluation within the Organization and shows how the Evaluation Manual complements the Charter of the Office of Evaluation and Internal Oversight and UNIDO's updated Evaluation Policy. It also describes who the manual is for and outlines the rationale behind evaluations in UNIDO. Finally, the section closes with an explanation on how to minimize the misuse of evaluation.

UNIDO’s governing body, the Industrial Development Board (IDB), recognizes the importance of a strong, effective, efficient and comprehensive independent evaluation function that is aligned with international standards and practices for providing assurance and assessing the results and impact of UNIDO’s work. This is in line with United Nations General Assembly Resolution 59/250 of 2004 calling for the systematic evaluation of operational activities of the United Nations System by assessing their impact on poverty eradication, economic growth and sustainable development.

In light of the 2030 Agenda for Sustainable Development and its central aim to “Leave No One Behind”, at its forty-fourth session in 2016, the IDB adopted a decision “... inviting the Director General to (...) continue reporting on UNIDO’s contribution to the realization of the 2030 Agenda and its Sustainable Development Goals in the Annual Report”.¹ In recognition of the ongoing role of evaluation in independently assessing and providing assurance on the progress of UNIDO’s contribution to the SDGs, the IDB also adopted a decision that, inter alia, “encouraged evaluations on results at outcome and impact levels, and the incorporation of information on performance and lessons into management and strategic planning processes”.²

In support of the IDB decisions outlined above, UNIDO produced an Evaluation Manual in 2018 in order to provide guidance on how to implement UNIDO’s Evaluation Policy and Charter in line with evaluation documents and tools as presented in **Figure 1**.

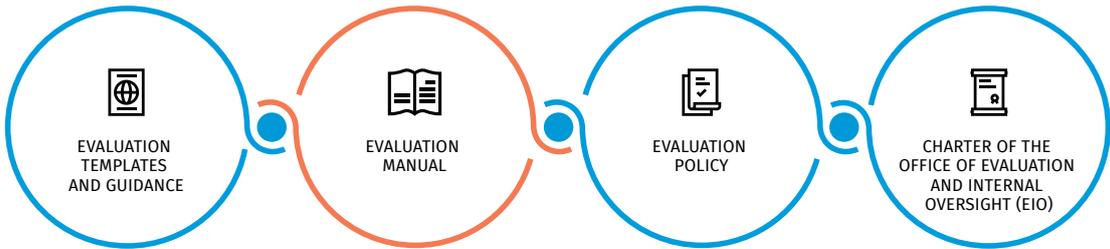


Figure 1: The connecting role of UNIDO’s Evaluation Manual

This current version of the Evaluation Manual is a necessary revision to the 2018 edition, following the publication of the new Charter of the Office of Evaluation and Internal Oversight (EIO)³ issued in 2020 and an update to UNIDO’s Evaluation Policy⁴ in 2021. It also provides an opportunity to incorporate emerging good practices from the international evaluation community as part of continuous efforts to strengthen the evaluation process in UNIDO.

1.1 Who is the manual for?

This manual is aimed primarily at UNIDO staff and consultants involved directly in evaluations and at UNIDO’s management and their teams. By consulting the manual, UNIDO management will obtain a better understanding of their roles and responsibilities in car-

¹ UNIDO, 2016, Industrial Development Board. Forty-fourth session. Vienna, 22 – 24 November 2016. Report of the Industrial Development Board (document GC.17/2, decision IDB.44/Dec.2(c)). (https://www.unido.org/sites/default/files/2016-12/GC_17_2_Report_of_the_IDB_44_e_0.pdf)

² UNIDO, 2016, Industrial Development Board. Forty-fourth session. Vienna, 22–24 November 2016. Report of the Industrial Development Board (document GC.17/2, decision IDB.44/Dec.12(c)). (https://www.unido.org/sites/default/files/2016-12/GC_17_2_Report_of_the_IDB_44_e_0.pdf)

³ UNIDO, 2020, Director General’s Bulletin. Charter of the Office of Evaluation and Internal Oversight. (document DGB/2020/11)

⁴ UNIDO, 2021, Director General’s Bulletin. Evaluation Policy. (document DGB/2021/11)

rying out an evaluation. Evaluation consultants working for UNIDO will gain from getting a clear and consistent view of the importance of evaluation and how it is practiced at UNIDO.

It may also be useful for UNIDO senior management and stakeholders in Member States who will gain insight into how evaluations are carried out, giving them a deeper understanding of the robustness of the process, and the consistently high quality of evaluation products.

The manual’s main purpose is to ensure consistency, rigour and transparency across independent evaluations and, ultimately, to enhance the effectiveness of the independent evaluation function at UNIDO.

1.2 Why are evaluations done?

The three main aims of evaluation as set out in the 2021 UNIDO Evaluation Policy are: i) to promote accountability; ii) to support results-based management (RBM); and iii) to drive learning and innovation. As such, evaluation provides UNIDO management and stakeholders with valuable information and contributes to improved policymaking based on evidence-based decision-making.

Figure 2 summarizes the three key purposes of the evaluation process and how different groups of stakeholders benefit from it both within and outside UNIDO.

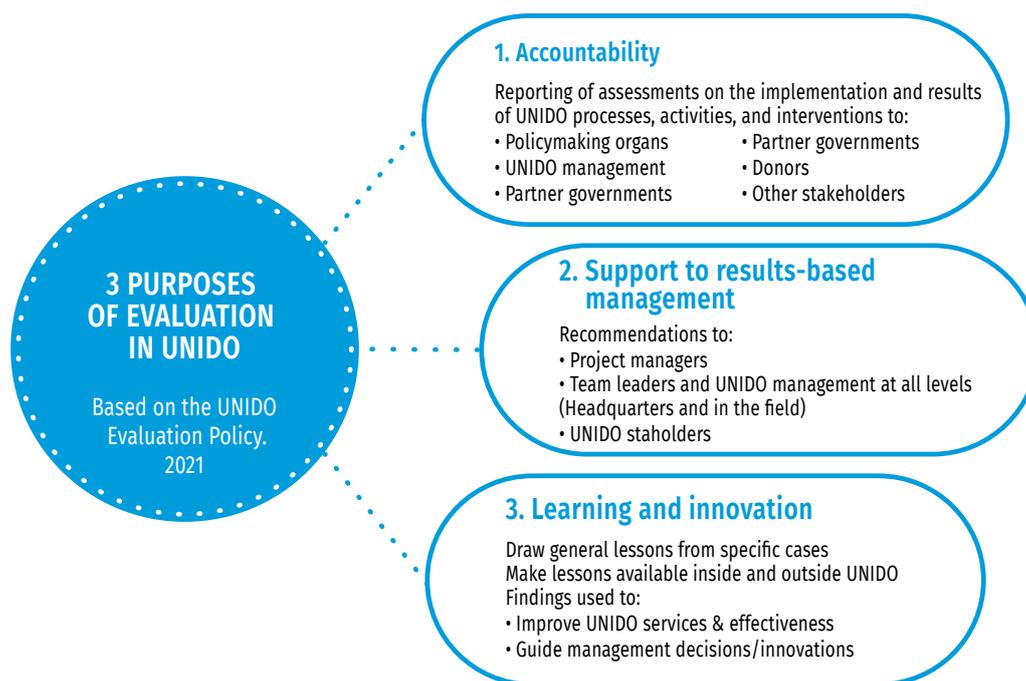


Figure 2: The purposes of evaluation in UNIDO

Evaluation promotes accountability by reporting the results of assessments of UNIDO processes, activities and interventions to the policymaking organs, management, partner governments, donors, and other stakeholders of UNIDO.

Evaluations support RBM through recommendations to project managers, team leaders, UNIDO management at headquarters and in the field, as well as to UNIDO stakeholders. Recommendations should be practical and evaluation users should actively participate

in the evaluation process. Recommendations, when accepted by management, should be developed into management action plans that are followed up regularly.

Evaluations drive organizational learning and innovation. To this end, they attempt to draw general lessons from specific cases, which are then made available to all those both inside and outside UNIDO who may benefit from their conclusions. In this way, findings are used to improve the services and effectiveness of the Organization, as well as to guide management decisions and/or innovation.

Drawing lessons is related to learning from experience. Evaluation enhances learning and innovation by assessing to what extent particular interventions, strategies or policies have worked or not, and why. Understanding the reasons for success or failure supports UNIDO in scaling up successful approaches and avoiding repeating mistakes.

The contribution of evaluation to organizational learning goes beyond UNIDO. The UNIDO evaluation function shares lessons learned from the evaluations it conducts and manages with other organizations of the United Nations System and relevant stakeholders.⁵

1.3 When should evaluations not be undertaken?

The majority of evaluations adhere to strict guidelines, however, the occasional intentional or unintentional misuse of evaluation is a reality in evaluating international development in the UN system and beyond.

Figure 3⁶ sets out some of the ways in which this can occur in order to help staff, consultants and stakeholders to identify cases where evaluations are not the appropriate tool:

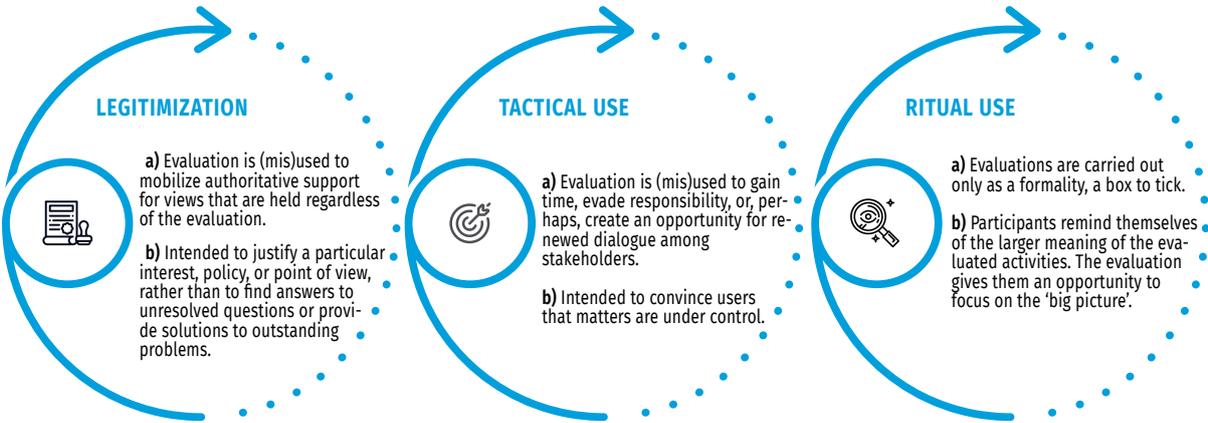


Figure 3: Categories and examples of misusing evaluation

While opportunities to reengage in a dialogue between stakeholders (tactical use of evaluation) or focusing stakeholders on the bigger picture of an intervention (ritual use) are valuable in themselves, the use of scarce evaluation resources for those purposes would be inappropriate. Instead, joint workshops with UNIDO and its partners or other engagement processes to seek a dialogue with stakeholders are recommended as part of the project intervention roll out and the monitoring process.

⁵ UNIDO, 2021, Evaluation Policy, page 4.
⁶ Molund, Stefan, and Göran Schill. 2007. *Looking Back, Moving Forward: Sida Evaluation Manual*. 2nd rev. ed. Stockholm: Swedish International Development Cooperation Agency (Sida), adapted from Vedung, E., 1997: *Public Policy and Program Evaluation*. New Brunswick: Transaction Publishers.

2. What is evaluation?

Section 2 defines evaluation and its principles. The role of evaluation in UNIDO's results-based management system is clarified, highlighting the differences between the monitoring and the evaluation function. The section also explains the types and dimensions of evaluations in UNIDO, key factors for success and how quality standards are ensured.

2.1 Definition

The UNIDO Evaluation Policy (2021), provides a definition of evaluation based on that set out by the United Nations Evaluation Group (UNEG, 2016).⁷

*Evaluation is an assessment that is as systematic and impartial as possible of a project, programme or entire strand of activities under a single thematic or institutional heading. An evaluation should provide evidence-based information that is credible, reliable and useful, thereby permitting the timely incorporation of findings, recommendations and lessons into the decision-making processes at the corporate, programme and project levels. Evaluations aim at determining the relevance, coherence, effectiveness, efficiency, impact and sustainability of UNIDO interventions and activities. They focus on the analysis of expected and achieved accomplishments, examining the results chain, processes, contextual factors and causality in order to ascertain the degree of achievement or the lack thereof.*⁸

The terminology used in the UNIDO Evaluation Manual is aligned to the Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC), 2002, *“Glossary of key terms in evaluation and results-based management,”* Paris, OECD

2.2 Evaluation and related concepts: role within results-based management

Evaluations are distinct from other forms of assessments, according to the UNIDO Evaluation Policy (2021). They differ from: **appraisals** (a critical assessment of the potential value of an undertaking before a decision is made to implement it); **monitoring** (management’s continuous examination of progress achieved during the implementation of a project or a programme in order to track compliance with the plan and to take necessary decisions to improve performance); **audits** (an assessment of the adequacy of management controls to ensure the economical and efficient use of resources, the safeguarding of assets, the reliability of financial and other information, the compliance with regulations, rules and established policies, the effectiveness of risk management, and the adequacy of organizational structures, systems and processes); and **research** (a systematic examination designed to develop or contribute to knowledge).

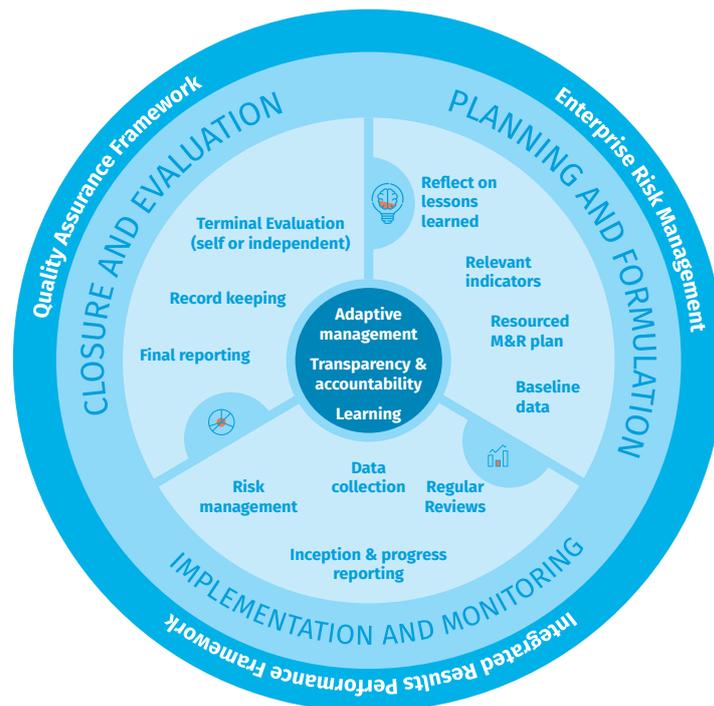
As the project cycle above indicates evaluations provide evidence for decision-making and feed into the project cycle management process. Thus, evaluation is not a decision-making process per se; however, it serves as an input to provide decision-makers with knowledge and evidence about performance and practice.

Findings and recommendations from evaluations will be used to inform strategies, thematic areas and internal processes, such as the project and programme identification and formulation process, as well as for learning and continuous improvement purposes.⁹

⁷ United Nations Evaluation Group (UNEG), 2016 UNEG Norms and Standards for Evaluation.

⁸ UNIDO, 2021, Evaluation Policy, page 3.

⁹ UNIDO, 2021, Evaluation Policy, pages 3- 4.



2.2.1 Results-based management

UNIDO defines RBM as “a broad management strategy aiming at improving management effectiveness and accountability by defining realistic expected results, monitoring progress toward their achievement, integrating lessons learned into management decisions and reporting”.¹⁰

As outlined in the forty-second session of UNIDO’s IDB in 2014 on the implementation of the Lima Declaration towards inclusive and sustainable industrial development (ISID), UNIDO has devoted great attention to implementing RBM-oriented strategies and exploring mechanisms to further enhance RBM coordination.¹¹

This shall be supported through demonstrating and providing evidence of the development impact across all its Technical Cooperation (TC) programmes and projects through an adequate UNIDO monitoring and reporting (M&R) system, as a core management responsibility. The evaluation function shall complement and provide assurance through its independent assessment, findings, recommendations, as well as lessons learned. The forty-fourth session of UNIDO’s IDB in November 2016 encouraged evaluations on results at outcome and impact levels, and the incorporation of information on performance and lessons into management and strategic planning processes.

Given that UNIDO started to link RBM more closely to the achievement of results from its 2010–2013 Medium-Term Programme Framework (MTPF) to the latest 2022–2025 framework, this highlights the strategic importance of evaluation in UNIDO. The forty-fourth session of the IDB unambiguously called for applying RBM for tracking UNIDO’s contribution to the 2030 Agenda and the Sustainable Development Goals (SDGs).¹²

¹⁰ UNIDO, 2007, “Results-based management and logframe approach. An introduction to the concept and to the way they are used in UNIDO.” (Draft)

¹¹ UNIDO, 2014, Industrial Development Board. Forty-second session. Vienna, 25–27 November 2014. Implementation of the recommendations relating to the management criteria contained in the “Strategic Guidance Document”. Report by the Director General (document IDB.42/16). (https://www.unido.org/fileadmin/user_media/PMO/IDB/IDB42/idb42_16e.pdf)

¹² UNIDO, 2016, Industrial Development Board. Forty-fourth session. Vienna, 22–24 November 2016. Report of the Industrial Development Board (document GC.17/2). (https://www.unido.org/sites/default/files/2016-12/GC_17_2_Report_of_the_IDB_44_e_0.pdf)

2.2.2 UNIDO Three Lines Model

The UNIDO Accountability Framework (Document DGB/2021/03) and the Internal Control Framework (Document DGB/2021/02) have adopted the “UNIDO Three Lines Model” (3LM) as the institutional approach for governance. The model aims to clarify the essential roles, responsibilities and lines of accountability within the Organization, including those governing risk management and internal controls.

The UN High-Level Committee on Management (HLCM), endorsed the 3LM as the reference “Risk Management, Oversight and Accountability Model for common positioning in the UN System with Governing Bodies” at its 28th session in 2014.

As shown in **Figure 4** below, evaluation is one of the 3rd line independent oversight functions in the Organization.

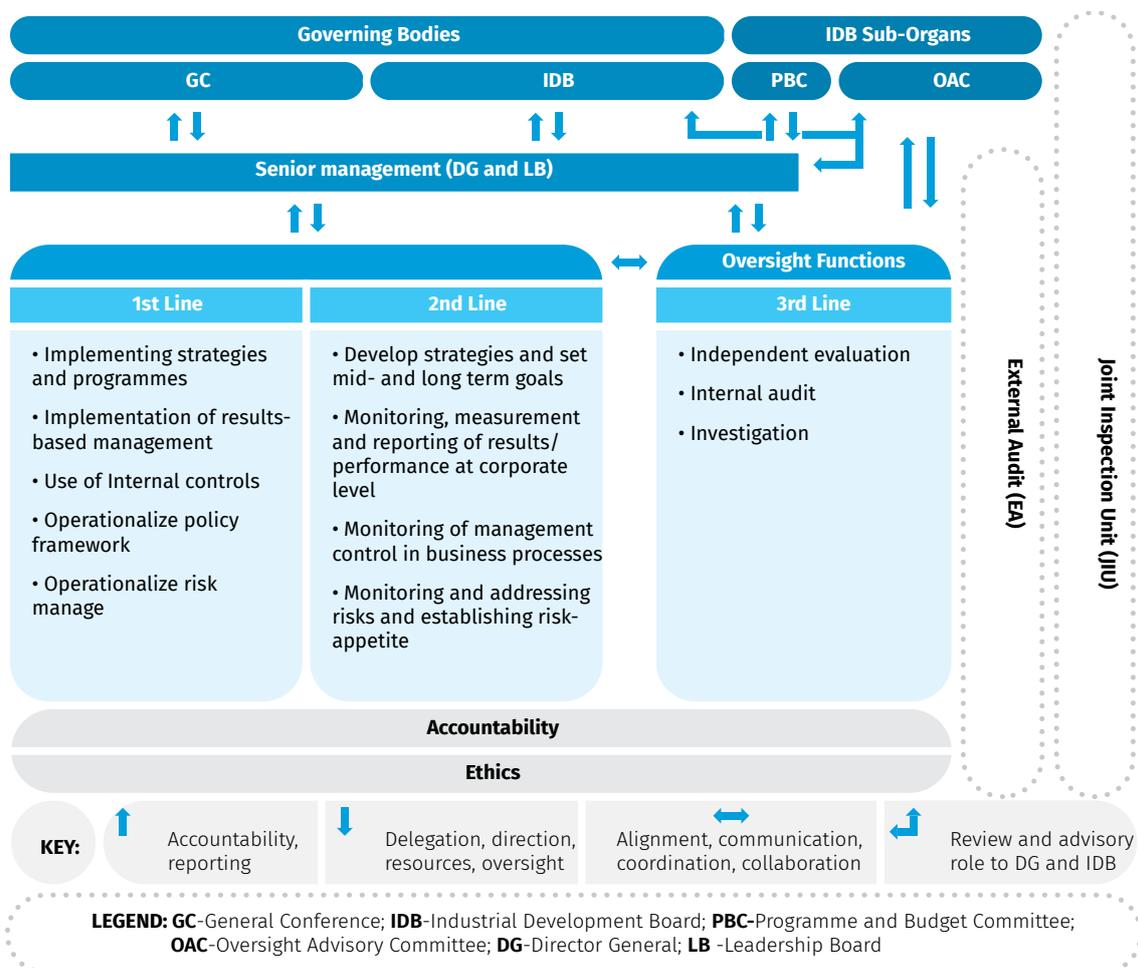


Figure 4: Three Lines Model (3LM)

2.2.3 Monitoring

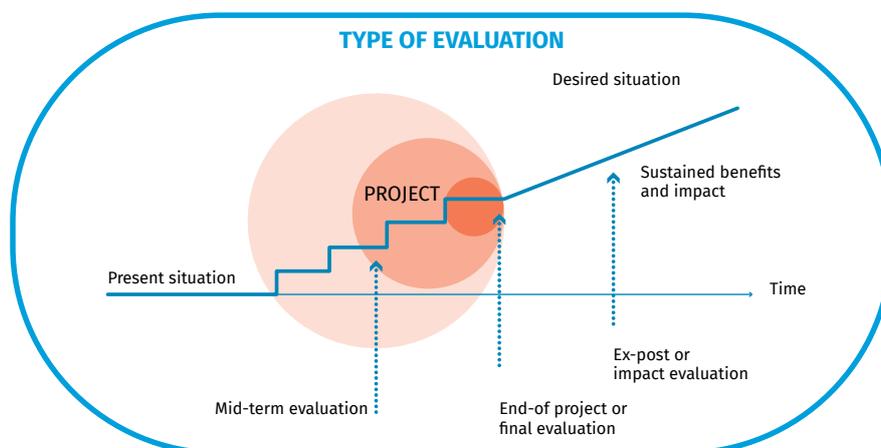
Monitoring (1st and 2nd Line function) and independent evaluation (3rd Line function) are different functions. Evaluation depends to a large extent on good monitoring. **Figure 5** highlights the main differences between these two functions.

MONITORING	EVALUATION
Continuous	Periodic
Management function	Independent from management
During implementation; it includes reporting and self-evaluations	Possible before (ex ante), during (e.g. mid-term), at the end (terminal/final), and after implementation (ex post)
Intervention objectives are given and used to “keep the ship on course” corrective action to improve performance	Objectives assessed about in terms of higher-level goals achieved or development problems solved; reconstruction of intervention logic
Pre-defined indicators of progress assumed to be appropriate	Validity and relevance of pre-defined indicators open to question
Focus on intended results	Identifies both unintended and intended results, positive and negative
Data routinely collected	Multiple sources of data and triangulation
Does not answer causal questions	Provides answers to causal question
Not independent	Independent

Figure 5: The concepts of monitoring and evaluation (M&E) in UNIDO
Source: UNEG, 2016, Sida, 2007, adapted.¹³

2.3 UNIDO evaluation types and dimensions

The UNIDO Evaluation Policy distinguishes two complementary types of evaluation in the Organization: self-evaluations and independent evaluations. Both the Evaluation Policy and the revised Evaluation Manual provide guidance for independent evaluations. However, the general principles of evaluation mentioned in the policy can also apply to self-evaluations,¹⁴ in the absence of specific guidance from UNIDO management on the subject.



¹³ Molund, Stefan, and Göran Schill. 2007. *Looking Back, Moving Forward: Sida Evaluation Manual*. 2nd rev. ed. Stockholm: Swedish International Development Cooperation Agency (Sida), adapted from Vedung, E., 1997: *Public Policy and Program Evaluation*. New Brunswick: Transaction Publishers

¹⁴ UNIDO, 2021, Evaluation Policy, page 5.

2.3.1 Self-evaluations

According to the UNIDO Evaluation Policy (2021), self-evaluations are an integral part of the project or programme M&R function, which is a management function. They take the form of a systematic, mid-term or final review of projects or programmes. As such, they are carried out or managed by officials who are responsible for their implementation, i.e., management. Independence is not a requirement for self-evaluations, although in keeping with good practice they are often undertaken by external evaluation consultants. Self-evaluations build upon M&R and should take place according to the rules established in project management guidelines. They are the vehicle for steering corrective action by line management, and therefore a management responsibility (under 1st and 2nd Line of the UNIDO 3LM).

2.3.2 Independent evaluations

Independent evaluations provide an independent, credible and evidence-based assessment on a given entity under evaluation, such as a project, programme, or an entire strand of activities under a thematic, geographical or institutional heading. Independent evaluations are conducted and/or managed by staff members of the UNIDO Independent Evaluation Unit and conducted by external independent evaluation consultants.

Figure 6 highlights the two main evaluation types within UNIDO, including their main characteristics, based on the UNIDO Evaluation Policy.

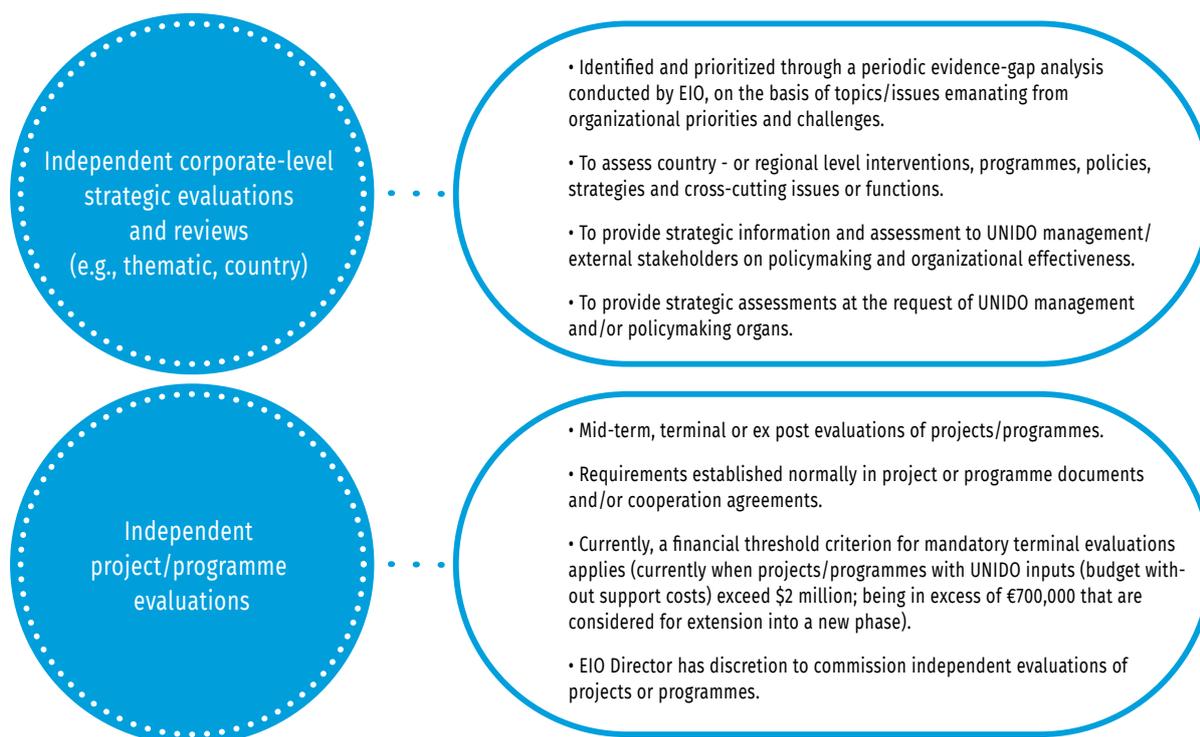


Figure 6: Types of independent evaluations in UNIDO
Source: UNIDO, 2021, Evaluation Policy, page 5.

2.4 Key factors behind successful evaluations

Experience suggests that a number of factors determine whether evaluations are influential in driving change in UNIDO,¹⁵ as shown in the **Box 1** below.

Box 1.

(a) The extent of consultation and inclusiveness – how far relevant stakeholders were engaged in the evaluation design, implementation and formulation of recommendations, with the ultimate aim of increasing the ownership of the evaluation process and evaluation results.

(b) The extent to which the work being assessed can be linked to a performance framework, logical framework or theory of change.

(c) The existence of baseline information, counterfactuals, targets, etc. to further strengthen the evidence base.

(d) The use of targeted evaluation questions. Realistically, an evaluation can answer about 15 core evaluation questions in depth to provide robust evidence for decision making.

(e) Length of the evaluation report (not exceeding 30 pages, excluding annexes).

(f) Recommendations are: Realistic (resource implications), Clear (what issue should be addressed), and Actionable (Who will be responsible, by when?). Do the recommendations make sense in the context of the Organization and for the programme/project or subject under evaluation?

(g) Management is engaged in the response process and commits to taking action Management action plans (MAPs) or management responses are established.

(h) Lessons learned drive organizational learning and innovation.

(i) The existence of a Recommendations/MAPs implementation tracking and reporting process.

(j) Whether a clear link can be drawn between evidence-based findings, conclusions, and recommendations, and lessons learned from evaluations.

2.5 Quality standards

The UNIDO Evaluation Quality Checklist aims to ensure UNIDO evaluations meet good quality standards. These standards are based on the UNEG Norms and Standards for Evaluation (2016), the UNEG Quality Checklist for Evaluation Terms of Reference and Inception Reports (2010)¹⁶ and the UNEG Checklist for Evaluation Reports (2010)¹⁷. They serve as concrete tools to transparently and consistently assure the high quality of UNIDO evaluations at different stages in the evaluation process.

For evaluation consultants it is good practice to use the checklists for evaluation inception reports and evaluation reports during the respective reporting periods to deliver a product of high quality. This should happen as early as at the first draft stages of the reports. The guidelines and checklists should be shared with the evaluation teams.

¹⁵ Based on: World Intellectual Property Organization (WIPO), Internal Oversight Division, 2016 Evaluation Manual, amended.

¹⁶ United Nations Evaluation Group, 2010: UNEG Checklist for Evaluation Terms of Reference and Inception Reports. Guidance document.

¹⁷ United Nations Evaluation Group, 2010, UNEG Checklist for Evaluation Reports. Guidance document.



2.6 Evaluation principles

10 EVALUATION NORMS

1. Internationally agreed principles, goals, and targets
2. Utility
3. Credibility
4. Independence (behavioural and organizational)
5. Impartiality
6. Ethics
7. Transparency
8. Human rights and gender equality
9. National evaluation capacities
10. Professionalism

The UNIDO Evaluation Policy comprises the following evaluation principles: evaluation ethics, involvement and consultation of stakeholders, impartiality and independence, evaluation skills, competencies, and professional standards.

UNEG has developed 10 evaluation norms:

- **Independence** (impartiality) - Independence is necessary for credibility, influences the ways in which an evaluation is used and allows evaluators to be impartial and free from undue pressure throughout the evaluation process. The independence of the evaluation function comprises two key aspects – behavioural independence and organizational independence. Behavioural independence entails the ability to evaluate without undue influence from any party. Evaluators must have the full freedom to conduct their evaluative work impartially, without the risk of negative effects on their career development, and must be able to freely express their assessment. The independence of the evaluation function underpins the free access to information that evaluators should have on the evaluation subject.
- **Credibility** - Evaluations must be credible. Credibility is grounded on independence, impartiality and a rigorous methodology. Key elements of credibility include transparent evaluation processes, inclusive approaches involving relevant stakeholders and robust quality assurance. Evaluation results (or findings) and recommendations are derived from – or informed by – the conscientious, explicit and judicious use of the best available, objective, reliable and valid data, and by accurate quantitative and qualitative analysis of evidence. Credibility requires that evaluations are ethically conducted and managed by evaluators that exhibit professional and cultural competencies.

- **Utility** - In commissioning and conducting an evaluation, there should be a clear intention to use the resulting analysis, conclusions or recommendations to inform decisions and actions. The utility of evaluation manifest through its use in making relevant and timely contributions to organizational learning, informed decision-making processes and accountability for results. Evaluations could also be used to contribute beyond the organization by generating knowledge and empowering stakeholders.

Evaluation ethics include taking due care of issues of confidentiality, sensitivity towards the social and cultural environment, discrimination, and wrongdoing.

The principle of involving and consulting stakeholders promotes management buy-in and ownership of the evaluation results, as well as supporting transparency; while the principle of impartiality and independence encourages objectivity, professional integrity, and absence of bias.

Finally, the principle of competencies and professional standards, relates to the required skills and experience for evaluation managers and evaluators, based on the “UNEG Evaluation Competencies Framework” and UNEG Norms and Standards for Evaluation, amongst others. The Info box below provides further insights into the evaluation principles.

**Section Info box:
evaluation principles**



UNIDO Evaluation Policy

3. Methodology

Section 3 describes the methodology behind the planning and execution of an evaluation in UNIDO. It includes elements such as evaluation criteria and questions, data collection and analysis, and the use of a rating system to assess different types of findings and conclusions, including cross-cutting issues. The learning dimension of evaluations, benchmarking, and stakeholder participation in the evaluation process and evaluation management are also outlined. The section closes by addressing how evaluations are used.

UNIDO evaluations align with the UNEG Norms and Standards for Evaluation, which state that “methodologies should be chosen with a clear intent to provide credible answers to the evaluation questions. The methodology should ensure that the information collected is valid, reliable and sufficient to meet the evaluation objectives and that the analysis is logically coherent and complete (and not speculative or opinion-based). Triangulation principles (utilizing multiple sources of data and methods) should be applied in order to validate findings”.¹⁸

3.1 Evaluation criteria

The UNIDO Evaluation Policy¹⁹ Annex 2 contains a full list of evaluation criteria and related evaluation questions which could be used by the evaluation teams to prepare an evaluation matrix. As with most United Nations specialized agencies, UNIDO uses internationally agreed, revised and updated evaluation criteria based on the OECD-DAC principles,²⁰ subsequently adapted to UNIDO’s needs, as shown below in **Figure 7**.

These evaluation criteria apply mainly to project and programme evaluations. For strategic or corporate-level evaluation (thematic, country-level, reviews), the main OECD-DAC criteria may apply (C1-C5 from **Figure 7**), and other criteria can be used and should be defined on a case-by-case basis.

¹⁸ UNEG, *UNEG 2016 Norms and Standards for Evaluation, Standard 4.2, page 23.*
¹⁹ UNIDO, 2021, Evaluation Policy, page 3.
²⁰ Organisation for Economic Co-operation in Europe (OECD) Development Assistance Committee (DAC), 2019: Better Criteria for Better Evaluation. Revised and updated evaluation criteria.

#	EVALUATION CRITERIA	DEFINITION	MAN-DATORY RATING
A	Progress to impact	Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended, including redirecting trajectories of transformational process and the extent to which conditions for trajectory change are being put into place.	Yes
B	Project design	Formulation of the intervention, the plan to achieve a specific purpose.	Yes
1	Overall design	Assessment of the design in general. (adequate description of the problem/need, stakeholders analysis, context, contribution to national programmes or other projects in the country).	Yes
2	Project results framework/ Logframe	Assessment of the logical framework to plan the intervention.	Yes
C	Project performance	Functioning of a development intervention.	Yes
1	Relevance	The extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor.	Yes
2	Coherence	The extent to which other interventions and policies support or undermine the intervention and vice-versa at global, country, sector or institutional level.	Yes

3	Effectiveness	The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.	Yes
4	Efficiency	The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.	Yes
5	Sustainability of benefits	The extent to which the net benefits of the intervention continue, or are likely to continue.	Yes
6	Impact	The extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.	Yes
D	Gender mainstreaming	The extent to which UNIDO interventions have contributed to better gender equality and gender related dimensions were considered in an intervention.	Yes
E	Project implementation management		Yes
1	Results-based management (RBM)	Assessment of issues related to results-based work planning, results-based M&E and whether decision-making on project/programme management are based on results.	Yes
2	Monitoring and reporting (M&R),	Refers to management's continuous examination of progress achieved during the implementation of a project or a programme in order to track progress of the plan and to take necessary decisions to improve performance (monitoring), how they are recorded and communicated to stakeholders (reporting), and whether desired results are being achieved and tracked.	Yes
F	Performance of partners	Assessment of partners' roles and responsibilities engaged in the intervention.	Yes
1	UNIDO	Assessment of the contribution of partners to project design, implementation, monitoring and reporting, supervision and backstopping and evaluation. The performance of each partner will be.	Yes
2	National counterpart		Yes
3	Implementing partner (if applicable)		Yes
4	Donor		Yes
G	Environmental and Social Safeguards (ESS), disability and human rights	Assessment of partners' roles and responsibilities engaged in the intervention.	Yes
1	Environmental safeguards	How UNIDO avoided or minimized/mitigated potential adverse impacts and risks	Yes
2	Social safeguards, disability and human rights	Assessment of social inclusiveness and human rights aspects, such as indigenous peoples' rights, land rights or other social and economic rights. Includes the assessment of disability-related vulnerabilities	Yes

Figure 7: Definition of evaluation criteria

The six evaluation criteria are perceived as a set of lenses through which interventions can be analyzed. Their use provides a number of potential advantages, including the ability:

- harmonize evaluation frameworks within the international evaluation community;
- employ consistent terminology which is comprehensible and used in the development field and beyond.
- offer guidance to differentiate between policy and operational levels in the evaluation.
- allow for comparison and learning across interventions; and
- align terms of reference and evaluation reports.

Even though the evaluation criteria offer certain advantages, some criticism has also been voiced. These are:

- The six evaluation criteria and the respective evaluation questions are applied in a more “blueprint modality (copy and paste)” without sufficient reflections about the purpose and utility of the evaluation.
- The evaluation criteria are perceived as a “corset” with too little flexibility.
- The same evaluation questions are frequently asked again, focusing on short-term project or programme results rather than on wider positive and negative impacts.
- The evaluation scope (number of criteria, evaluation questions) and its planned budget may not match.
- When interventions are more complex or volatile, the interconnections between inputs, outputs, outcomes, and impacts may not be as linear as expected. Therefore, using the evaluation criteria might result in missing important factors when analyzing “systems” in particular.

It is acknowledged that some issues covered by one criterion might overlap with others. However, to enable UNIDO to learn from the deeper evaluation analyses and lessons on a number of areas, separate criteria are included, such as those on monitoring and reporting and RBM. The consistent use of the criteria allows for comparability of UNIDO’s performance over time. The criteria provide a basis for the formulation of evaluation questions, as specified in the following section.

3.2 Formulating evaluation questions

The formulation of evaluation questions determines the breadth and depth of an evaluation. Questions should be clearly targeted, focusing on the purpose and objectives of the evaluation. Restricting questions to a limited number of key issues would ensure more targeted data collection and produce more in-depth analysis, resulting in an overall better evaluation product. It would also allow more efficient use of evaluation resources. UNIDO suggests the use of a maximum of 10 questions in its evaluation guidance.

“There are technical limitations that make it impossible to answer multiple questions or, more precisely, to provide quality answers to an excessive number of questions. This guide recommends a maximum of 10 questions.”

“Focusing an evaluation on a few key questions is all the more necessary when the intervention concerned is multidimensional and when the evaluation itself is multi-

²¹ European Commission (2005). Evaluation Questions. Methodological bases Evaluation process (How?) Evaluation questions Examples. https://europa.eu/capacity4dev/evaluation_guidelines/wiki/focus-evaluation-key-questions-0

dimensional. In that case, if one wanted to evaluate all the dimensions of the aid and all the dimensions of the evaluation, the work would be extremely costly or very superfluous. It is therefore necessary to make choices.”

Source: EC, 2005.

Additionally, when carrying out an evaluation, it is advisable to double check whether the answers to specific evaluation questions are already known or whether another ongoing study (audit, review, appraisal, or evaluation) is already working on providing the answer. Experience shows that the larger the number of evaluation questions, the less depth an evaluation can achieve. Hence a balance needs to be struck in identifying the most relevant key evaluation questions. Evaluators will further identify key issues and questions for the evaluation team to focus on in the inception report.

HOW TO SELECT THE EVALUATION QUESTIONS

It is good practice to follow a two-step process for developing evaluation questions.

A first set of evaluation questions should be based on:

- a) analysis of the intervention logic and rationale*
- b) issues justifying the launch of the evaluation*
- c) current known issues relating to the intervention*
- d) issues raised in previous evaluations of the intervention*

In the second step, normally during the evaluation inception phase, the questions are verified or fine-tuned by:

- e) issues raised by key informants at the start of the evaluation*
- f) expectations of members of a reference group*
- g) subjects raised by the evaluation team*

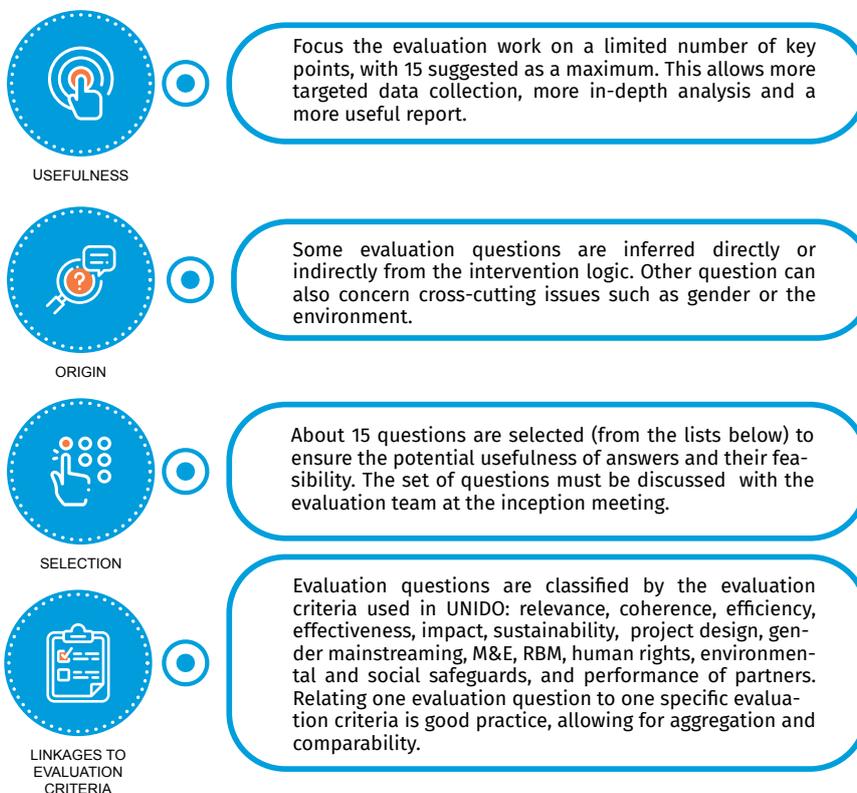


Figure 8: Good practices for the formulation of evaluation questions
Source, EU, 2006, amended.

The OECD provides a useful checklist of questions to help evaluation managers or team leaders to create more focused evaluation questions.²² It includes:

- If we could only ask one question about this intervention, what would it be?
- Which questions are best addressed through an evaluation, and which might be addressed through other means (such as a research project, evidence synthesis, monitoring exercise or facilitated learning process)?
- Are the available data sufficient to provide a satisfying answer to this question? If not, will better or more data be available later?
- Who has provided input to the list of questions? Are there any important perspectives missing?
- Do we have sufficient time and resources to adequately address all of the criteria/evaluation questions of interest, or will focusing the analysis on just some of the criteria/evaluation questions provide more valuable information?

A selection of examples of questions for some of the criteria is presented below. It is good practice to be selective and use the most pertinent evaluation questions rather than the entire sample of questions listed below. A full list of sample questions for all evaluation criteria stated in Section 3.1 is provided in Annex 2.

Relevance

- How does the project fulfil the needs and priorities of the target group?
- To what extent is the project aligned with the national development needs and priorities of the country (e.g. national development plans, industrial or sector development strategies)?
- How does the project reflect donor policies and priorities?
- To what extent does the project correspond to UNIDO's mandate and comparative advantages?
- Is the project a technically adequate solution to the development problem? Does it eliminate the cause of the problem? Is the design of the interventions/strategy and the targeting strategies adopted are consistent with the objectives?
- How realistic is the design in terms of suitability to the context and implementation capacity?
- Are the original project objectives (expected results) still valid and pertinent to the target groups? If not, have they been revised? Are the revised objectives still valid in today's context?

Coherence

- To what extent does the UNIDO intervention support or undermine other interventions (particularly policies), and vice versa?

Internal coherence

- What are the synergies and interlinkages between the intervention and other interventions carried out by the same institution/government?
- To what extent is the intervention consistent with the relevant international norms and standards to which that institution/government adheres?

External coherence

- To what extent is the intervention consistent with other actors' interventions in the same context (complementarity, harmonization and coordination with others)?
- How is the intervention adding value while avoiding duplication of effort?
- How economically are the project resources/inputs (concerning funding, expertise, time...) being used to produce results?
- Are the results being achieved at an acceptable cost? Would alternative approaches accomplish the same results at less cost? Could more have been achieved with the same input? Could the same have been achieved with less input?
- What has been the relationship and ratio of costs and benefits (e.g. net present value, internal rate of return)? How does it compare with similar interventions (if the comparison is plausible)? Are the unit costs of specific interventions consistent with recognized practices and the results achieved?
- To what extent were expected results achieved within the original budget and time frame? If no, please explain why.
- Is the timeframe of the intervention design and implementation justifiable, considering the results achieved, the specific context and unforeseeable events?
- What measures have been taken during planning and implementation to ensure that resources are efficiently used? Were the project expenditures in line with budgets?
- How timely was the project in producing outputs and outcomes? Comment on the delay or acceleration of the project's implementation period.
- To what extent were the project's activities in line with the schedule of activities as defined by the project team and annual work plans?
- Are project management cost ratios justifiable in terms of intervention objectives and results achieved, considering contextual aspects and unforeseeable events?
- Have the inputs from the donor, UNIDO and government/counterpart been provided as planned, and were they adequate to meet the requirements?

Effectiveness

- What are the main results (mainly outputs and outcomes) of the project? What have been the quantifiable results of the project?
- To what extent did the project achieve its objectives (outputs and outcomes), against the original/ revised target(s)?
- What are the reasons for the achievement/ non-achievement of the project objectives?
- What is the quality of the results? How do the stakeholders perceive them? What is the feedback of the beneficiaries and the stakeholders on the project effectiveness?
- To what extent is the identified progress result of the project attributable to the intervention rather than to external factors?
- What can be done to make the project more effective?
- Were the right target groups reached?

Progress to impact

- What difference has the project made to the beneficiaries?
- To what extent has the intervention had the anticipated impact on the target group, institutions and policies? Why?
- What are the observed changes in the target group, institutions and policies in terms of advancing economic competitiveness, safeguarding the environment and creating shared prosperity (social inclusiveness)?
- What kind of economic, social and environmental impact has the project helped bring about? The changes could be further broken down to consist of the following elements:

ECONOMIC IMPACT	ENVIRONMENT IMPACT	SOCIAL IMPACT
<ul style="list-style-type: none"> • Improved technologies, practices, concepts, techniques, approaches • Improved internal process and practices (production, management...) • Better economic performance of existing business (saving, higher revenue, market access...) • Income (improved/worsened) • Development of viable business models 	<ul style="list-style-type: none"> • Reduction in environmental threats • Shifts towards low-emission and resilient development path • Phase out, disposal and reduction of POPs, mercury and other chemicals 	<ul style="list-style-type: none"> • Gender equality • Human rights • Employment • Health and safety (reduced exposure to environmental risks and better working condition) • Access of basic services/capital • Opportunities for more sources of income

- What explains those changes? What are the challenges?
- What are the changes attributable to the project? To what extent?
- Were there any unintended effects (positive or negative)?
- From an equity standpoint, have very poor/marginalized groups and/or special categories substantially benefited?
- To what extent is the project contributing to the long-term transformation (objective/goal of the project/programme)? To what extent the project puts in place mechanisms to support the durability and expansion of enabling conditions after the project ends.
- To what extent the project contributions have been mainstreamed, replicated and scaled up?
 - **Mainstreaming:** To what extent has the project been incorporated into broader stakeholder mandates and initiatives such as laws, policies, regulations and other projects?
 - **Replication:** To what extent are the project's specific results or initiatives (e.g. methodology, technology or lessons learned) reproduced or adopted at a comparable administrative, geopolitical, or ecological scale, often in another geographical area or region?
 - **Scaling-up:** To what extent are the project's initiatives and results implemented at larger geographical, economic or ecological scale which are likely to feedback and

further strengthen lower level enabling conditions to the desired development trajectory scale?

Sustainability

- Will the project results and benefits be sustained after the end of donor funding (including technical capacities)?
- Does the project have an exit strategy?
- To what extent have the outputs and results been institutionalized and ownership ensured?

Financial risks:

- To what extent was the project able to diversify founding sources?
- To what extent are funding streams stable?
- What is the likelihood of financial and economic resources not being available once the project ends?

Socio-political risks:

- Are there any social or political risks that may jeopardize the sustainability of project outcomes?
- What is the risk that the level of stakeholder ownership (including ownership by governments and other key stakeholders) will be insufficient to allow for the project outcomes/benefits to be sustained?
- Do the various key stakeholders see that it is in their interest that project benefits continue to flow?
- Is there sufficient public/stakeholder awareness in support of the project's long-term objectives?

Institutional framework and governance risks:

- Do the legal frameworks, policies, and governance structures and processes within which the project operates pose risks that may jeopardize the sustainability of project benefits?
- Are requisite systems for accountability and transparency and required technical know-how in place?

Environmental risks:

- Are there any environmental risks that may jeopardize the sustainability of project outcomes?
- Are there any project outputs or higher-level results that are likely to have adverse environmental impacts, which, in turn, might affect the sustainability of project benefits?

3.3 Rating systems

UNIDO introduced a six-point rating system for the evaluation criteria in 2015, in line with the practice adopted by other development agencies. The aim of the system is to quantify the judgment of evaluators, identify good and poor practices, to facilitate aggregation within and across projects and enable tracking performance trends over a period. The six point rating system, with six (6) representing the best and one (1) the worst score, allows for nuanced assessment of all evaluation criteria. The same rating scale is used for all rating areas as shown in **Figure 9** below.

SCORE	DEFINITION*
6 HIGHLY SATISFACTORY	Level of achievement presents no shortcomings (over 90% achievement rate of planned expectations and targets).
5 SATISFACTORY	Level of achievement presents minor shortcomings (70% - 89% achievement rate of planned expectations and targets).
4 MODERATELY SATISFACTORY	Level of achievement presents moderate shortcomings (50% - 69% achievement rate of planned expectations and targets).
3 MODERATELY UNSATISFACTORY	Level of achievement presents some significant shortcomings (30% - 49% achievement rate of planned expectations and targets).
2 UNSATISFACTORY	Level of achievement presents major shortcomings (10% - 29% achievement rate of planned expectations and targets).
1 HIGHLY UNSATISFACTORY	Level of achievement presents severe shortcomings (0% - 9% achievement rate of planned expectations and targets).

Figure 9 : UNIDO evaluation rating scale

Note: * For impact, the assessment will be based on the level of likely achievement, as it is often too early to assess the long-term impacts of the project at the project completion point.

Figure 10 below contains the formula applied to transform the results of UNIDO’s six-point rating scale to the Global Environment Facility (GEF)’s four-point scale for sustainability.²³

UNIDO rating	UNIDO rating: sustainability	GEF rating: sustainability
6	Highly likely	Likely
5	Likely	Moderately likely
4	Moderately likely	Moderately likely
3	Moderately unlikely	Moderately unlikely
2	Unlikely	Moderately unlikely
1	Highly unlikely	Unlikely

Figure 10: Formula transforming UNIDO ratings into GEF ratings

3.3.1 Project design

Criteria for rating project design are related to the logical framework (logframe) approach and the overall quality of the project design. These criteria include:

Overall design quality

- Alignment with and clear contribution to country priorities, the needs of target groups and UNIDO and donor strategies
- Consideration and use of lessons and evaluative evidence from other projects
- Adequacy of stakeholders’ analysis and engagement
- Consideration of and alignment with other projects, programmes or actions that contribute to the same priorities
- Technical feasibility and validity of project design
- Budgeted M&R and evaluation plans with clear timelines, roles, and responsibilities
- Adequacy of risk assessment (for example financial, sociopolitical, institutional, environmental and implementation aspects)

Quality of the project’s logical framework or theory of change (TOC)

- Clarity and logic of results chain, including impacts, outcomes and outputs
- Description of external factors (assumptions/risks) that can be monitored
- Evidence that demonstrates validity of pathways (cause-effect links) suggested in the TOC
- Logical framework, including SMART indicators, targets and means of verification

²³ GEF uses a four-point scale for the criterion of sustainability.

3.3.2 Project performance

Project performance corresponds broadly to DAC criteria and should be customized according to the context of the intervention to be evaluated.

- Relevance
- Coherence
- Effectiveness
- Efficiency
- Progress to impact
- Sustainability of benefits

3.3.3 Performance of partners

UNIDO's projects are characterized by a group of main partners with specific roles and responsibilities. UNIDO itself acts as project implementer and supervisor. Though supplemented by implementation performance criteria listed above, the criteria to assess UNIDO as a partner are more specific and help to address frequent issues in its performance. Governments are local executors and owners of the project and donors provide project funding. Rating the partners is therefore a key part of UNIDO project/programme evaluations.²⁴ The six-point rating scale applies.²⁵

The key issues to be addressed to rate UNIDO's performance are:

Support to project design

- Mobilization of adequate technical expertise for project design
- Inclusiveness of project design (with national counterparts)
- Previous evaluative evidence shaping project design
- Planning and budgeting sufficiently for monitoring, reporting and evaluation

Implementation

- Timely recruitment of project staff
- Project modifications following changes in context or after the Mid-Term Review
- Timely follow-up to address implementation bottlenecks
- Role of UNIDO country presence (if applicable) in supporting the project
- Convening role and support to government engagement in policy dialogue to ensure up-scaling of innovations
- Supporting the project coordination function
- Exit strategy, planned together with the government
- Overall effectiveness of project management as outlined in the project

²⁴ As practiced by the World Bank and the International Fund for Agriculture Development.

²⁵ 6 = Highly satisfactory; 5 = Satisfactory; 4 = Moderately satisfactory; 3 = Moderately unsatisfactory; 2 = Unsatisfactory; 1 = Highly unsatisfactory.

- document
- Project's governance system
- UNIDO HQ-based management, coordination, monitoring, quality control and technical input

In the case where UNIDO is not responsible for project/programme implementation, the implementing partner/s are assessed based on the criteria above.

To assess the performance of national counterpart institutions, the evaluation looks into the following issues:

Project/programme design

- Ownership and commitment to designing the project/programme
- Ensuring alignment to national development priorities

Implementation

- Ownership of the project/programme
- Financial contributions (cash or in-kind)
- Support to the project/programme, based on actions and policies
- Counterpart funding
- Internal government coordination
- Exit strategy, planned together with UNIDO, or arrangements for continued funding of certain activities
- Facilitation of the participation of non-governmental organizations (NGOs), civil society, academia and the private sector where appropriate
- Suitable procurement procedures for timely project/programme implementation
- Engagement with UNIDO in policy dialogue to promote the up-scaling or replication of innovations

For the assessment of donor performance, the following issues should be included:

- Level of engagement and contribution to the project/programme design and/or during the implementation
- Timely disbursement of project/programme funds
- Donor monitoring and feedback to progress reports
- Support by the donor's country presence (if applicable) to the project/programme, for example through engagement in policy dialogue

3.3.4 Human rights and gender mainstreaming

The universally recognized values and principles of human rights and gender equality

²⁶ UNEG 2014: "Integrating Human Rights and Gender Equality in Evaluations" at: <http://www.unevaluation.org/document/detail/1616>

should be integrated into all stages of an evaluation, underpinning the commitment to the principle of ‘No One Left Behind’. All evaluations must therefore take gender mainstreaming, human rights and disability concerns into consideration.

Each evaluation should take a gender-responsive approach, even when assessing projects that were not designed as gender-related. The UNEG guidance document, “Integrating Human Rights and Gender Equality in Evaluations”²⁶ provides examples of how to incorporate these elements into the purpose, objectives, context and scope of the evaluation, and to incorporate a gender dimension into the standard evaluation criteria.

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

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

A gender-responsive evaluation promotes accountability on commitments to gender equality, human rights and women’s empowerment by providing information on the way in which development programmes affect women and men differently and how they contribute towards the achievement of those commitments. It is applicable to all types of development programming, not just gender-specific work²⁷

A gender-responsive evaluation should specify how gender issues are addressed in the methodology, including how data collection and analysis methods integrate gender considerations and ensure that data collected are disaggregated by sex.

Such methodology should employ a mixed-methods approach, appropriate to evaluating gender equality and women’s empowerment considerations. The evaluation methods and sampling frame should also address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate.

The current UNIDO policy on gender equality and the empowerment of women, (Document UNIDO/DGB/2019/16) provides overall guidelines for establishing a gender mainstreaming strategy and action plan to guide the process of addressing gender issues in UNIDO interventions. The Strategy for Gender Equality and the Empowerment of Women 2020–2023 published in 2019 states that evaluation and internal oversight in UNIDO will do the following:

“Coordinate the establishment of common tools and approaches for integrating gender equality and human rights-based approaches in evaluation and audit, conduct gender-related evaluations and gender audits in line with the four-yearly strategies for GEEW, and take part in United Nations Evaluation Group (UNEG) UN-SWAP Evaluation Performance Indicator-related working groups and peer learning exchanges.”

Source: UNIDO, 2019: Policy on Gender Equality and the Empowerment of Women 2020-2023, page 8.

²⁷ See UN Women Independent Evaluation Office, 2015, How to Manage Gender-Responsive Evaluation. Evaluation Handbook <https://www.unwomen.org/en/digital-library/publications/2015/4/un-women-evaluation-handbook-how-to-manage-gender-responsive-evaluation>.

In line with this strategy, all UNIDO technical assistance projects have been assigned a gender marker, with project designs screened based on a gender mainstreaming checklist before gaining approval. UNIDO’s gender marker is in line with UN System-wide Action Plan (UN-SWAP) requirements, comprising four categories: 0 – no attention to gender; 1 – some/limited attention to gender; 2a – significant attention to gender; and 2b – gender is the principal objective.

The UNIDO Strategy for Gender Equality and the Empowerment of Women (2020–2023) foresees the following role for evaluation:

Excerpt from the UNIDO Gender Strategy 2020-2023

<p>STRATEGIC OBJECTIVE 2: Strengthen UNIDO institutional capacity and effectiveness to enhance the delivery of results on gender equality and the empowerment of women.</p>	
<p>PRIORITY ACTION AREA 7: Gender-responsive evaluation and audit (UN-SWAP performance indicators 4 and 5)</p>	<p>Gender-responsive oversight activities, such as evaluation and audit functions, provide much-needed quality assurance and accountability for the Organization’s gender equality and women’s empowerment results. UNIDO commits to “exceeding” the UN-SWAP 2.0 requirements by 2023, specifically through the following actions:</p> <ul style="list-style-type: none"> • The Office of Evaluation and Internal Oversight will ensure continued improvement in integrating gender and human rights-based approaches throughout the evaluation cycle, particularly in terms of mainstreaming gender in findings, lessons learned and recommendations in evaluation reports. • The Office of Evaluation and Internal Oversight will require specific experience/knowledge on gender integration issues in the profiles of the evaluation team members where appropriate.

To assess gender mainstreaming, terminal evaluations should assess if the gender marker is correctly assigned at entry, if appropriate outputs/activities/indicators and monitoring are put in place during implementation, and what results can actually be observed at the time of the terminal evaluation (in line with UNIDO’s organizational results reporting to SWAP). The gender mainstreaming six-point rating scale should then be used accordingly. Together with the budget, the time required to adequately carry out a gender-responsive evaluation will need to be taken into account.

For projects and programmes awarded 2a or 2b ratings at the design stage, at least one member of the evaluation team should have experience in evaluating GEEW projects. For other projects/programmes, evaluators are encouraged to further familiarize themselves with the UNIDO policy framework and guidelines related to gender mainstreaming.

3.3.5 Environmental and social safeguards (ESS)

As part of its commitment to ISID and to achieving the 2030 Agenda for Sustainable Development, UNIDO has long supported environmentally sound and sustainable development in the full range of its project activities. UNIDO first introduced a set of Environmental

and Social Safeguard Policies and Procedures (ESSPP) in January 2015 to projects and programmes funded by the GEF and Green Climate Fund (GCF). In July 2017 and again in June 2021, the ESSPP was revised and has since been applied to all UNIDO projects and programmes submitted to the Adaptation Fund (AF), GEF, GCF, and all other UNIDO projects and programmes when relevant. The ESSPP draws on the safeguard requirements and policies of key funding partners, as well as on the guidance provided by the United Nations Environment Management Group, of which UNIDO is a member.

The main purpose of the ESSPP is to ensure that UNIDO projects and programmes comply with the Organization's safeguards by assessing environmental, climate change and social risks and impacts as early as possible in the project cycle. This is achieved through disclosure of relevant information and effective monitoring and supervision of agreed environmental and social management and mitigation measures during project implementation. As such, ESS policies and procedures have been systematically mainstreamed into UNIDO's project cycle through the application of an environmental and social screening and assessment procedure from project programming and, formulation to implementation, monitoring and evaluation.

UNIDO is committed to full compliance with the 12 Operational Safeguards outlined in the **Box 2** below.

Box 2. UNIDO's 12 Operational Safeguards

- 1. Environmental and Social Assessment Safeguard:** ensures that projects are environmentally and socially sound and sustainable by providing an overarching framework for the required screening and assessment processes that all UNIDO projects undergo.
- 2. Protection of Natural Habitats and Biodiversity Safeguard:** ensures that biological diversity is conserved, and that sustainable use of natural resources is promoted.
- 3. Involuntary Resettlement and Land Acquisition Safeguard:** ensures that UNIDO projects do not result in involuntary resettlement or physical and/or economic displacement of anyone through project-related land use.
- 4. Indigenous Peoples Safeguard:** ensures that projects foster full respect for indigenous peoples and their dignity, human rights, and cultural uniqueness.
- 5. Pest Management Safeguard:** ensures that environmental and health risks associated with the use of pesticide are minimized and managed, and that safe, effective, and environmentally sound pest management is promoted and supported.
- 6. Cultural Heritage Safeguard:** ensures that UNIDO does not engage in any projects that adversely impact upon tangible and intangible cultural heritage.
- 7. Safety of Dams Safeguard:** ensures compliance with relevant international quality and safety standards in the design, construction, operation, and maintenance of new dams and the rehabilitation of existing dams.
- 8. Labour and Working Conditions Safeguard:** ensures that the pursuit of poverty reduction and economic growth through employment creation and income generation should be accompanied by protection of the fundamental rights of workers.

9. Resource Efficiency and Pollution Prevention Safeguard: ensures that a project-level approach to resource efficiency, cleaner production processes and pollution management in line with internationally disseminated technologies and practices is adopted.

10. Community Health, Safety and Security Safeguard: recognizes that project activities, equipment, and infrastructure can increase community exposure to risks and impacts, and ensures that the health, safety, and security risks and impacts on project-affected communities are avoided or minimized

11. Information Disclosure and Stakeholder Consultation Safeguard: ensures that UNIDO's approach to information disclosure is guided by openness and open and transparent consultation with project stakeholders.

12. Accountability and Grievance System: ensures that a mechanism for dispute resolution and for ensuring accountability and compliance with UNIDO's environmental and social safeguards exists at UNIDO and, if required, also at project/programme level.

The ESSPP requires that projects and programmes undergo environmental and social (E&S) risk assessments before being approved. This helps UNIDO decide whether a project or programme should be supported and, if so, the way in which environmental and social issues should be addressed in its development and implementation.

The objective of the E&S screening is to help UNIDO project development teams to: (i) determine the appropriate environmental and social risk category of a proposed UNIDO project or programme, and (ii) assess whether project or programme activities pose any specific risk of breaching the safeguards.

To incorporate environmental and social considerations into evaluations, these two dimensions should be considered during the development of the evaluation ToR and the subsequent evaluation process. It is important to state that only safeguards relevant to the intervention under evaluation should be taken into consideration in order to maintain a focused assessment.

For projects that are in category²⁸ A and B at project design, at least one evaluation team member should have sufficient expertise in environmental or social areas to appropriately assess their impact. For projects in category C, evaluators are encouraged to further familiarize themselves with the environmental and social safeguards from the UNIDO ESSPP.²⁹ **Box 3** below displays a list of key guiding questions to that will help evaluation teams to assess environmental and social considerations. In keeping with good practice, evalua-

²⁸ Category A: A proposed project is classified as Category A if it is likely to induce significant and/or irreversible adverse environmental and/or social impacts that are sensitive, diverse, or unprecedented, or that affect an area broader than the sites or facilities subject to physical works, i.e. the geographical scope of the project.

Category B: Category B projects often differ from Category A projects of the same type only in scale. They are likely to have fewer adverse impacts on human populations or environmentally important areas than Category A projects. Likely impacts will be few in number, site-specific, and few, if any, will be irreversible. In most cases impacts can be readily minimized by applying appropriate management and mitigation measures or incorporating internationally recognized design criteria and standards.

Category C: A proposed project is classified as Category C if it is likely to have either minimal or no adverse social and/or environmental impacts.

²⁹ https://www.unido.org/sites/default/files/files/2021-07/AI_2021_03_UNIDO_ENVIRONMENTAL_AND_SOCIAL_SAFEGUARDS.pdf

tors should ensure that questions selected areas targeted and relevant as possible.

Box 3. Guiding key evaluation questions to address ESS

ToR and inception phase

1. Was the project/programme proposal screened against the environmental and social safeguards based on UNIDO ESSPP during the preparation phase?
 - a. If no screening was conducted, were there any environmental considerations that should have been subjected to screening? If so, what were they and how can they be evaluated?
 - b. If screening was conducted, what environmental considerations were raised and how were they addressed during implementation?
2. What E&S category was assigned to the project?

During the evaluation process

1. To what extent were environmental and social safeguards applied during project preparation and implementation?
2. To what extent has the intervention's theory of change included environmental and social aspects?
3. To what extent has the project tracked relevant E&S risks and applied appropriate E&S safeguards?
4. How have these processes affected the achievement of project results?
5. Has the intervention caused any harm to the environment or to any stakeholders?
6. Has the project affected marginalized communities, including indigenous peoples?
7. To what extent has the intervention contributed towards environmental and social effects (positive or negative) in the following areas?
 - biodiversity conservation and sustainable natural resource management
 - climate change and disaster risk
 - pollution prevention and resource efficiency
 - community health, safety and security
 - displacement and resettlement
 - indigenous people

SECTION INFO BOX:



Further reading about environmental and social safeguards, including social inclusion/human rights considerations:

UNIDO, 2021: The UNIDO environmental and social safeguards policies and procedures (ESSPP)

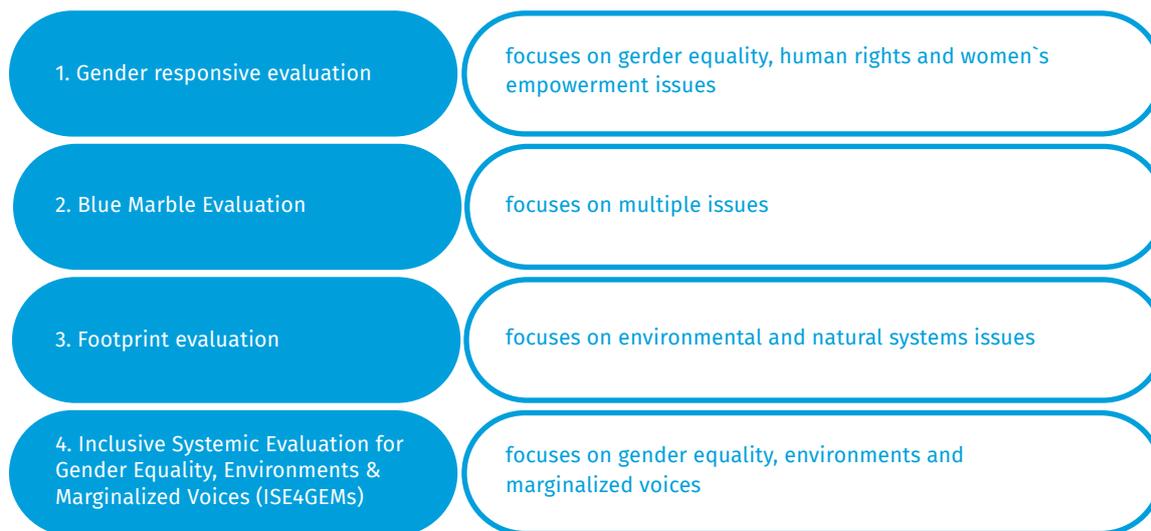
Further reading on the evaluation of disability:

UNEG, 2022: Guidance on Integrating Disability Inclusion in Evaluations and Reporting on the UNDIS Entity Accountability Framework Evaluation Indicator.

3.3.6 Other evaluation approaches

The UNIDO Evaluation Policy, UNEG Norms and Standards for Evaluation, the six OECD/DAC criteria and other cross-cutting criteria primarily guide evaluations in UNIDO. Nonetheless, the UNIDO Office of Evaluation and Internal Oversight recognizes new evaluation practices that have evolved since the launch of the 2030 Agenda for Sustainable Development. More recent evaluation approaches reflect the need for major and rapid system transformations and include the notions of transformative change and complex systems thinking in their assessments. **Box 4** below provides an example on related concepts.

Box 4.



Assessing the contribution to transformational changes.

Blue Marble evaluation takes its name from the first photograph of the whole Earth taken from space in 1972 and epitomizes the need for a global consciousness. The Blue Marble evaluation approach means looking beyond political and geographical borders and across sectors, as well as connecting the global and the local. It puts evaluation in the context of global crises, such as pandemics, climate emergencies, social unrest and other interrelated challenges to the Earth and humankind. Against this backdrop, Blue Marble evaluation aims to transform evaluation to evaluate transformation, with the ultimate objective of strengthening the contribution that evaluation can make to address universal challenges. It calls for reframing traditional evaluation concerns about attribution and accountability and distances itself from the traditional six evaluation criteria. Blue Marble evaluation offers new principles, such as global thinking or the transformative engagement principle, and alternative criteria (e.g. resilient and adaptive sustainability) to drive and contribute to global systems transformation. The premise of Blue Marble is that “evaluation must be transformed if it is to be part of the solution rather than part of the problem”.

In 2022, UNIDO conducted an independent strategic evaluation on the Organization's capacity to contribute to transformational change. The idea behind this formative evaluation was, inter alia, to trigger new strategic thinking around UNIDO's development cooperation. The key entry point for this assessment was the systemic nature of development interventions. In this context, the evaluation team developed a tailor-made set of review criteria for the thematic approaches currently used in UNIDO. This strategic evaluation highlights the sig-

nificance of understanding complex and dynamic systems and recognizes that evaluations must be flexible and adaptable.

The current revised UNIDO *Evaluation Manual* aims to reflect advances in evaluation practices and seeks to incorporate new perspectives, elements and methodologies into those practices, demonstrating that UNIDO is responsive, progressive and innovative in its approach to evaluation.

The following evaluation questions are examples of questions relevant to assessments of processes and/or interventions designed to contribute to transformational change:

- To what extent are the different UNIDO delivery modalities [such as Country Programmes (CPs), Programmes for Country Partnership (PCPs), project and thematic approaches] systemic and conducive to transformational change?
- What are the key assumptions and risks that influence progress towards advancing the transformational impact of UNIDO's work?
- What are the strengths and weaknesses of UNIDO in contributing to transformational change?
- Which are good international practices and lessons that UNIDO could learn to enhance contributions to transformational change?
- What needs to be improved to enhance UNIDO's contributions to transformational change?

3.3.7 Logframe

The rationale behind an intervention should be part of the project document. In order to identify the aims of any intervention, UNIDO uses the logframe planning tool, which contains indicators, baselines and targets that are designed to provide details on project objectives and outcomes. For thematic or country evaluations, the intervention rationale could be included in the related strategy document.

A verification of the intervention logic is advisable or, if missing from planning documents, it should be reconstructed. For this purpose, theory of change can be used.

Why is it important to have an intervention logic?

- It helps staff and consultants to gain a common understanding of how the intervention is supposed to work and how the main components relate to one another.
- It helps to communicate the intervention to outsiders in a concise and compelling way.
- It helps to make the changes and results you expect to see more explicit.
- It helps to define and collect data for monitoring, review and evaluation purposes.

3.3.8 Theory of change

A theory of change describes the mechanisms through which change is expected to occur under particular interventions. It shows pathways to changes in a narrative and graphic

³⁰ Vogel, I. Department for International Development (DFID), 2012: "Review of the use of "Theory of Change" in international development." Review report, page 44.

form. In this context pathways refer to the “sequence or hierarchy of changes and events that map out how things will change”.³⁰

“A project’s theory of change provides a basis for evaluation of the theory and results. The terminal evaluation report will include a description of the project’s theory of change including description of: the outputs, outcomes, intermediate states, and intended long-term environmental impacts of the project; the causal pathways for long-term impacts; and, the implicit and explicit assumptions. The project’s objective(s) should also be included within the theory of change.”

Source: Global Environment Facility (GEF), 2017, Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects.

This is to be tested in the evaluation process. The added value of a TOC is that it further elaborates the often implicit assumptions behind the intervention and makes them explicit. Those assumptions can be related for example to the expected behaviour of a target group to benefit from the results of an industrial development intervention.

TOC depicts the causal and transformational pathways from project outputs to outcomes and longer-term impacts. It also identifies the drivers and barriers to achieving results. Learning from this analysis will be useful for the design of future projects so that the management team can effectively use the theory of change to manage the project based on results. In this respect, TOC complements the logframe approach, which tends to exclude context regarding barriers and drivers of change. Yet understanding the why, how, when and under what conditions specific actions lead to certain results is important to correctly interpret the evaluation results.

What is a theory of change?

- A theory of change is a comprehensive description and illustration of how and why a desired change is expected to happen in a particular context
- It is a planning, monitoring and evaluation tool

When is it developed?

- As a planning tool, the TOC is developed during the design phase
- As an evaluation tool, it is reviewed or reconstructed during the evaluation phase

How does it help?

- Specifies how an intervention leads to results and impact (behaviour change)
- Presents pathways to change in a narrative and graphic form
- Elaborates implicit assumptions behind the intervention, making them more explicit
- Articulates linkages between outputs, outcomes, and impacts
- Identifies barriers to achieving the objectives, and highlights pathways and drivers of change

The theory of change should therefore be complemented by a logframe to ensure that indicators and targets are specified and, if possible, quantified. Experience shows that “using

³¹ Ibid., page 43.

theory of change thinking to bring ‘evaluative thinking’ into an (intervention) at an early stage is one of the key benefits of working with the approach. It helps to identify progress markers, and where focused evaluation questions can provide insights as a (intervention) is implemented”.³¹

What to do if the UNIDO intervention does not benefit from a theory of change?

If no theory of change for the intervention’s evaluation is available, the evaluators could reconstruct such a theory using the planning documentation, such as the project or programme document. Information gathered from a logframe matrix could also help to reconstruct a theory of change.

To pose the right questions, the logic behind the intervention should be clarified, with goals and objectives (impact and outcomes), indicators and assumptions known before the evaluation starts. However, in reality, the evaluation process itself often serves the purpose of reconstructing the rationale for the intervention. This puts the evaluation team in a situation where difficult choices have to be made. The evaluation team needs to balance the time required to verify or reconstruct the intervention logic before the start of an evaluation with the time required for the same task during an evaluation, which might risk widening the scope of the evaluation by introducing new evaluation questions. During the evaluation process, validating the theory of change with key stakeholders would enrich the evaluation team’s understanding and assessment of an intervention and help stakeholders assess the contribution of the intervention to transformational change of the complex system within which the projects and programmes operate.

There is a wide range of M&E methodologies that are “valid for building an evidenced case, qualitative as well as quantitative. What links them all is the importance of having a theory of change that lays out the expected story in advance of the changes happening. This then provides the basis for collecting evidence, checking other possible explanations as counterfactuals and presenting a case from which cause can be reasonably inferred and linked back to the programme”.³²

Evaluators can use the TOC at various levels, for example by: i) validating the intervention’s assumptions; ii) assessing the causal pathway of the results chain from outputs to outcomes and impact; and iii) validating the intervention’s situation analysis, including the identification of main problems and drivers of change.³³ When using a theory-based evaluation approach, evaluators are expected to use at least one of those aspects listed above, and ideally all.

When evaluating an intervention’s theory of change, evaluators need to bear in mind that those interventions operate within a system, and systemic change takes time, rarely taking place within the lifetime of an intervention.

“However, the TOC is a model to help assess the extent to which project activities and outcomes are helping to steer change in the direction of the desired transformation. Most importantly, TOCs are tools that help evaluators understand how a project be-

³² White, H. and Phillips, D., 2012, “Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework.” International Initiative for Impact Evaluation Working Paper 15.

³³ External to the UNIDO intervention but important to consider and leverage as catalysing effects for the intervention’s results achievements.

comes part of the broader process that the project seeks to influence and that can help derive lessons to provide recommendations for the future.”

UNIDO, 2019: Independent Terminal Evaluation Indonesia. SMART-Fish programme: Increasing trade capacities of selected value chains within the fisheries sector in Indonesia. UNIDO Project ID: 120110, page 9.

As an illustrative example, **Figure 11** shows how the achievement of an intervention could be presented by using a theory of change, as recently applied in an UNIDO evaluation.

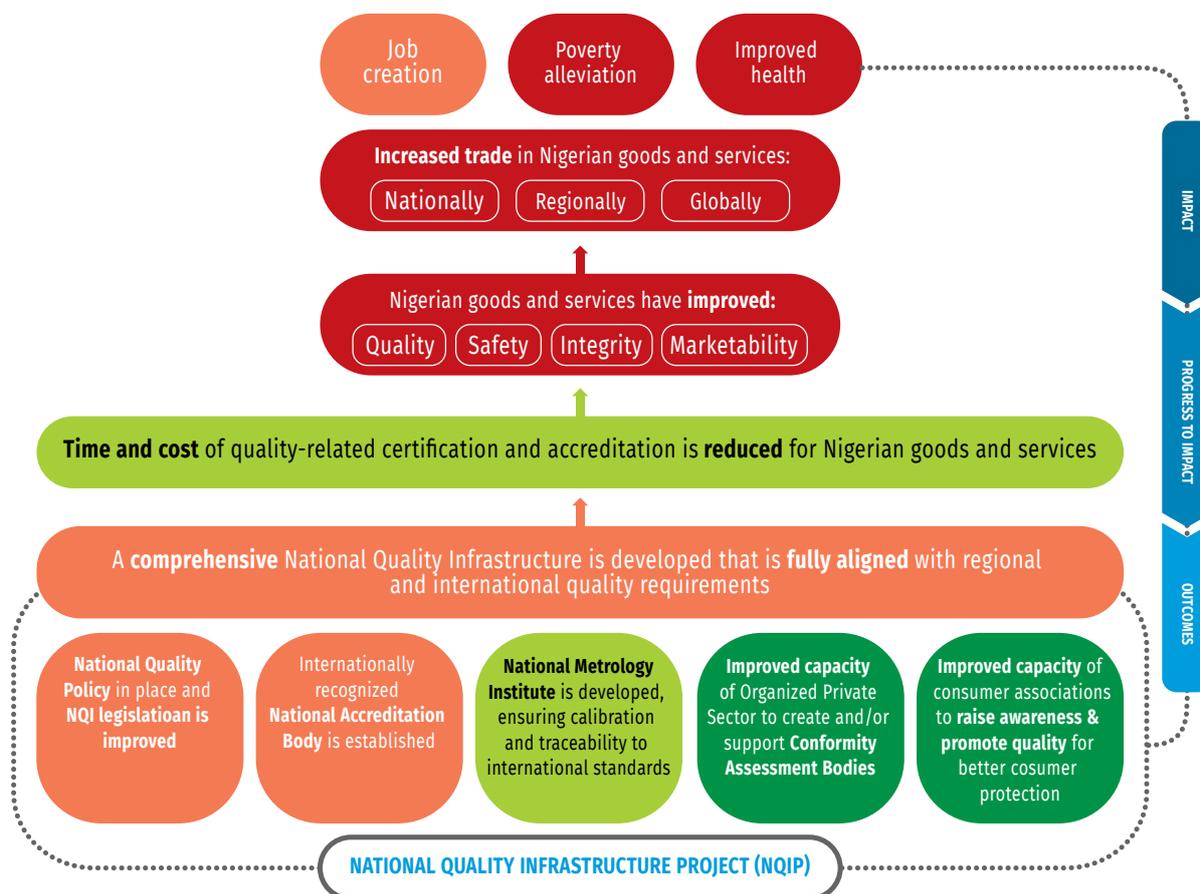


Figure 11: Assessment of progress against a theory of change with colour coding.

Source: UNIDO, 2019: Independent Terminal Evaluation. National quality infrastructure project – Nigeria. UNIDO project ID: 130220, page 27.

Using the theory of change to summarize project performance in a graphical way is good practice and should therefore be replicated where possible. Ideally, evaluators could also assess the output achievement, the validity of the intervention’s assumptions and intervention’s situation analysis, including the identification of main problems and drivers of change.

Figure 12 presents an analysis of a UNIDO theory of change which was developed at design by the programme management team. This table constitutes a good practice due to its completeness when assessing the quality of the theory of change. In the specific example, the evaluators provided this table as a summary, accompanied by a detailed annex.

CRITERIA	ASSESSMENT	RATING
1. Impact	While partly ambiguous, a long-term objective has been established.	satisfactory
2. Outcomes	Three of the four identified root causes are addressed by the three expected outcomes. Only the challenge regarding the financial mechanisms is not addressed.	satisfactory
3. Outputs	Outputs have been well defined, but not reflected in the TOC figure	satisfactory
4. Causal linkages	The causal linkages between components 1 and 2 and the causal linkages between outputs, outcomes and impact are explained in the narrative of the project document including the logical framework. The TOC figure does only to a limited extent reflect the narrative.	satisfactory
5. Assumptions	Assumptions are an important component in the project document and they are integral part of the GEIPP theory of change. The reflection of the assumptions in the TOC figure is limited.	satisfactory
6. Drivers of change	The drivers of change are mentioned: partly also how to influence them. They are not shown in the TOC figure.	satisfactory
7. Pre-conditions	The "necessary preconditions" are conceptually not quite clear and their location in the TOC figure appears to be rather arbitrary.	weak
8. Narrative and visualization	There is no easy-to-understand narrative in the project document which captures the theory of change in a concise manner and the visualization of the theory of the change is not comprehensive and partly ambiguous.	weak
OVERALL ASSESSMENT		SATISFACTORY

Table 7: GEIP Theory of change-methodological analysis

Figure 12: Example of rating a theory of change in tabular form

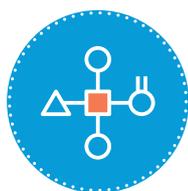
Source: UNIDO, 2021: Mid-term evaluation. Global. Global Eco-industrial Parks Programme (GEIPP) UNIDO Project ID: 170222, page 16.

Steps for undertaking an evaluation based on a theory of change approach include:

- *Expressing a theoretical idea of what changes are anticipated in advance, what the project manager considers to be the effects and changes that are likely to be seen as a result of an activity or strategy.*
- *Expressing the different explanations of how and why the actions will influence that effect.*
- *Documenting the analytical or worldview perspective on the theoretical links between effect and cause, with reference to other sources, being alert to alternative explanations of how changes could be influenced.*
- *Summing up the theory about the link, for example: 'If we take x action, then y change will result because...'*
- *Collecting evidence and information through a range of appropriate methods at key points throughout the programme cycle, in order to understand to what extent/whether observed changes can be linked back to the theory.*
- *Testing the explanation and evidence by checking if the changes seen could also be explained by other influences.*

Source: Vogel, I./Department for International Development (DFID): 2012: Review of the use of "Theory of Change" in international development. Review report, page 45 (amended).

SECTION INFO BOX:



Further reading about “Theory of change”:

<http://www.learningforsustainability.net/evaluation/theoryofchange.php>

http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/James_ToC.pdf

http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/DFID_ToC_Review_VogelV7.pdf

http://www.3ieimpact.org/media/filer_public/2012/05/07/Working_Paper_3.pdf

3.4 Data collection methods and tools

This section provides an overview of data collection techniques and tools typically used in UNIDO evaluations. Advantages and disadvantages of the techniques and practical guidance on good practices are also provided. These techniques and instruments are to be used transparently and in consultation with major stakeholders, as stipulated in UNIDO’s Evaluation Policy.³⁴

The data collection tools used in UNIDO evaluations depend on variables such as the evaluation type, availability of data, the local context, resources and time available. Frequent problems in data collection in UNIDO and possible solutions are identified below.³⁵

Access to informants

- a) *The sampling process proves to be difficult*
 - *Decide whether or not a reduced sample size is likely to provide statistically valid findings. If not, apply another technique such as the focus group.*
- b) *An informant does not express him/herself freely*
 - *Focus interviews on facts rather than opinions.*
 - *Propose to keep the collected information anonymous and explain how this will be secured.*
- c) *An informant expresses him/herself in a way which seems purposely biased*
 - *Focus demands on facts, not on opinions.*

Cultural gap

- d) *An informant or an information source can be accessed in the local language only*
 - *The evaluation team should include at least one member who is fluent in the local language (translation and interpretation always generate important information losses).*
- e) *There is a large cultural gap between the evaluation team and the surveyed group*
 - *The evaluation team should include one or several members capable of bridging the gap between the two cultures.*

Lack or weakness of data

- f) *An information source proves to be incomplete*
 - *If possible, extrapolate missing data and cross-check with other sources.*

³⁴ UNIDO, 2021, Director General’s Bulletin. Evaluation Policy (document DGB/2021/11).

³⁵ Based on: European Union, Directorate General External Relations, Directorate General Development, EuropeAid Co-operation Office, Joint, 2006: Methodological bases for evaluation, page 72.

g) An information source proves to be unreliable

- If possible, understand the biases, adjust data and cross-check with other sources.

Experience tells us that a “mixed-methods approach” that takes account of the issues addressed and the data available enhances the robustness and credibility of an evaluation, complemented with interdisciplinary and multicultural teams (when appropriate).

To the extent possible, evaluation data should be disaggregated by sex as well as by ethnicity, age, disability and any other relevant category, especially if the project is people-centered and it promotes gender equality and/or women’s empowerment.

Biases of data collection tools:

Not all, but many data collection tools can also be applied remotely, online or virtually, but the following aspects and biases need to be taken into account:

- Technology bias: Access to IT technology such as hardware, software, general IT skills, stable internet connections, etc.
- Selection bias: Issue of participation of project beneficiaries (marginalized, poorer groups, women/men, etc.).
Who will be invited? Who will be left out? How will this affect the analysis, findings, conclusions, recommendations of the evaluation?”

Triangulation is the use of multiple data sources and evaluation methods in order to strengthen the validity of data. For example, findings from key informant interviews are supported by data from a monitoring report and the online evaluation survey. Combining multiple data sources seeks to overcome the bias emanating from single informants, single methods, single observations, or single theory studies.

Figure 13 below summarizes the main data collection methods used in UNIDO.

Methods	Description	Advantages	Disadvantages
Research/ collection of documents	Investigative approach to trace and assess secondary data sources	Provides a robust basis for the evaluation; allows for systematic use of regular reporting and existing evaluative evidence	No immediate disadvantages
Structured interviews	Based upon a standardized set of questions following a pre-determined order	Facilitate greater comparability of data collected	Little or no space for improvisation
Semi-structured interviews	Cover broadly the same set of questions as structured interviews, but not necessarily in the same pre-determined order	Provide a good balance between rigor and flexibility, as spontaneous questions are possible	Comparability and aggregation of results can suffer
Unstructured interviews	No pre-defined set of questions and are similar to a conversation on a particular topic. Questions are spontaneous	Questions are spontaneous. Allows for maximum flexibility	Very limited comparability and aggregation of results

Key informant individual interviews	The interviewer is in control of the event, directing the pace of the interview, as well as providing clarification to items in the interview guide and responses to these items, if necessary. Usually conducted to ensure that important voices are heard in the evaluation	Adds to the overall validity of the data collected; face-to-face interaction helps build rapport	Large number of such interviews requires careful and often time-consuming coding; restricts the interview to the most important issues
Focus groups	Interaction of a relatively small group of people (normally 6-12) on a limited set of topics, facilitated by a moderator. Beneficiaries agree on some topics such as preferences, conclusions, beliefs or attitudes	Valuable for understanding interactions and areas of agreement and disagreement	Generalization of findings can be an issue
Memory recall	Entails interviews with beneficiaries and other stakeholders, individually or in groups, who reconstruct their situation before the project	Can be necessary where baselines are not available	Strong limitations especially with long recall periods
Participatory techniques such as wealth ranking and problem ranking	Participants are requested to come up with their own criteria and indicators to assess a situation, a process or a distribution of resources and how it has changed over time	Indicators and parameters are elicited from people rather than pre-selected by evaluators	Generalization of findings can be an issue. Challenge of aggregation
Case studies	In-depth assessment of a very limited number of observations	Can provide the information on general trends across cases that can be used to assess how an intervention is working or has worked. Allow the evaluator to experience "real" intervention examples in their entirety, which can give added insight for the evaluation	Increased opportunity for bias to be introduced into the results, because of the extent of comprehensive involvement in the cases; heavy focus on context makes it difficult to generalize the results to the larger universe of interventions
Self-administered surveys	Collects information from a large number of respondents on a series of distinct questions or issues	Surveys are a potentially powerful tool for establishing the magnitude of a given phenomenon or sentiment that can be generalized to the larger survey population; collect comparable data in a systematic manner; generally less costly than interviews	Significant amount of time required for designing, pre-testing and conducting survey. Response rates can be low, rendering follow-up a labour-intensive endeavor and making it difficult to generalize results for the survey population
Field-based surveys	"Local population surveys" are used to obtain information from intervention beneficiaries	Generate up-to-date primary data on the direct and indirect impacts of the intervention and the extent to which the intervention is making a difference in the lives of the beneficiary population(s)	Time-consuming and potentially costly nature if they are to achieve an adequate sample size for extrapolating survey data to the wider population. Potentially difficult accessibility to beneficiaries. Quality and reliability of data cannot be directly controlled if outsourced to a third party

Direct observation	Process by which data are generated through the direct observation of a situation, group or event	Relies less on the perceptions of respondents and more on the real situation being observed, including actual behaviours and body language; observer can obtain from the observation information that people would be reluctant to discuss (sensitive issues)	Can be labour-intensive and costly, unless sites are carefully sampled, there might be bias in site selection; can be difficult to identify or exclude observer bias
Field missions	Visits to intervention sites in partner countries are often indispensable to the credibility and utilization of UNIDO evaluations. However, UNIDO evaluations try to collect as much data as possible through remote means to be cost-conscious and to consider its carbon-footprint	Provide an opportunity to interview stakeholders who would otherwise be difficult or impossible to reach by remote data collection means; only way to undertake direct observation	Expensive and time-consuming to conduct, and leaves a carbon footprint; cannot be used to generate data that can be generalized to all locations or countries; can introduce bias into the evaluation

Figure 13: Main data collection methods used in UNIDO evaluations

Sources: Office for Internal Oversight Services (OIOS), 2014, International Fund for Agricultural Development, 2015, amended by Achim Engelhardt (2016).

3.5 Data analysis

Data analysis follows data collection. At this stage, the evaluation processes information gathered during data collection with the aim of answering the evaluation questions specified in the terms of reference. UNIDO Evaluation Policy highlights the importance of the “analysis of expected and achieved accomplishments, examining the results chain, processes, contextual factors and causality in order to ascertain the degree of achievement or the lack thereof”.³⁶

A sound and methodologically rigorous data analysis is key to fulfilling the criteria set out on quality evaluations under the UNIDO Evaluation Policy. This is the basis for soundly grounding evaluation findings, conclusions, recommendations, and lessons learned on evidence.³⁷

Once the validity, accuracy, and completeness of data have been verified, different data sets can be analyzed. If the evaluation identifies shortcomings in the data, this should be stated in the methodology section of the evaluation report under “limitations”, accompanied by measures taken to mitigate those limitations. **Figure 14** below summarizes generic steps in data analysis for selected data collection methods.³⁸

³⁶ UNIDO Evaluation Policy (DGB/2021/11, 21 September 2021), page 3.

³⁷ See also: OIOS, 2014: Inspection and Evaluation Manual, page 84.

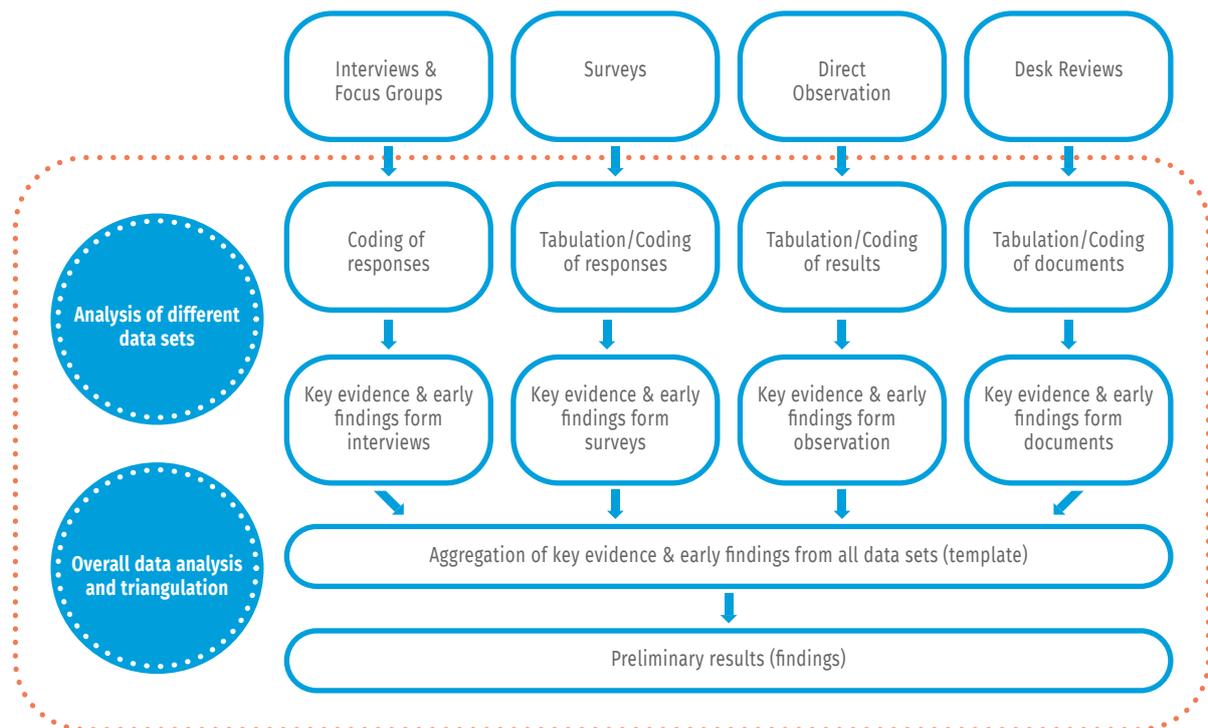


Figure 14: Examples of data analysis in UNIDO evaluations

How to analyse quantitative data

Financial, human resource and monitoring data, or direct measurements of quantitative parameters such as CO² emissions, are examples of quantitative data. Analysing time series for such data or comparing the data with available benchmarks are options for analysing this kind of quantitative data.

How to analyse qualitative data

Evaluators are encouraged to use ratings in evaluation surveys and interviews to allow for the collection of quantitative data which complements qualitative data. The latter is likely to dominate data collection through interviews. The systematic use of such ratings can support the data analysis concerning stakeholder perceptions and strengthen the evidence base. Evaluators can subsequently use charts and graphs to present the results of quantitative data analysis.

In the case of complex evaluations with a large number of interviews, coding completed interview questionnaires or protocols using appropriate software is helpful to ensure the systematic use of the collected qualitative data. Coding is a process of categorizing qualitative data into common themes to assist in identifying patterns within the evaluators' dataset.

Evaluators can again strengthen the evidence base by quantifying qualitative data, an important step during data analysis and recommended by the research community.³⁹ This can be done by organizing, reading and coding the data, then presenting and interpreting it with the use of visual aids such as charts and graphics.

³⁸ Ibid., page 89.

3.6 The “why” and “how” questions: Understanding change causes and processes

To capitalize on lessons learned from UNIDO evaluations, efforts are required to understand the factors influencing good performance as well as those that limit the achievement of UNIDO development objectives. Therefore, evaluations should ensure sufficient coverage of the “why” question in complementing “what” the performance was. This is required both to understand the rationale behind performance ratings and to enable UNIDO to consolidate lessons learned and feed learning back into the planning and implementation cycle. Given the increasing importance of theories of change, the understanding of change processes, “how” results are achieved, is equally important.

Understanding change causes and processes: Examples of change in industrial and environment policies

Successful coalition building: *Policy change enabled coordinated activity among a range of individuals with the same core policy beliefs, for example in achieving COP21 agreements on climate change in Paris in 2015.*

Opening policy window: *Policy can be changed during a window of opportunity when advocates successfully connect the way a problem is defined with the policy solution to the problem, for example in the energy policies of some European countries after the nuclear accident in Fukushima.*

Progress in advocacy and campaigning from the grassroots: *Policy change as the result of collective action by members of the community who work on changing problems affecting their lives, for example on stopping the expansion of coal mine exploitations.*

Source: Politics and ideas, 2013, adapted (<http://www.politicsandideas.org/?p=950>).

3.7 Evaluating impact: progress to impact at completion point

As highlighted above, it is often too early to assess the impact of projects at their completion date as the scaling-up and replication required to bring transformational changes take time. In addition, while it is possible to observe the direct impact on beneficiaries at local level when a project ends, it is normally necessary to look at several projects over time to understand what the long-term, sustainable impacts, or potential impacts, are. Therefore, given that almost all project terminal evaluations at UNIDO take place around completion of the project, terminal evaluations focus on assessing the impact of progress by the end of the project.

Figure 15 below illustrates the broader results and impact chain, highlighting the challenges involved in transforming results on direct beneficiaries into impacts on the whole system through transformation mechanism initiatives (e.g. mainstreaming and sustaining changes, replication, scaling-up and market change). The assessment of project progress to impact consists of two components: 1) analysis of different elements of impact dimensions, and 2) assessment of transformation mechanism embedded into the project at completion.

³⁹ United Nations, Office of Information and Communications Technology, 2021: “How to Quantify Qualitative Data: A Look at Research on Trust in the UN”.



Figure 15: Progress towards impact for UNIDO

Source: The UNIDO Integrated Results and Performance Framework (IRPF). UNIDO Annual Report 2019.

1. Impact dimensions

At corporate level, under the motto of “progress by innovation,” UNIDO has three thematic priorities corresponding to three impact dimensions:

1. Economic impact: Advancing economic competitiveness
2. Environment impact: Safeguarding the environment
3. Social impact: Creating shared prosperity (social inclusiveness)

Based on UNIDO’s most common areas of operation, its strategic documents, its focus on achieving ISID and the SDGs, the three impact dimensions can be further broken down, as presented in **Figure 16**.

To make a meaningful assessment, evaluation teams could use the elements under each impact dimension to analyse and assess progress toward impact of the evaluated projects/programmes, whether the impacts observed are at local level on direct beneficiaries, or at system level. It should be noted that not all projects will contribute to all three im-

pact dimensions, or lead to results in all the elements. Similarly, some projects may be designed to target one impact dimension, but in practice may contribute to others.

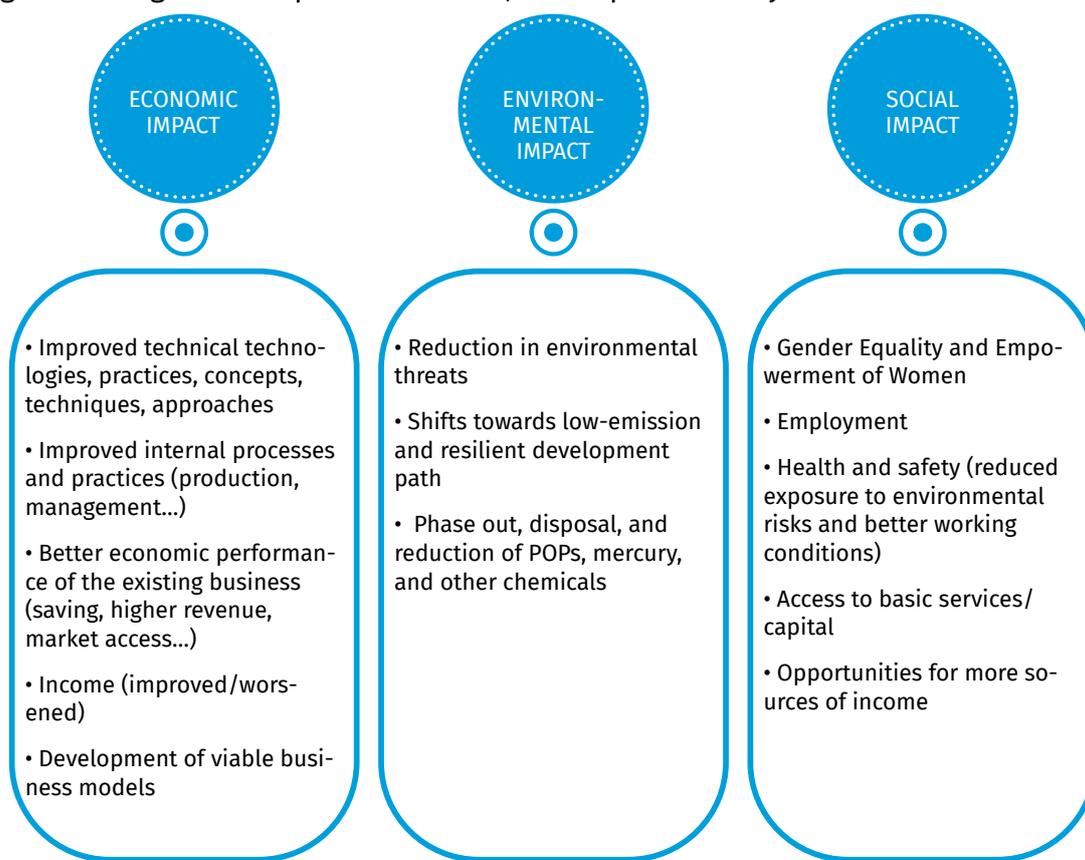


Figure 16: Impact dimensions

Source: UNIDO, 2019: Synthesis of UNIDO Independent evaluations 2015-2018.

2. Transformation mechanism (Broader adoption)

Many UNIDO interventions are pilot or demonstration projects, thus to ensure that project outcomes and local results transform into system impacts, broader adoption of UNIDO initiatives by governments and others needs to take place during project implementation or at completion. This transformation takes place in the following ways:⁴⁰

1. **Mainstreaming/sustaining:** Information, knowledge or specific results of projects are incorporated into laws, policies, regulations and programmes by governments, development agencies or private sectors.
2. **Replication:** UNIDO initiatives are reproduced/adopted in other geographical areas or regions.
3. **Scaling-up:** UNIDO initiatives are implemented at larger geographical scale.
4. **Market change:** UNIDO initiatives catalyze market transformation by influencing the supply and demand for goods and services contributing to global environmental, economic and social benefits.

Beyond the completion of a project, if these steps are not planned, adopted or imple-

⁴⁰ Progress towards impact and broader adoption concept used by the Synthesis team are based on the GEF Independent Evaluation Office's evaluation methodology (GEF 2013. Progress Towards Impact, OPS5 Technical Document #12 and GEF 2017. Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects).

⁴¹ UNIDO Evaluation Group, 2010: Independent Evaluation Sri Lanka. Impact of SMTQ projects in Sri Lanka.

mented, it would be difficult for any pilots to be mainstreamed, replicated or scaled up.

Broader adoption or transformation can be measured by assessing the extent to which contributions have been mainstreamed, replicated or scaled up. It is necessary for the evaluation team to assess the progress of a project, based on the following categories:

1. The steps have been planned but not adopted
2. The steps have been partially implemented
3. The steps have been adopted and implemented

3.8 The challenge of evaluating impact

UNIDO evaluations should endeavour as far as possible to assess the results available from an intervention at the time of the evaluation.

As indicated above, broader impact assessment, that measures the effects of interventions on a target group, examines the longer-term consequences of UNIDO interventions, and the issue of wider socioeconomic change, is not fully feasible under the current structure and mechanisms for project evaluations.

Since its first impact evaluation in 2010, UNIDO has carried out a number of ad hoc initiatives to evaluate impact. For example, in 2010 it made the following methodological observations, which are still valid today and which should be considered for any future impact evaluation in the Organization:⁴¹

- **Baseline data:** *Because the impact evaluation has not been planned in advance, baseline data are not available, which limits the possibilities to thoroughly compare ex ante and ex post situations.*

- **Time-lag:** *It has been more difficult than expected to make all critical information available because considerable time has elapsed since the projects were designed about 10 years ago, started about 7 to 8 years ago and ended about 4 to 5 years ago. Therefore, a comparison between services delivered before and after the intervention would not be possible.*

- **External factors:** *During the period under evaluation, the “competitiveness” of the Sri Lankan economy was heavily influenced by a number of external factors and macro-events. Hence the methodological challenges of isolating the hypothetical influence of improved Standards, Metrology, Testing and Quality services from other much more powerful factors such as:*

- i) *socioeconomic effects of the ups and downs of the civil war and post-conflict situation of the country;*

- ii) *destructive effects of the 2004 tsunami on the fisheries sector combined with the extraordinary recovery efforts and donor support that followed this major natural disaster;*

- iii) *improved export opportunities to the European Union due to Sri Lanka accessing to the privileged status of a GSP+ country (Generalized System of Preferences); and*

- iv) *turbulences of the world textile market due to the end of the Multi-fibre Agree-*

ment affecting heavily the textile sector, Sri Lanka's number one foreign currency earner.

- **Stakeholder willingness to cooperate:** *The collection of company data turned out to be particularly difficult and time-consuming. Private sector companies in Sri Lanka were not at all keen to spend time on surveys and even less to share sensitive data on their specific position in highly competitive export markets.*

Source: UNIDO Evaluation Group, 2010: Independent Evaluation Sri Lanka. Impact of SMTQ projects in Sri Lanka, pages 3-4.

Impact evaluations require considerably larger budgets than other evaluation types due to more complex methodology, earlier preparation time, and the need to ensure that adequate and reliable data are collected.

■ 3.8.1 Impact evaluation techniques

To assess the impact of UNIDO's work in a credible manner, planning for impact assessment **must start at the project or programme development stage**. The use of baselines and comparator groups is recommended to ensure questions about impact are answered robustly.

Before and after: From a methodological point of view, impact can be assessed using the "before and after" technique to attribute effects to a particular UNIDO intervention, requiring, however, a robust baseline and adequate M&E systems. The technique of memory recall is applied for such purposes but tends to be rather unreliable.

The "**with and without**" technique is another option to assess the impact of UNIDO interventions. For this purpose, evaluators need to define a plausible counterfactual. (A counterfactual is a situation or scenario that would hypothetically prevail if there was there no UNIDO intervention.) This can be accomplished by assessing the results of interventions on target groups (institutions, companies, communities) and by comparing them with the situation of "populations" outside the target group in a similarly situated business sector, institutional setting or geographic region. UNIDO has learned from its peers that the identification of comparator groups at the outset of a development intervention requires utmost attention to ensure the comparability. This includes large similarities between both groups. Once control groups are identified, the issue of incentives for those groups emerges to form part of the impact evaluation design. Including the comparators in the UNIDO intervention at a later stage (following the intervention phase) might be such an incentive. Otherwise, the control groups are unlikely to invest their time in interviews and allow for extensive data collection. This is particularly true for the private sector.

A **counterfactual** which is more feasible to systematically take across all UNIDO evaluations is perception based, easy to establish and a proven proxy measure for UNIDO impact. Essentially, the perception-based counterfactual does not use target and control groups but asks stakeholders what would have happened if the UNIDO intervention had not taken place.⁴²

⁴² For more information, consult: <https://plato.stanford.edu/entries/causation-counterfactual/>

The basic idea of counterfactual theories of causation is that the meaning of causal claims can be explained in terms of counterfactual conditionals of the form “If A had not occurred, C would not have occurred”.

Source: Stanford University, 2019.

Impact towards policy and institutional change: For the evaluation of progress towards impact of UNIDO’s work on policy and institutional change, UNIDO builds on research to apply an evaluation framework of normative work, as presented in **Figure 17** below.⁴³

The steps by which influence on policy lead to behaviour change, and ultimately to advancements in ISID, are not necessarily linear and should be treated during the evaluation accordingly.

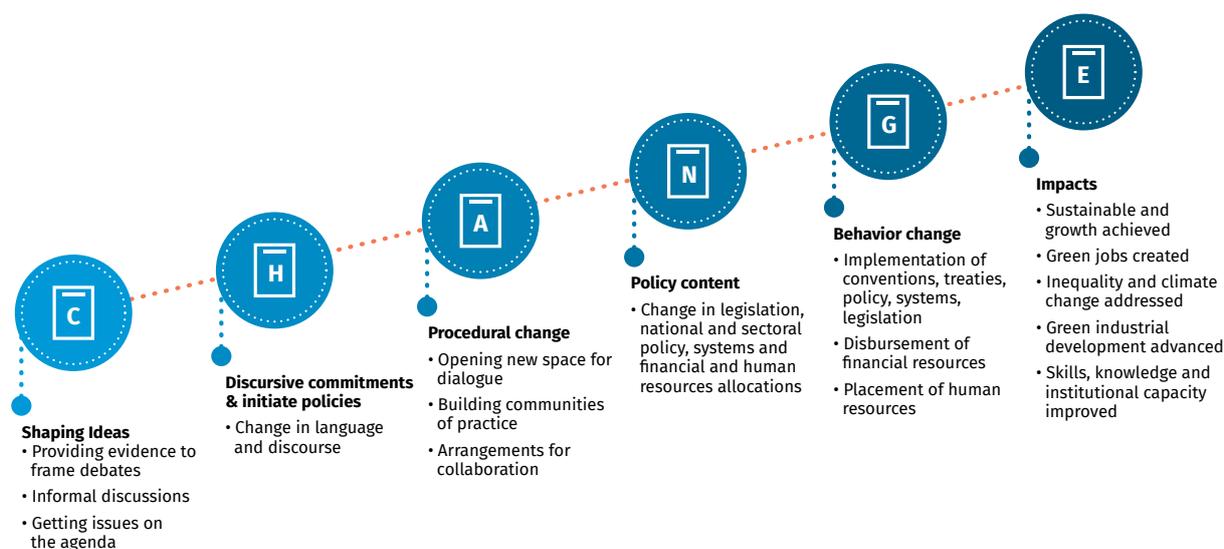


Figure 17: CHANGE framework for evaluating impact towards policy and institutional change

Sources: Jones, N. with Villar, E. (2008), Keck, M. and Sikkink, K. (1998), UN Evaluation Group, (2013), Aarvar et al (2012), Engelhardt, A. 2015, amended 2022.⁴⁴

1. First steps in a policy influencing process can be described as “shaping ideas”, “discursive commitments” and “procedural change”. At times those processes happen in parallel. Activities for shaping ideas could include for example informal discussions, the framing of debates, and getting issues on the agenda of a policy partner or an international forum concerning ISID.

⁴³ This framework has been further developed under the global DFID Forest Governance Climate and Market Programme in 2015 and as part of an institutional evaluation of policy influencing as part of work for the Swiss Development Cooperation in 2015 by the consultant supporting the development and subsequent review of the UNIDO Evaluation Manual.

⁴⁴ Jones, N. with Villar, E., 2008, “Situating children in international development policy: challenges involved in successful evidence-informed policy influencing”. Evidence and Policy, vol.4, no.1: p 53-73.
Keck, M. and Sikkink, K., 1998, Activists beyond borders: advocacy networks in international politics. Ithaca, NY: Cornell University Press.
Jones, Harry, 2011, A guide to monitoring and evaluating policy influencing. Overseas Development Institute (ODI), UNEG, 2013, UNEG Handbook for Conducting Evaluations of Normative Work in the UN System, page 25.
Aarva P, Zukale S, Magnusson A. and Nogueira de Morais, I., 2012, Evaluation report 2012:6 Nordic Influence in Multilateral Organizations: A Finnish Perspective.. Ministry for Foreign Affairs of Finland, Kopijyvä Oy, Jyväskylä. Engelhardt A., et al., 2015, Evaluation of SDC’s Global Programmes Climate Change, Water Initiatives, Food Security, Migration and Development and Health, external institutional evaluation commissioned by the Swiss Agency for Development and Cooperation (SDC).

2. Changes in the language and terminologies used during consultations or in position papers are examples of discursive commitment and the initiation or revision of policies.
3. Procedural change tends to follow discursive commitments and relates to opening new spaces, for example through coalitions for decision-making committees or arrangements for joint financing.
4. Changes in policy content are often a next step in the policy influencing process. This can be achieved by providing expertise during consultations such as replenishment or treaty negotiations. As a result, changes might be achieved in conventions, international treaties, policies, legislation or budget allocations.
5. The implementation of the latter will eventually show real behaviour change of the responsible bodies.
6. Ultimately, the implementation, including the disbursement of related financial resources and placement of human resources, is aimed at advancing inclusive and sustainable industrial development, for example by creating green jobs and sustaining green growth to foster a green economy.

Evaluators can use the CHANGE framework to identify how far a UNIDO intervention may have advanced along the change pathway, for example in sectoral policy influencing. This seems particularly relevant for thematic or country-level evaluations.

3.9 Benchmarking

Benchmarking allows comparing the performance of a specific project or programme, sector, geographic area, or of UNIDO as a whole, with other comparable units. Good practices from other specialized UN agencies show the value of using benchmarking as an instrument to learn from the processes or approaches of peers. This can be the case when UNIDO's experience concerning a specific approach or process might be limited. Despite the value of benchmarking, comparators have learned that benchmarking results should always be interpreted carefully.⁴⁵

3.10 Formulating recommendations and management action plans (MAPs)

UNIDO uses UNEG guidance on the formulation of evaluation recommendations. UNEG advises that “recommendations should be firmly based on evidence and analysis, clear, results-oriented and realistic concerning implementation”.⁴⁶ It is of particular importance to ensure that recommendations are not opinion-based, as endorsed by UNEG and practiced by UNIDO.

Evaluators will also be also expected to “give special attention to analyzing lessons and proposing recommendations on aspects related to factors that contributed to or hindered attainment of project objectives, sustainability of project benefits, innovation, catalytic effect and replication, and project M&R”.⁴⁷

⁴⁵ The International Fund for Agricultural Development (IFAD), 2009: Evaluation Manual. Methodology and processes.

⁴⁶ UNEG, 2016 *UNEG Norms and Standards for Evaluation*, page 26.

⁴⁷ GEF, 2008, *Guidelines for GEF Agencies in Conducting Terminal Evaluations*, page 14.

UNIDO using UNEG Standard 4.10: Recommendations

Utilization-focus:

- *Recommendations should be formulated with their use in mind. Depending on the subject of the evaluation, recommendations could indicate strategic directions or be more focused on operational matters.*

Implementation in mind:

- *Recommendations should be clear on who needs to implement them. In formulating recommendations, it is good practice to consult the likely implementers of the recommendations to secure their acceptance and feasibility.*
- *Recommendations should not be overly prescriptive. However, it may be helpful to include operational details and proposals for practical actions for implementation to enhance the understanding, ownership, and commitment of those who will respond.*

Source: UNEG, 2016 UNEG Norms and Standards for Evaluation, page 26, amended.

Recommendation checklist:

- *Recommendations are firmly based on evidence and conclusions.*
- *Recommendations are within the scope of action of the project/programme evaluated (not a generalization to a systemic issue).*
- *Recommendations are relevant to the object and purposes of the evaluation.*
- *Recommendations clearly identify the target group for each recommendation.*
- *Recommendations are actionable and reflect an understanding of the commissioning organization and potential constraints to follow-up.*

Since 2021, the UNIDO Independent Evaluation Unit has adopted the use of MAPs for all the strategic evaluations.

MAPs are formulated by relevant management functions to make explicit “how the Organization will implement the recommendations and address the respective findings. They should provide explicit verifiable deliverables, responsibilities and a timeline for their implementation”.

MAPs are validated by EIO in order to ensure consistency and alignment with the respective recommendations and to become a mechanism for further follow-up on strategic evaluations.

Normally, MAPs are also included in the final evaluation report in order to demonstrate management ownership and commitment, and as an indicator of the use of the evaluation. Monitoring and implementation of MAPs are under the explicit responsibility of the assigned senior manager (normally in coordination with other business owners). EIO follows up on each MAP at the due date for its implementation. MAPs are closed by EIO upon validation of their implementation through the verification of the specific deliverables expected by the MAP.

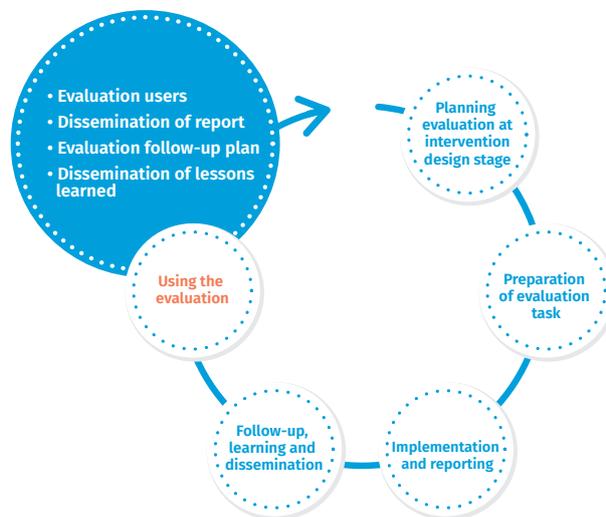
When MAPs cannot be implemented by management within the allocated timeframe consideration can be given to a deadline extension and/or to reformulation of the MAP. In

exceptional cases, closing a MAP without full implementation can be justified; in such cases, EIO will refer the issue to senior management as needed.

3.11 Participation

Participation is a key principle in UNIDO evaluations to ensure: i) sharing understanding; ii) ownership of the evaluation products; iii) useful evaluation results; iv) recommendations; and v) lessons learned. Participation also increases the credibility and quality of evaluations. This is anchored in UNIDO’s Evaluation Policy and in line with UNEG’s Standard 4.6 on “stakeholder engagement and reference groups”.

UNEG’s Standard 4.6 specifies that: “processes should be in place to secure the participation of individuals or parties who may be affected by the evaluation, can influence the implementation of recommendations, or who would be affected in the long term. Stakeholders should be consulted in the planning, design, conduct and follow-up of evaluations. A variety of mechanisms can be used to consult with a broad range of stakeholders (e.g. consultation meetings on evaluation design, validation workshops on preliminary findings and post-evaluation learning workshops). Besides, different types of stakeholder groups could be formed for their continued engagement (reference groups, learning groups, and advisory groups)”⁴⁸



3.12 Using the evaluation

UNIDO’s Evaluation Policy focuses the use of the evaluation on knowledge building and, ultimately, organizational learning. This is achievable given that evaluation recommendations and lessons learned target the needs of the relevant target audiences in UNIDO.

EIO is responsible for facilitating the use of evaluations and the adoption of lessons learned through targeted briefing materials beyond the actual evaluation report. These can include:

- policy briefs for UNIDO’s Member States or evaluation briefs for all stakeholders; and
- written evaluation summaries for decision makers among UNIDO’s evaluation stakeholders.

“The contribution of evaluation to organizational learning goes beyond UNIDO. It also shares lessons learned among stakeholders, including with other organizations of the United Nations System, using UNEG as a knowledge network.”

Source: UNIDO, 2021, Evaluation Policy, page 12.

⁴⁸ UNEG, 2016 Norms and Standards for Evaluation, page 24.

3.12.1 Evaluation users

The use of any evaluation should be planned for from the outset of the exercise. When considering the purpose of the evaluation, the ToR should clearly indicate who the main and potential users of the evaluation and lessons learned from an evaluation are. The interest of potential users can be strengthened through involving them in the process, as needed. This could be achieved, for example, by including their inputs into the evaluation ToRs and through a presentation of preliminary findings, as specified in the relevant sections above.

3.12.2 Dissemination of the evaluation report

Evaluation teams should provide an electronic version of the report. All evaluation reports are made publicly available on the UNIDO website and on the Organization's Intranet.

A key stage in the process is the dissemination of report findings, conclusions, recommendations and lessons learned, As outlined in the UNIDO Evaluation Policy,⁴⁹ the evaluation office should ensure that an effective and proactive communication and dissemination strategy is carried out, as suggested by UNEG's Standard 4.11.

UNEG's Standard 4.11: Communication and dissemination

"Communication and dissemination are integral and essential parts of evaluations. Evaluation functions should have an effective strategy for communication and dissemination that is focused on enhancing evaluation use."

UNEG, 2016 Norms and Standards for Evaluation.

3.12.3 Lessons learned and good practices

Learning is one of the dual purposes of the evaluation in UNIDO. Evaluations can capture "what works for whom and in which situation" to achieve UNIDO strategic objectives. UNIDO shares the definition of lessons learned as used in other specialized UN agencies:

"A lesson learned is an observation from project or programme experience which can be translated into relevant, beneficial knowledge by establishing clear causal factors and effects. It focuses on a specific design, activity, process or decision and may provide either positive or negative insights on operational effectiveness and efficiency, impact on the achievement of outcomes, or influence on sustainability. The lesson should indicate, where possible, how it contributes to 1) reducing or eliminating deficiencies; or 2) building successful and sustainable practice and performance."⁵⁰

⁴⁹ UNIDO, 2021, Director General's Bulletin: Evaluation Policy, page 6.

⁵⁰ International Labour Organization (ILO) Evaluation Unit, 2014: "Guidance Note 3: Evaluation lessons learned and emerging good practices".

UNIDO lessons learned should fulfil the following quality criteria:

- Context: the economic, social or political context from which the lesson has been derived.
- Challenges: difficulties, problems or obstacles encountered, and solutions found. Positive and negative aspects should be described.
- Causal factors: evidence for “how” or “why” something did or did not work.
- Target users affected by the lessons learned.
- Success or failure: The lessons learned should cite any decisions, tasks, or processes that constitute reduced or eliminated deficiencies or built successful and sustainable practice and performance; or have the potential of success.
- Lesson learned should not be mistaken for a recommendation or conclusion.

Lessons learned and good practices are part of a continuum that allows their reuse in the UNIDO project, programme, and policy cycle. Some main differences between a lesson learned and good practice are:

- Successful interventions or strategies that performed well
- Through establishing a clear cause-effect relationship, the practice has achieved marked and measurable results or benefits
- Related strategies are assessed to be specifically useful for replication or upscaling

Successful lessons are presented as emerging good practices when also the following other criteria are fulfilled:

- Potential for replication
- Link to UNIDO policy goals

4. The project evaluation process: Step-by-step

This section leads readers through a standardized project terminal evaluation process in UNIDO: planning and budgeting for the evaluation at project design, planning the conduction of the evaluation, conducting the evaluation, reporting and disseminating, follow-up, learning and using the evaluation.

4.1 Background

The flowchart in Figure 18 provides an overview of specific steps, roles, and responsibilities during the evaluation processes for independent project evaluations. This is applicable to independent mid-term evaluation and terminal evaluations.

Figure 18 : Flowchart for Independent project evaluations

The following sub-sections describe the process of an independent project evaluation.

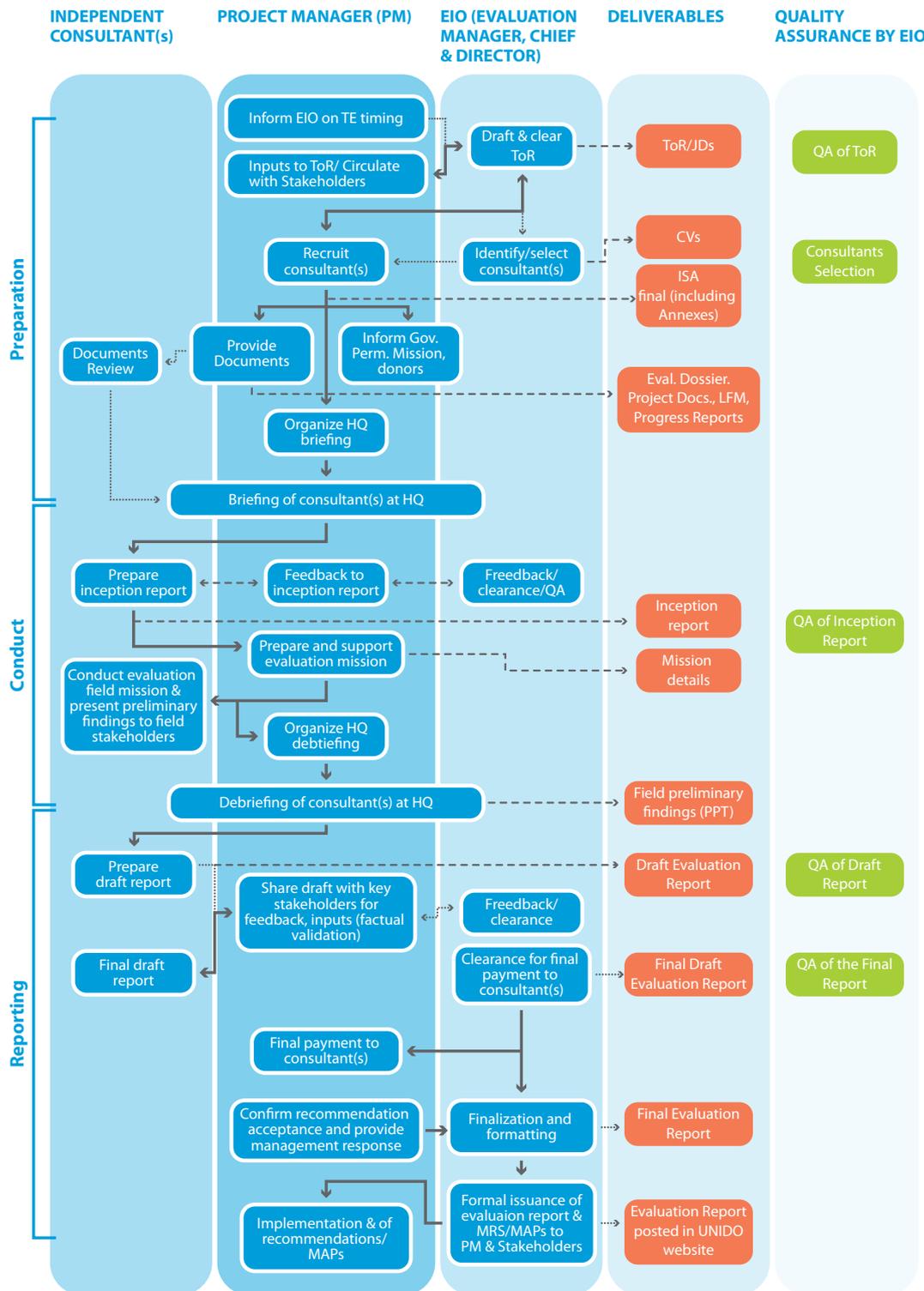


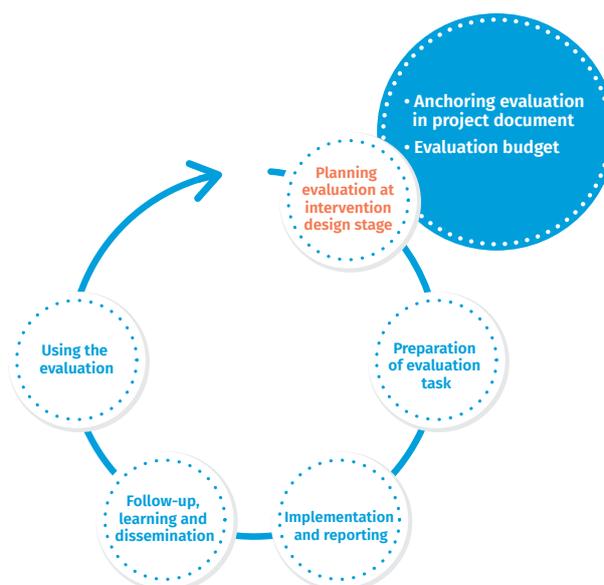
Figure 18: Flowchart for Independent project evaluations

4.1 Planning an evaluation at the design stage

To assess development results and impact, a project evaluation needs to be planned for from the design stage of the intervention.

The **project document** should incorporate the evaluation process for all UNIDO interventions. Even if the evaluation is planned for at this early stage of an intervention, changes are still possible throughout the lifecycle of the intervention.

Those changes should be justified and properly recorded in the project documentation and referred to in the evaluation ToRs.



4.2 Evaluation planning and budget

As part of standard good management practice, all projects must have their respective monitoring & reporting (M&R) and evaluation provisions and their respective budgets (commensurate to size and complexity). Budgets for M&R activities and evaluation activities should remain separate.



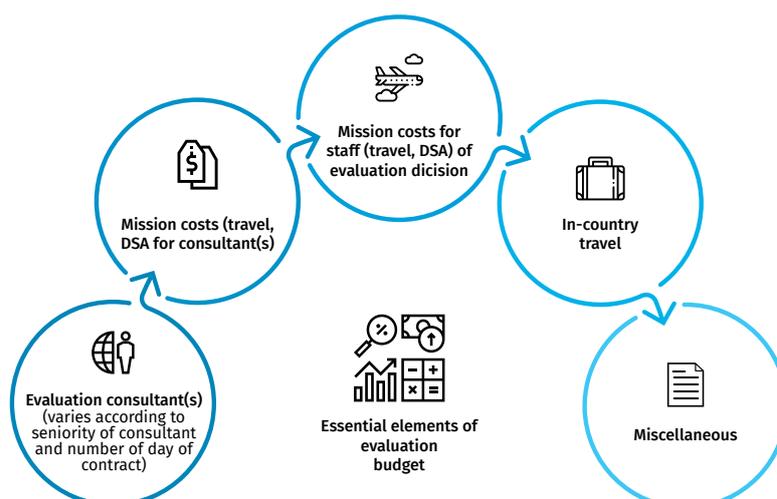
PM to-do list:

- Design M&R, and evaluation plan
- Set M&R budgets
- Set evaluation budget

All projects must have at least one evaluation conducted at the end of the project. This evaluation could be a self-evaluation or an independent evaluation.

Independent project evaluations will be conducted on selected projects on the following basis:

- When required by donors, which would be reflected in the project document at the time of project design and formulation.
- When selected by EIO during its periodic risk and evidence gap assessment.



In general, EIO is responsible for conducting independent project evaluations, with funding allocated from within the project budget. In some cases, donors can choose to conduct external independent evaluations themselves.

In cases where there is no independent project evaluation, **self-evaluations** must be conducted under full responsibility of project management by the end of the project. Final self-evaluation reports are to be submitted to EIO for record-keeping.

In terms of budget, the average cost for a terminal project evaluation is around €40,000 to €70,000 (or around 2-4 per cent of the total project budget), depending on the project size, complexity, and geographic coverage. In case of regional projects, the evaluation budget would be around €70,000 to €100,000.

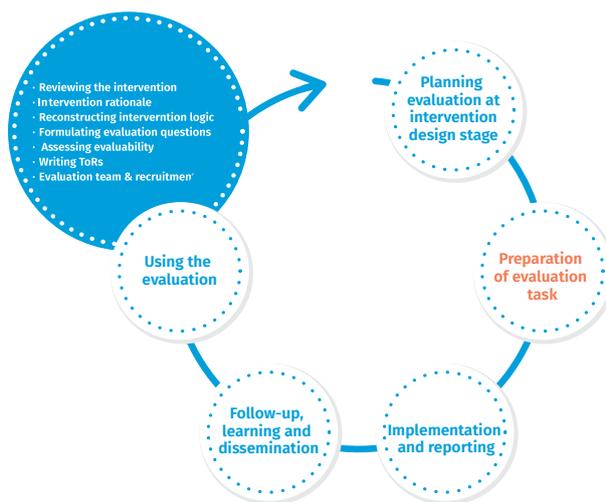
4.3 Preparation of evaluation task

While keeping the principle of independence, UNIDO evaluations should be participatory and inclusive (see Section 3.9). In this spirit, stakeholders should be involved as early as possible in the preparation of evaluations. The extent of stakeholder engagement should be balanced with practical considerations such as time and cost.

Stakeholders can contribute by helping to clarify the purpose of the evaluation, one of the most important tasks at the outset of preparing an evaluation. A clear purpose facilitates the formulation of evaluation questions and ensures a relevant and useful evaluation product.

As a matter of good practice, the timeliness and intentionality of evaluations are addressed from the outset of the evaluation planning. Providing timely information that meets the needs of intended users should facilitate decision-making processes.

Every independent evaluation in UNIDO is managed by a responsible evaluation officer (EO) from EIO, who will manage, facilitate, quality assure and monitor the evaluation process.



4.3.1 Developing terms of reference (ToR)

The terms of reference establish the purpose and objective, methodology, process and contractual basis for an evaluation in UNIDO. Well-written ToRs determine the quality of an evaluation, as they contain the evaluation requirements and expectations. ToRs are the point of reference during any stage of the evaluation and help to maintain a mutual understanding of the evaluation.

Evaluation consultants are referred to the ToR as the first step to learning about evaluation requirements.

The key components for any evaluation ToR are:

- the evaluation context and purpose;



PM to-do list:

- Provide inputs to the draft ToR (on the basis of ToR template provided by EIO)
- Circulate ToRs with stakeholders

EO to-do list

- Feedback, clearance & QA to the ToR
- Publication of the ToR in evaluation website.

- a description and a clear definition of the subject to be evaluated;
- the scope of the evaluation;
- the evaluation objectives with key evaluation questions and criteria;
- evaluation methodology;
- management arrangements;
- expected deliverables; and
- the evaluation process and timetable.

UNIDO Guidelines for Terms of Reference specify that writing ToRs is a consultative process led by the EIO. For project evaluations, the initial drafting of the ToR is delegated to the project manager (PM) (on the basis of a template provided by EIO). The final review and clearance of the ToR is under the responsibility of EIO.

The importance of the consultative nature of writing ToRs is emphasized in the **Box 5** below.

Box 5.

“If the intended users of the evaluation participate in the process right from the beginning, the usefulness of the evaluation outputs is likely to be enhanced. (...) As the intended users know their information needs better than anyone else, including them in defining evaluation questions is clearly very important.”

Source: UNIDO, 2016, UNIDO evaluation tools. Guidelines for Terms of Reference for Evaluations, page 2.

UNIDO embraces the UNEG standards, stating that, “Evaluations should be designed to ensure that they provide timely, valid and reliable information that will be relevant to the subject being assessed and should clearly identify the underlying intentionality. (...) Timeliness is thus an important factor in ensuring evaluation utility.”⁵¹

Regarding evaluation ratings, UNIDO Guidelines for Terms of Reference clarify that, “Evaluation ToR are meant to guide evaluation teams to producing reports that include a systematic rating system for design quality and implementation performance of individual projects and/or of an entire programme. This is to be achieved through applying consistent rating criteria.”⁵² Those rating criteria are described in the relevant section above.

Complying with UNIDO guidelines ensures high-quality ToRs. The latter are the cornerstone of a quality evaluation. Evaluation managers also benefit from both a UNIDO and UNEG checklist for evaluation ToRs to ensure the consistently high quality of ToRs in UNIDO.⁵³

In principle, all evaluations conducted or commissioned by UNIDO must integrate human rights and gender equality, and aim to meet the requirements of the UN-SWAP on Gender Equality and the Empowerment of Women Evaluation Performance Indicators.

⁵¹ UNEG, 2016 UNEG Norms and Standards for Evaluation, UNEG Standard 4.1: Timeliness and intentionality, page 21.

⁵² UNIDO, 2016, UNIDO evaluation tools. Guidelines for Terms of Reference for Evaluations, page 2.

⁵³ United Nations Evaluation Group, 2010, UNEG checklist for evaluation terms of reference and inception reports. Guidance document.

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

**SECTION INFO BOX:
INTRANET**



HYPERLINK

UNIDO Guidelines for Terms of Reference

UNIDO, 2016: UNIDO evaluation tools. Guidelines for Terms of Reference for Evaluations.

4.3.2 Intervention rationale

To understand the evaluation object, its main features should be reviewed early in the evaluation process. Understanding the intervention and its intended logic facilitates the formulation of evaluation questions.

UNIDO’s project documents should contain the rationale for the intervention, which should be summarized in the project logframe and addressed in the project’s theory of change (see Section 3.4). The intervention rationale is defined as follows: “The rationale of an intervention is to satisfy the needs, solve the problems or tackle the challenges that are considered to be priorities in a particular context and that cannot be addressed more effectively in another way.”⁵⁴

For thematic or country evaluations, such a project document might not be available. However, thematic or country strategies and related ToRs can contain the rationale for UNIDO’s engagement.

The evaluators should highlight the main points of the rationale, note changes in the rationale during the intervention lifecycle or reformulate the rationale in case the project document is ambiguous.

4.3.3 Assessing evaluability

As part of preparing the evaluation, it is good practice for the evaluation manager to double check the evaluability of the project or programme. This helps to increase the likelihood of the evaluation delivering timely and credible information for decision-making in UNIDO.⁵⁵

The evaluation manager should verify: i) that timing of the evaluation is relevant in relation to the intervention cycle; ii) the availability of sufficient budget for the planned evaluation; iii) the specificity of the intervention logic; iv) the existence and quality of baseline, monitoring, and implementation data, disaggregated by gender and other relevant criteria; v) opportunities to make use of planned stakeholder meetings for evaluation interviews; and vi) knowledge of aspects of possible challenges hindering an impartial evaluation process.

The evaluability assessment helps to find out to what extent the proposed evaluation ques-

⁵⁴ European Union, Directorate General External Relations, Directorate General Development, EuropeAid Co-operation Office, Joint, 2006: Methodological bases for evaluation, page 37.

⁵⁵ In line with the UNEG’s Standard 4.2 on evaluability assessment.

tions can be answered, potentially leading to modifications of the evaluation design, the reformulation of evaluation questions or the timing of the evaluation.

The evaluation of a cluster of projects requires a specific approach to assess evaluability.

Why evaluate a cluster of projects?

Cluster evaluations allow for more strategic focus when evaluating projects in a similar thematic area, or with a strategic scope. In addition to the potential efficiency gains, one of the most significant benefits of consolidating multiple assessments into a single assessment is that commonalities and differences between thematically or geographically similar projects can be analyzed. This can help identify critical success factors and potential risks, providing valuable information for the performance of ongoing and future UNIDO interventions.

UNIDO evaluates a cluster of projects, for example, when undertaking a country evaluation.

In addition to the criteria listed above, for evaluating a cluster of projects the evaluability assessment also needs to consider the following aspects: i) common planning framework of the cluster; ii) contribution of the projects to similar UNIDO objectives, for example in the Medium-Term Programme Framework; iii) timing of the evaluation considering the implementation status of the projects (“level of maturity”, e.g. mid-term or final).

■ 4.3.4 Evaluation team composition and recruitment

The UNIDO Guidelines for Terms of Reference elaborate on the composition of the evaluation team (ET). The specific ET requirements are laid out in the respective ToRs for the evaluation, based on the evaluation subject, focus, methods, and analyses. A good skills mix is necessary with interdisciplinary teams providing the best results. It is important to keep in mind that independent terminal project evaluation is more about assessing actual results, changes and effects of the interventions and learning from good and bad practices, and should not be mixed or considered as a “technical” study (which can always be conducted by the project manager as part of the project). The evaluation team, in particular the evaluation team leader (TL) should have an adequate background to be able to understand the technical issues in the project, but it is not necessary to have a technical expert for conducting an evaluation.

A typical evaluation team is composed of one international evaluation expert and a national consultant. The composition, including the total number of team members, depends on the complexity of the intervention to be evaluated.

Evaluators should be independent of all project activities and have no stake in the outcome of the evaluation. The latter is particularly important in the case of highly specialized technical experts active in sub-sectors with very limited human resources available for evaluations.



EO to-do list:

- Final formatting & MRS
- Submission of evaluation report & MRS to: PM & stake-holders, MRS respondents
- MAPS

PM to-do list:

- Implement & follow-up recommendations in MRS

The qualifications and skill areas for evaluation teams should include:

- Evaluation skills appropriate to the subject area
- Technical competence
- Ability to address relevant cross-cutting thematic issues, including gender, human rights and environmental and social safeguards (ESS)
- Adequate understanding of local social and cultural issues
- Appropriate language skills
- Process management skills, including facilitation skills
- Writing and communications skills
- Good interpersonal skills
- Adequate mix of national and international expertise and of women and men
- Knowledge of the ethical evaluation standards as stated in the UNEG Norms and Standards

Source: UNIDO, 2018: UNIDO evaluation tools. Guidelines for the Preparation of Terms of Reference for Evaluations. Evaluation team composition, pages 9-10.

UNEG Evaluation Competence Framework:

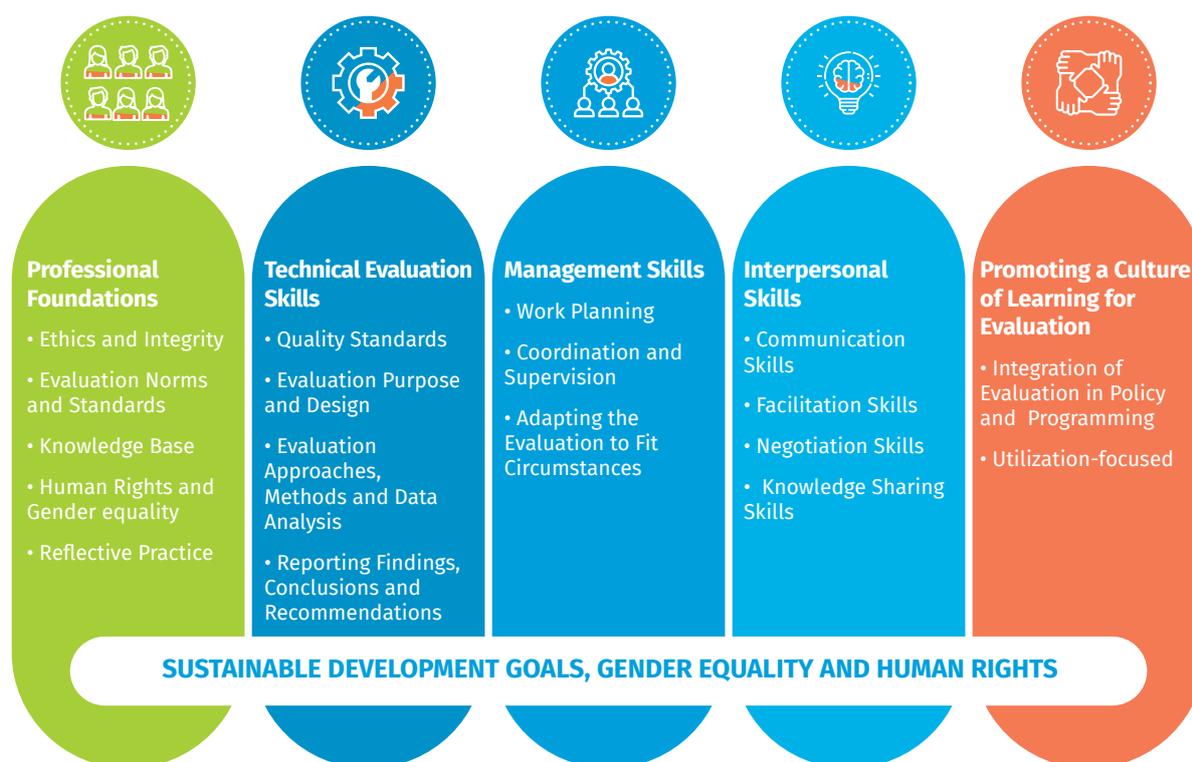


Figure 19: UNEG Evaluation Competence Framework

UNEG’s Standard 4.8 on the selection and composition of evaluation teams stresses the requirement for an open and transparent selection process, as well as the use of professionals from the countries or regions concerned. UNEG calls for strong evaluation expertise in evaluation teams: **“the core members of the evaluation team must be experienced**

evaluators with appropriate methodological expertise”⁵⁷. This pledge is backed up by evidence,⁵⁸ as shown in the **Box 6** below.

Box 6.

“UNIDO is a specialized UN agency. In this house we need thematic specialists for evaluations. Why would we need professional evaluators?”

The United Kingdom-based Centre for Development Impact (CDI), a joint initiative between the Institute of Development Studies at the University of Sussex, the University of East Anglia and Itad, undertook a study of the factors influencing the quality of evaluations. The study found that: “there is a specific set of skills unique to evaluation that are required to deliver a quality evaluation product”. As part of this skills set Australia’s Office of Development Effectiveness identified technical knowledge of different evaluation methodologies; knowledge of how to lead an evaluation and the management of both international and local consultants; strong diplomatic and interpersonal skills; expertise in collecting, analysing and presenting data; and writing credible reports in a tight timescale as key evaluation skills.

The CDI found that this finding is supported by other research in this area. The United States Agency for International Development (USAID) found that USAID evaluations employing an evaluation specialist as part of the team were statistically of significantly higher quality. This was the result of reviewing the quality and coverage of 340 randomly selected evaluations completed between 2009 and 2012 (USAID 2013) (USAID 2013: 119). Australia’s Office of Development Effectiveness and a recent evaluation of the Norwegian Aid Administration confirms this finding (DFAT 2014: 35, Itad/Chr. Michelsen Institute 2014: 81)”

Source: Centre for Development Impact, Institute of Development Studies: Improving Quality: Current Evidence on What Affects the Quality of Commissioned Evaluations. CDR Practice Paper No. 9, March 2015, Lloyd, R., Schatz, F.

In UNIDO evaluations, technical experts can support evaluations as team members.

In the evaluation process, the selection of evaluators is an important step. Evaluators qualified for the assignment enhance the likelihood of a satisfactory evaluation process and quality deliverables.

Experience tells us that sufficient time needs to be allowed for the recruitment of the evaluators.

The recruitment process itself may also add to the time required before the evaluation can start and should be included in the evaluation planning.

The evaluation manager has the responsibility for selecting the evaluator(s). This might be done following prior consultation with the donor or other partners.



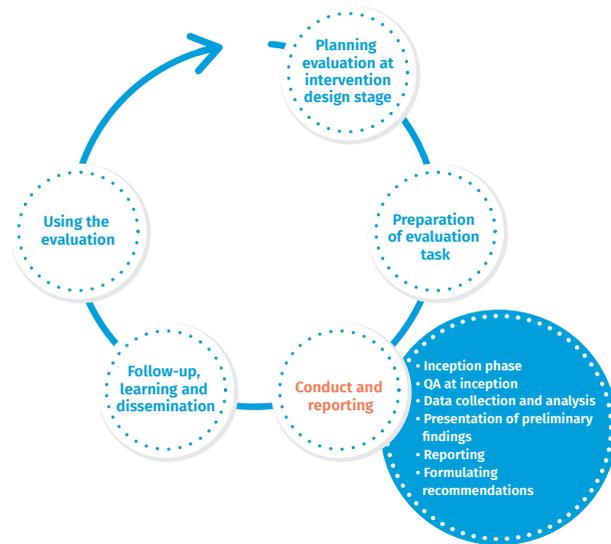
PM to-do list:

- Provide documents
- Inform Gov., donor(s), stakeholders in HQ and in the field and respective area office(s)
- Organize HQ briefing of consultants

⁵⁷ UNEG, 2016 UNEG Norms and Standards for Evaluation, page 25.
⁵⁸ Centre for Development Impact, Institute of Development Studies: Improving Quality: Current Evidence on What Affects the Quality of Commissioned Evaluations. CDR Practice Paper No. 9, March 2015, Lloyd, R., Schatz, F.

4.4 Evaluation conduct and reporting

This section guides the implementation of the evaluation. Different evaluation phases are discussed, including inception, field visit, preliminary findings, reporting, follow-up, and management response. Guidance on data collection and data analysis is provided in Section 3 “Methodology”.



4.4.1 Inception phase

UNIDO’s guidance on evaluation inception report⁵⁹ describes the inception phase as the point during the evaluation process where the evaluation team reviews project/programme documents, analyses the ToRs, and develops a detailed proposal for the implementation of the evaluation. The results are laid down in an inception report that is submitted to the UNIDO evaluation manager for review and approval. The continuation of the evaluation process is conditional on the UNIDO Independent Evaluation Unit’s approval of the report.

During the inception phase, the evaluation team can fine tune the evaluation methodology, further develop evaluation questions and focus further defined or reviewed, and sampling strategies and criteria clarified. The intervention logic or a theory of change (if developed in the project document) can be reconstructed during the inception phase.



PM to-do list:

- Prepare & support evaluation mission

The inception report contains, among other elements, the evaluation work plan with a concrete timetable, a project/programme logframe and the evaluation framework. The latter includes the evaluation criteria, related evaluation questions, indicators, sources for information and information about the methodology for data collection and analysis.

UNIDO provides detailed guidance on the format of an evaluation inception report, as shown in the box below.

SECTION INFO BOX: INTRANET



UNIDO guidelines for the preparation of inception reports

4.4.2 Quality assurance of inception report

Experience shows that the inception stage can be used to refine the evaluation to implement the ToR but also to deviate from the original ToR. The evaluation manager is required to ensure that the scope and objectives of the ToR are adhered to. If needed and justified, any deviations should be clearly explained. Once the inception report is approved by the

⁵⁹ UNIDO evaluation inception report: standard format, June 2023

evaluation manager, the evaluators move to the main phase of data collection.

Example of how an inception report can deviate from the original ToR

The ToR call for an impact evaluation of sector support in a given country.

However, the inception report does not mention the term “impact” in the methodology section of the inception report.

A proposed field visit to the country does not include meetings with the final beneficiaries of the intervention but with the sector ministry only.

A rationale for the omission of addressing impact in this “impact evaluation” is not given in the inception report.

In this real case example from a bilateral donor, the evaluation report failed to deliver on evaluating impact as expected after carefully reading the inception report. Quality assurance of the evaluation manager failed at inception.

4.4.3 Field work

Field work is an essential part of the evaluation for triangulating evaluation findings to validate and verify outputs, to observe and assess ownership and to get direct information and data from the object of evaluation.

It is important to be familiar with the country and intervention context. Carrying out field work can also allow the evaluation manager to obtain the latest inputs from partners, such as an updated list of intervention stakeholders, and brief partners about the evaluation methodology. Assessing the availability of data and identifying data sources are additional benefits from the preparatory work before the field mission. Experience during the COVID-19 pandemic led to further development and use of online tools and approaches in conducting evaluations and these should continue to be applied where possible.

The core objectives of the evaluation field mission are to:

- i) collect data on the ground to build the evidence trail;
- ii) validate and supplement desk review results;
- iii) interact with project stakeholders, including beneficiaries; and
- iv) gain direct observation of stakeholder behaviour and project results.

Field missions often face the challenge of time and budgetary constraints. As a result, it is even more important that project managers and evaluation managers ensure that the visit takes place at the right place with the right stakeholders. Visits to more remote project sites often allow for insights into the reality on the ground beyond capital cities.

In this context, it is important to remember that field visits aim to fill any information gaps and to validate existing information. They are not expected to assess all activities undertaken by the project or programme, nor are they expected to conduct large-scale household surveys.

To manage time and budget constraints common when carrying out field missions, it is recommendable to undertake case studies. It is also good practice to concentrate on components which require verification, or that are innovative or problematic.

At the end of the field visit, the evaluation team is required to debrief key stakeholders in

the field with emerging evaluation findings. This constitutes good practice and allows for onsite feedback and initial factual validation. The opportunity to meet with the evaluation team again tends to further increase the interest in and ownership of the evaluation results. Key field stakeholders in this case would include the project management team, governmental counterparts, Chief Technical Advisor (CTA), beneficiaries, and donor representatives.

4.4.4 Presentation of preliminary findings

After the field mission, and when data analysis and preparation of the draft report is at an advanced stage, it is a good practice to hold an internal debriefing at headquarters to allow the evaluation team to present preliminary evaluation findings, conclusions and recommendations. The debriefing should be attended by the PM, line managers of the related division and departments, and other staff working with the same or similar projects/programmes at UNIDO Headquarters in Vienna.



- PM to-do list:**
- Organize HQ debriefing
 - Debriefing of evaluators

It can take place before or after the evaluation team drafts the evaluation report and be carried out online or in-person. Each option has its own advantages and disadvantages and can be decided by the evaluation manager upon consultation with the key stakeholders. Sharing preliminary findings before finalizing the draft evaluation report helps to prepare the ground for the written report, clarifies misunderstandings, and avoids surprises later in the reporting stage and helps to create further ownership of the evaluation results.⁶⁰ Presenting the evaluation findings and recommendations after the draft report is shared with the project management team provides the evaluation team more time to internalize the information and data collected from the field, allowing the critical and analytical thinking to sink in and making the analysis more rigorous. At the same time, the project management team also has an opportunity to review the draft report carefully before discussing its findings and recommendations with a wider audience.

Lessons from the evaluations during the COVID-19 pandemic showed that such debriefings can be undertaken remotely in most cases.

“The debriefing session will be instrumental in ensuring ownership.”

Source: Vienna-based specialized UN agency.

4.4.5 Reporting

UNIDO’s Evaluation Policy specifies that evaluation reports provide complete and objective information on the methodology followed when carrying out independent evaluations, the limitations of the evaluation and key concerns, the evidence-based findings, conclusions and recommendations, as well as dissenting views, and agreed MAPs, if any. They are brief, to the point and easy to understand. They include an executive summary that encapsulates the essence of the information contained in the report. They may also include a chapter on lessons learned for wider applicability. Evaluation reports are prepared in accordance with the specific evaluation ToR and guid-

⁶⁰ UN Office on Drugs and Crime (UNODC), 2012, Evaluation Handbook.

ance provided by EIO. Draft reports are formally shared with key stakeholders for comments and factual validation. The evaluation team is responsible for the reliability and quality of the information contained in the report, which should reflect any factual corrections brought to the team’s attention prior to being finalized. In the event of differing views being expressed, these might be reflected in the analysis and in the report.⁶¹



PM to-do list:

- Share draft report with key stakeholders

EO to-do list:

- Feedback
- Clearance/ approval
- QA check, compliance with UNIDO standard format
- Clearance for final payment to consultants

PM to-do list:

- Final payment to consultants

Once all data are captured and analyzed, reporting can start based on the agreed reporting outline. The reporting outline is specified either in UNIDO’s evaluation ToR, or in the evaluators’ inception report. The reporting outline is helpful for organizing and focusing evidence found in the evaluation. From a practical point of view, it can also assist in distributing responsibilities for specific sections of the report to the evaluation team members.⁶² The report must comply with the UNIDO Quality Standards for Evaluation Reports.⁶³

The responsibility for delivering a draft report on time according to the timeframe agreed in the inception report lies with the evaluation TL. EIO is responsible for ensuring that the draft report is shared (normally through the project manager) with key stakeholders in-house, and among relevant project stakeholders and partners for factual validation and feedback. Feedback and comments received are sent to the evaluation team leader.

The evaluation officer submits EIO feedback on the draft report to the evaluation team leader on time.

As a next step, the evaluation team leader is responsible for addressing all the feedback received. Due to the independent nature of UNIDO independent evaluations the feedback is not binding for the evaluation TL with the exception of factual errors, which have to be addressed. To uphold transparency, it is good practice to keep a log of how the feedback was considered in the final report, which again comes under the responsibility of the evaluation TL. The TL is also responsible for editing and formatting the final report in line with the specified report structure.

The final report (presented in one MS Word document, including all annexes) is submitted by the TL to the evaluation project manager for final review and for EIO to conduct the final quality assurance. EIO will proceed with the final document preparation (cover page, registry), final adjustments as needed, and preparation of the e-book for formal distribution and dissemination.

Reporting: good practices

Experience shows that evaluation reports often benefit from significant efforts during data collection and analysis but face challenges at the time of reporting. This is particularly true when evaluators are faced with identifying key findings among the main emerging evaluation findings, followed by the need to draw conclusions which in turn

⁶¹ UNIDO, 2021: Evaluation Policy, page 10.
⁶² A good practice stipulated by World Intellectual Property Organization (WIPO), 2010: WIPO Independent Evaluation Guidelines.
⁶³ UNEG, 2010, UNEG Quality Checklist for Evaluation Reports.

lead to recommendations.

It is good practice to identify and summarize key findings for each evaluation criterion at the beginning of each relevant section in the evaluation report. After the findings section, the report could contain a table with three main columns listing the key findings by evaluation criterion answering the main evaluation questions, followed by conclusions and the recommendations. Based on good practice in previous UNIDO evaluation reports, the use of such a table greatly supports evaluators in logically linking key evaluation findings with conclusions and recommendations.

Section 5.2.5 provides detailed guidance on reporting for strategic evaluations. The guidance is also valid for project evaluations.

SECTION INFO BOX: INTRANET



HYPERLINK

[UNIDO checklist on evaluation report quality](#)
(UNIDO Evaluation tools. Guidelines for Terms of Reference for evaluations: Annex 2)

4.5 Follow-up, learning, and dissemination

4.5.1 Management response to evaluations and internal follow-up

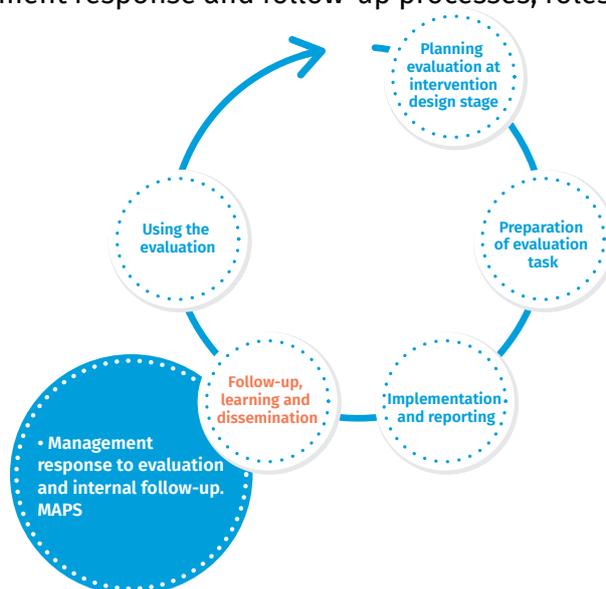
The UNIDO Evaluation Policy defines management response and follow-up processes, roles, and responsibilities by evaluation type.⁶⁴

A timely management response is required in line with the indicated deadlines for the follow-up process.

For independent evaluations, UNIDO Evaluation Policy outlines that:⁶⁵

For all independent project/programme evaluations, EIO issues the evaluation reports to relevant project managers, senior management, and/or the Director General as needed, together with a management response sheet (MRS). The MRS is assigned to a relevant UNIDO manager to enable tracking for each recommendation. Items tracked include comments of acceptance, partial or non-acceptance of evaluation recommendations, agreed management action plans, deadlines for implementation, and information on the actions taken to address recommendations.

UNIDO line managers ensure that those responsible for follow-up keep information in each MRS up to date.



⁶⁴ UNIDO, 2021, Evaluation Policy, page 10.

⁶⁵ UNIDO, 2021, Evaluation Policy, pages 10-11.

EIO monitors the information in the MRSs and compiles periodic reports on the level of acceptance of recommendations and on the status of implementation of the follow-up activities related to individual recommendations, and draws attention of UNIDO management to that status.



EO to-do list:

- Final formatting & MRS
- Submission of evaluation report & MRS to: PM & stake-holders, MRS respondents
- MAPS

PM to-do list:

- Implement & follow-up recommendations in MRS

For all strategic evaluations, the process now includes MAPs, formulated by the responsible manager/business owner, to address the recommendations and findings from the evaluation. EIO validates and clears the proposed MAPs.

Each MAP should specify appropriate actions that would address the identified issues and which would improve the process or processes in question. Each MAP should include specific verifiable deliverables or actions, and deadlines.

Once agreed between EIO and the responsible manager(s), MAPs are binding. It is the responsibility of managers who have agreed to a MAP to implement the MAP as soon as possible and within the specified deadline. Any amendment to a MAP is also subject to review/agreement between EIO and the responsible business owner.

EIO maintains an electronic management response system available on the evaluation's site on the UNIDO Intranet through which the response status can be tracked.

4.5.2 Learning and dissemination

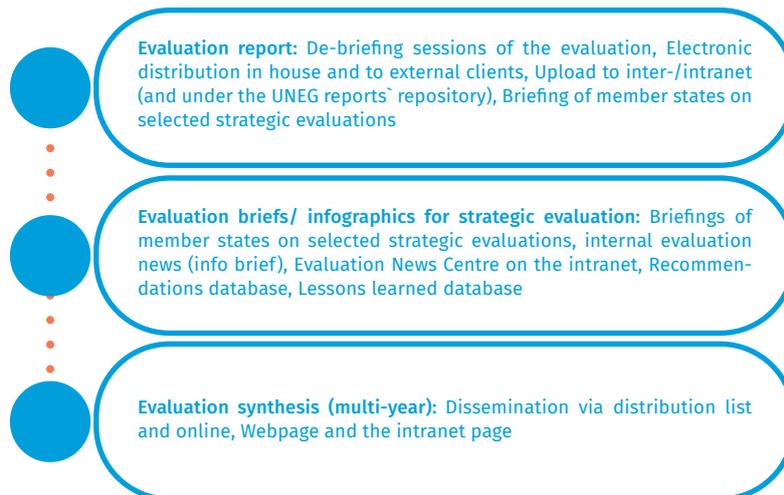
Learning

“Learning is not attained by chance, it must be sought for with ardour and attended to with diligence.”

Source: Abigail Adams, 2nd First Lady of the United States of America (1797–1801) on May 8, 1780 in a letter to her son John Quincy Adams, 6th President of the United States of America (1825 – 1829).

The UNIDO evaluation function uses three main learning products and processes, as presented in the graphic below: evaluation reports, evaluation briefs/infographics for strategic evaluations, and multi-year evaluation synthesis.

Dissemination



The UNIDO Evaluation Policy specifies that the biennial evaluation work plan, the ToR of independent evaluations and all evaluation reports are posted on the relevant pages of the UNIDO public website.⁶⁶ Other dissemination processes are listed in the previous section.

⁶⁶ UNIDO, 2021, Evaluation Policy, page 11.

5. The process for strategic evaluations: Country, thematic, corporate-level and impact evaluations

In the final section of the Evaluation Manual, relevant information is provided on strategic evaluations. The section clarifies objectives and methodology, describes the process and specifies reporting requirements.

5.1 Background

The UNIDO Evaluation Policy describes strategic evaluations as follows: “Independent corporate-level evaluations or reviews, such as country-level, thematic or strategic evaluations, are conducted to assess country or regional programmes, policies, strategies and cross-cutting issues or functions. They inform UNIDO management and external stakeholders on policymaking and organizational effectiveness.

The strategic evaluation process is effectively equivalent to the independent project evaluation process. However, it differs from it in two main respects: i) in terms of scope, it covers organizational priorities and issues, strategic areas, cross-cutting or thematic topics, key processes or country-level programmes, and ii) in evaluation planning, which for strategic evaluations is rooted in a biennial EIO evaluation work plan.

As regards preparation, conduct and reporting, strategic evaluations follow the workflow presented below.



5.1.1 Selection and timing.

Strategic/corporate-level evaluations are identified and prioritized by EIO through a periodic evidence-gap analysis, which considers, inter alia, organizational management priorities, challenges and/or requests emanating from UNIDO management or policymaking organs”.⁶⁷

An evidence gap assessment is carried out biannually.

The degree to which evidence from previous evaluations or external/independent assessment is documented and available.

Evidence gap

The degree to which it can be expected that evaluation results will be used by relevant management and business owners.

Likelihood of use of evaluation results

The degree to which the evaluation can be delivered on time to feed into decision making processes.

Timeliness

The degree to which an evaluation can draw on sufficient robust evidence within the budget constraints.

Evaluability

For country-level and thematic evaluation also the resources allocated or the technical cooperation portfolio volume is taken into account.

Portfolio size

Figure 20: Planning and selection of strategic evaluations

⁶⁷ UNIDO, 2021, Evaluation Policy, page 5.

The selection of a particular evaluation should be guided by the extent to which the project topics are relevant to the policy agenda of UNIDO, partner governments and donors. It should also be based on whether evaluative evidence is lacking for decision-making about future strategic decisions in the area under consideration. The latter also influences the timing of these evaluations.

The Independent Thematic Review of UNIDO Interventions in the Area of Enterprise Development for Job Creation, including for Women and Youth (2015), for example, states that the evaluation exercise was commissioned “to feed into the development and implementation of the Inclusive and Sustainable Industrial Development vision and strategy”.⁶⁸

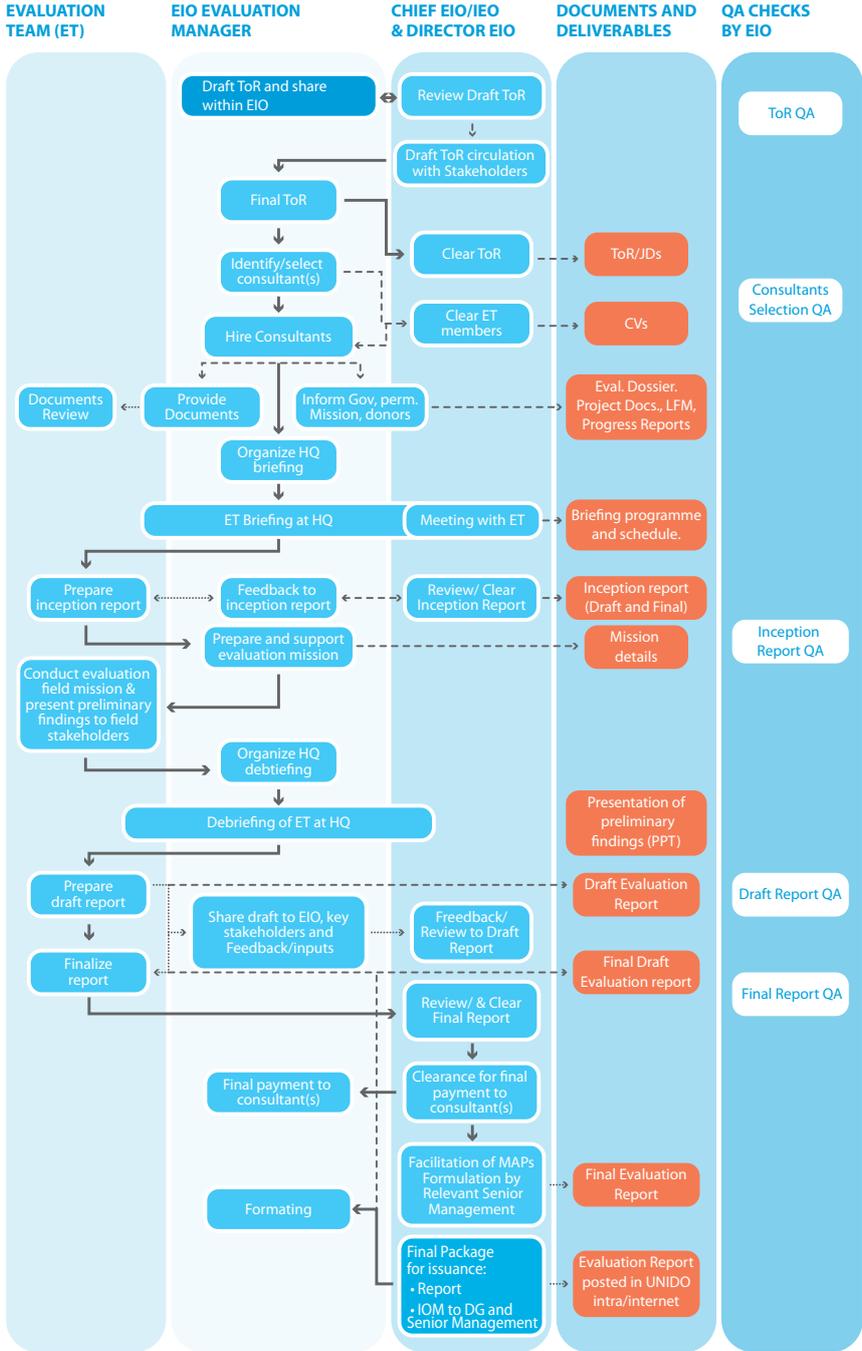


Figure 21: Flowchart for Strategic evaluations

⁶⁸ UNIDO Interventions in the Area of Enterprise Development for Job Creation, including for Women and Youth (2015), page 1.

5.2 Country-level evaluations

5.2.1 Purpose

UNIDO's country-level evaluations aim to provide an assessment of the Organization's aggregated country engagement, specific results and impact at country level, and its contribution to global transformational change in creating shared prosperity, advancing economic competitiveness and safeguarding the environment. Country evaluations are also forward-looking by generating findings and recommendations to shape the future partnership between UNIDO and the country concerned, along with any other country partners.

In this context, CP or PCP evaluations fall into the category of country-level evaluations. Country-level evaluations need increasingly to consider UNIDO's role as part of the wider UN Country Team (UNCT) in a country. Consequently, they should give more prominence to UNIDO's contribution to the Sustainable Development Cooperation Framework (UNSDCF) [or former UN Development Assistance Framework (UNDAF)]. At the same time, results from UNIDO country-level evaluations can feed into UN country framework evaluations by highlighting UNIDO's contribution to the Sustainable Development Cooperation Framework and to transformational changes at the country and global level.

5.2.2 Scope

In addition to assessing projects and programmes, which are the mainstay of UNIDO country engagements, country-level evaluations also focus on country-level specific results such as: i) UNIDO's analytical and policy advisory services; ii) standards setting and compliance; and iii) UNIDO's convening and partnership role.

Good practice example of UNIDO country evaluation: Scope

“As a Country Programme evaluation, the main focus was less on the performance of the specific projects but rather on the question: to what does it all add up, what difference did the set of interventions make to Tanzania?”

Still, projects constituted to building stones of the Country Programme. The evaluation team decided to group the projects in three thematic clusters:

- 1st cluster: Policy, national systems, statistics, trade
- 2nd cluster: Value-chain development, industrial upgrading, entrepreneurship
- 3rd cluster: Environment and energy
- Additional area: UNIDO office and UNIDO as part of Delivering as One”

Source: UNIDO, 2016, Independent UNIDO Country Evaluation: United Republic of Tanzania, page 5.

Ideally, evaluations should cover UNIDO's cooperation in a country over a period of six to 10 years. This allows sufficient time to build a critical mass of evidence to track anticipated changes resulting from a UNIDO intervention.

Country evaluations also represent a unique opportunity to assess the sustainability of previous UNIDO interventions and to verify to what extent national ownership and project results have created lasting impact.

The added value of country evaluations for UNIDO goes beyond summarizing project-level results. An example can be seen in the case of the Pakistan country evaluation (2014), as presented in the box below.

Added value of country evaluations

“The evaluation also assessed the performance of UNIDO’s Field Office with regards to its contribution to development results and through performing convening, normative and technical cooperation functions as well as the efficiency and effectiveness of the office in managing, coordinating and implementing projects and programmes.”

Source: UNIDO Evaluation Group, 2014, Independent UNIDO Country Evaluation Pakistan, page 10.

■ **5.2.3 Process and methodology**

The main purpose of country evaluations is to assess how UNIDO has contributed to national development programmes, and to countries’ industrial development strategies. It is important to evaluate against expected results. Project targets, with baselines and indicators, are one means by which to assess UNIDO’s contribution to countries’ industrial development strategies.

In the case of country-level evaluations, the logframe and/or TOC for CPs or PCPs serves as the main reference point for evaluators with its upstream focus on national industrialization agenda targets.

The process for country evaluations is theory-based, in line with good practice across the UN System. This allows building on a strong context analysis and show how the country engagement aimed to lead to results and impact, based on explicit assumptions. Section 4 of this Evaluation Manual explains the ToR approach in more detail.

Evaluations should be guided by their respective ToRs and inception reports, which should include a table indicating evaluation questions, information sources and methods for data collection and analysis. This forms the basis of the evaluation questionnaire. The evaluation framework can be used to define particular roles within the evaluation team, such as reporting or field visits.

A ToR template is available for on UNIDO Evaluation website. This can be customized to suit CP, PCP or any other type of country evaluation.

Project evaluation results for a specific country are aggregated as far as possible to include ratings to identify patterns and trends by clusters. As highlighted above, evaluating against expected results is important and has to date been one of the weaker areas of previous country evaluations.

Country evaluations should establish linkages to other UNIDO core functions at the country level where possible. In order to keep the evaluation focused and affordable, case studies could be carried out outlining the processes and results of UNIDO’s analytical and policy advisory services, standard-setting and compliance, and its convening and partnership role.

Figure 22 below summarizes common challenges faced in UNIDO country evaluations and lists possible solutions.

<ul style="list-style-type: none"> • Lack of documentation, including baseline data and monitoring data 	<ul style="list-style-type: none"> • UNIDO Independent Evaluation Unit to coordinate selection of project and country evaluations to increase the evaluative evidence for the given country
<ul style="list-style-type: none"> • Insufficient availability of evaluative evidence 	
<ul style="list-style-type: none"> • Large number of projects to be evaluated (average of 10) during 3-week field visit if no cluster approach is taken 	<ul style="list-style-type: none"> • i) Using the good practice of clustering projects thematically; ii) sampling within a cluster; iii) focus sampling on projects with no evaluative evidence
<ul style="list-style-type: none"> • Difficulties in creating a single theory of change for a CP 	<ul style="list-style-type: none"> • Unless a CP exists, it might not be possible to create a theory of change for the entire project portfolio. However, creating theories of change by cluster would still provide some benefit from using a theory-based evaluation approach

Figure 22: Common challenges in UNIDO country evaluations and possible solutions

5.3 Thematic evaluations

5.3.1 Purpose

Thematic evaluations aim to provide timely, credible and useful information on UNIDO’s thematic, programmatic and strategic dimensions, as well as enabling evidence-based decision-making.

5.3.2 Scope

UNIDO’s thematic evaluations address systemic issues and cut across projects/programmes, UNIDO core functions, countries and regions.

Thematic evaluation assesses the results of: i) UNIDO’s analytical and policy advisory services; ii) standard-setting and compliance; iii) UNIDO’s convening and partnership role on a specific theme; and iv) other specific technical themes or priorities that the Organization has been working on. This is complemented by relevant evaluative evidence from TC projects and past project/programme and corporate-level evaluations.

5.3.3 Process and methodology

The main focus points of thematic evaluations are UNIDO’s policies, its thematic strategies, or the thematic strategies of regional bodies and UNIDO partners.

As in the case of country evaluations, UNIDO’s thematic evaluations are theory-based. They should be based on a logic model, theory of change, and/or the expected results chain. If unavailable, these elements should be reconstructed based on the strategic documentation available and interviews with key stakeholders. As far as is possible, evaluators should aggregate existing evaluative evidence from thematically relevant project evaluations. A good practice example for aggregation emerges when a meta-evaluation approach is taken. The Independent Thematic Evaluation: “Post-Crisis Interventions” (2015) rigorously analyzed available project evaluation results along the lines of evaluation criteria, rather than opting for proliferating the analysis by project themes. The use of case studies can help to assess the performance of other UNIDO core functions, beyond TC projects.

Thematic evaluations in UNIDO usually rely primarily on past project evaluations. Given budget and time constraints, primary data collections related to TC projects through field visits are the exception rather than the rule.

Figure 23 below summarizes common challenges encountered in UNIDO thematic evaluations.⁶⁹ Suggestions on how those challenges can be addressed are also listed.

<ul style="list-style-type: none"> • Very broad subject to assess; concept and as an operational category is open to different interpretations or no formally adopted strategy or targets adopted 	<ul style="list-style-type: none"> • Systematic evaluability assessments required to judge feasibility of thematic evaluation • Use evaluation as an opportunity to reach consensus on the definition of the topic to be evaluated (e.g. for the evaluation ToR or approach paper)
<ul style="list-style-type: none"> • Gaps in information, no availability of monitoring data 	<ul style="list-style-type: none"> • UNIDO Independent Evaluation Unit to coordinate selection of project and thematic evaluations to increase the evaluative evidence for the given theme
<ul style="list-style-type: none"> • Evidence of accomplishments from secondary sources; i.e. no primary data collection in the field • Purposive sample of evaluated projects with “outdated” data 	<ul style="list-style-type: none"> • Primary data collection through surveys or interviews for up-to-date and first-hand information • Include results of o-going evaluations •Additional stocktaking and feedback workshop •Development of additional thematic working documents to be developed by the evaluation team •Use of benchmark/comparator analysis •If triangulations still remain limited, qualify and label assessment as ‘review’ and not as ‘evaluation’ as originally envisaged
<ul style="list-style-type: none"> • Amount of project field visits limited ability of the evaluation to assess fully the progress towards results / impacts 	<ul style="list-style-type: none"> • Focus evaluation on a portfolio of projects rather than in-depth findings for each project to get an overall picture •Select case studies based on the criterion of most significant change/ least significant change for targeted selection for country visits
<ul style="list-style-type: none"> • ERP system did not allow a clear-cut allocation of resources to the topic under evaluation 	<ul style="list-style-type: none"> • Review whether evaluation criteria of efficiency can be fully assessed. If not state this limitation transparently in the methodology section of the evaluation report

Figure 23: Common challenges in UNIDO thematic evaluations and possible solutions

For the verification of secondary data, the use of evaluation questionnaires in interviews or surveys is recommended. UNIDO thematic evaluations consult stakeholders. In that case evaluation questionnaires should be used systematically. Questionnaire results need to be appropriately reflected in the evaluation report to evidence triangulation. This can be accomplished by using quotes (while maintaining the anonymity of the interviewee), using questions with ratings or questions allowing for rankings. Qualitative data captured through the evaluation questionnaire can also be quantified. Unless interview results are used stra-

⁶⁹ Between 2012 and 2016

tegitally for data collection, the evaluation's evidence base won't be sufficiently strengthened.

Using comparators from other similar organizations to benchmark UNIDO's strategic approach and performance for specific thematic issues has been shown to strengthen the evidence base. The Independent Strategic Evaluation "Implementation of the Expanded UNIDO Medium-Term Programme Framework 2010-2013" (2015) and the Independent Thematic Evaluation: "Post-Crisis Interventions" (2015) are good examples in this respect.

5.4 Impact evaluations

Multilateral organizations are facing increasing demands to demonstrate results, in particular to show evidence of impact from development interventions. While UNIDO's management is responsible for evidence-based monitoring and reporting (mainly thorough systematic evidence-based monitoring and reporting), UNIDO's evaluation function play an important role in providing assurance on outcome, impact and transformational changes. (The challenges involved in evaluating impact and impact evaluation domains and techniques are described above in Section 3: Methodology)

5.4.1 Objectives

The objective of an impact evaluation is to assess whether UNIDO programmes or interventions have produced, or clearly contributed to, broader and sustainable change, and whether that change has had a lasting effect on the target population. In other words, the impact evaluation seeks to examine to what extent UNIDO intervention, or interventions, have created shared prosperity, advanced economic competitiveness, and/or safeguarded the environment.

5.4.2 Purpose

Impact evaluations in UNIDO focus on the impact criterion and assess the positive and negative, and intended or unintended attributable effects of an UNIDO intervention.

5.4.3 Scope

The unit of assessment is typically a programme (thematically and/or geographically, such as country). Single projects are rarely assessed unless they consist of several phases. The programme examined should be of a mature nature in order to allow sufficient time for change to have taken place. That means that impact evaluation should ideally focus on projects or programmes in the second or third phase of operation. Alternatively, impact evaluation should be carried out ex post, up to 10 years after the end of a UNIDO intervention. However, after such a time-lag tracing stakeholders and records/data may be more difficult.

5.4.4 Challenges of impact evaluations

Impact evaluations are rooted on the capacity to collect, validate and assess evidence and data to verify whether transformational, sustainable changes have occurred as a consequence of an intervention.

Impact evaluations should be prepared and planned together at the programme formulation (or portfolio or sequence of projects) stage, since the data necessary for impact evaluation has to be collected during the project implementation, by its monitoring and reporting system. SMART indicators and targets, at outcome and impact level, should also be clearly defined at the design and formulation stages of any programme or project.

■ 5.4.5 Process and methodology

There is no recipe or single methodology for conducting impact evaluations, with a variety of methods under discussion in the international development and evaluation community in recent years.

In the 2000s, randomized control trials (RCT) and other experimental and quasi-experimental approaches were considered the “gold standard” of impact evaluations. However, in practice they faced significant limitations in complex social and economic settings and in relation to budget. Methodological limitations of RCT and other experimental and quasi-experimental approaches include:

- Ensuring the representativeness of the sample size is challenging, as experimental and quasi-experimental approaches require large sample sizes for both the treatment group and the control group.
- The requirement for multiple sites in the partner country (for example factories across distinct geographic areas) to ensure validity is difficult to fulfill due to management challenges.
- The long trial run time may result in the loss of relevance as practice may have moved on by the time the evaluation is finalized.

Other specialized UN agencies and development organizations have faced several implementation challenges, combined with the burden of exploding budgets and the demand for vast amounts of data and long timeframes for this type of impact evaluation. UNIDO adopts a rigorous, pragmatic, and credible approach for impact evaluations that uses a mix of quantitative and qualitative methods. At the same time, the methodology takes account of the overall level of resources available within the Independent Evaluation Unit.

UNIDO uses cost-effective techniques for its impact evaluations, for example, a combination of counterfactual analysis (e.g. using control groups where possible), “before and after” techniques, case studies, broader theory of change, and triangulation methods, based on good practices carried out across UN specialized agencies.

Theory-based impact evaluations

A particularly cost-effective way to carry out impact evaluations is to use a theory-based approach, as successfully applied by one of UNIDO’s main partners, the GEF.⁷⁰ The approach entails collecting evidence about impact and the extent to which key assumptions listed in the theory of change hold. UNIDO’s contribution to changes in policy, legal or regulatory frameworks should be assessed. The latter includes observed changes in counterparts’ capacities (for example knowledge, attitude, practice, infrastructure or performance systems)

⁷⁰ Global Environment Facility, 2017, Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects.

and governance architecture such as the access to and use of information (for example laws, conflict resolution processes, document repositories or knowledge management systems).

To document changes being achieved at scale in sustainable industrial development beyond UNIDO's intervention area, the impact evaluation should record and analyze evidence of sustaining, mainstreaming, replication, scaling up and market change. In this context, it is important to identify to what extent UNIDO promoted approaches, frameworks, systems or policies that were adopted or implemented without direct support or involvement from UNIDO. This can provide evidence of progress towards impact. Since the publication of the first edition of the Evaluation Manual in 2018, UNIDO's Independent Evaluation Unit has applied the theory-based approach in many evaluations and has successfully demonstrated the contribution of some of the interventions to broader changes in the system and to long-term transformational changes.

An important dimension of assessing progress towards impact is behaviour change resulting to change in practices that are:

- i) Environmentally sound – Safeguarding environment
- ii) Economically competitive – Advancing economic competitiveness
- iii) Socially inclusive – Creating shared prosperity

Annex 2 contains sample questions to address the progress on impact of UNIDO interventions.

When analyzing contributions of UNIDO to observed change processes, the evaluation should also assess the contribution of other actors and factors. The contribution analysis includes assessing benefits from rival explanations of the observed change and provides a rationale for accepting or rejecting those explanations.

As in the case of the GEF, the broader adoption of UNIDO-promoted approaches, and/or technologies, typically take place through mainstreaming, replication, scaling-up and market change. The following definition applies for those dimensions:

- **Mainstreaming:** *Information, lessons learned, or specific results from UNIDO are incorporated into broader stakeholder mandates and initiatives such as laws, policies, regulations, and programmes. This may occur through governments and/or through development organizations and other sectors.*
- **Replication:** *UNIDO-supported initiatives are reproduced or adopted at a comparable administrative or ecological scale, often in another geographical area or region.*
- **Scaling-up:** *UNIDO-supported initiatives are implemented at larger geographical scale, often expanded to include new aspects or concerns that may be political, administrative or ecological in nature.*
- **Market change:** *UNIDO-supported initiatives help catalyze market transformation by influencing the supply of, and/or demand for, goods and services that contribute to global environmental, economic and social benefits. This may encompass technological changes, policy and regulatory reforms, and financial instruments.*

Annex 1: References to good evaluation practices



UNIDO Evaluation Policy

DAC RBM glossary

Checklist for evaluation terms of reference

United Nations Evaluation Group, 2010: UNEG checklist for evaluation terms of reference and inception reports. Guidance document

Gender

UNIDO gender guidelines

UNEG guidance on evaluating gender and human rights

Annex 2: Definition of evaluation criteria including key evaluation questions

#	EVALUATION CRITERIA	DEFINITION	MANDATORY RATING
A	Progress to impact	Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended, including redirecting trajectories of transformational process and the extent to which conditions for trajectory change are being put into place.	Yes
		<ul style="list-style-type: none"> • Mainstreaming: To what extent are information, lessons learned, or specific results of the project incorporated into broader stakeholder mandates and initiatives such as laws, policies, regulations and project? 	
		<ul style="list-style-type: none"> • Replication: To what extent are the project's specific results (for example methodology, technology or lessons learned) reproduced or adopted? 	
		<ul style="list-style-type: none"> • Scaling-up: To what extent are the project's initiatives and results implemented at larger geographical scale? 	
		<ul style="list-style-type: none"> • What difference has the project made to the beneficiaries? 	
		<ul style="list-style-type: none"> • What is the change attributable to the project? To what extent? 	
		<ul style="list-style-type: none"> • What are the social, economic, environmental and other effects, either short-, medium- or long-term, on a micro- or macro-level? 	
		<ul style="list-style-type: none"> • What effects are intended or unintended, positive or negative? 	
		The three UNIDO impact dimensions are:	
		<ul style="list-style-type: none"> • Safeguarding environment: Biophysical changes in reduction of threats emanating from action of humans and changes in the status of the environment. 	
		<ul style="list-style-type: none"> • Economic performance: Changes in the functioning and management of the resources, finances, income, and expenditure of, for example, a community, business or enterprise, contributed to by the intervention 	
<ul style="list-style-type: none"> • Social inclusiveness: Changes in the provision of certain rights to all individuals and groups in society, such as employment, education, and training. 			
B	Project design	Formulation of the intervention, the plan to achieve a specific purpose.	Yes
1	Overall design	Assessment of the design in general	Yes
		<ul style="list-style-type: none"> • Is the problem, need or gap to be addressed by the project clearly identified, with clear target beneficiaries? 	
		<ul style="list-style-type: none"> • Was the project design adequate to address the problems at hand? 	
		<ul style="list-style-type: none"> • Is the project consistent with the country's priorities, in the work plan of the lead national counterpart? Does it meet the needs of the target group? Is it consistent with UNIDO's Inclusive and Sustainable Industrial Development? Does it adequately reflect lessons learnt from past projects? Is it in line with the donor's priorities and policies? 	
<ul style="list-style-type: none"> • Is the applied project approach sound and appropriate? Is the design technically feasible and based on best practices? Does UNIDO have in-house technical expertise and experience for this type of intervention? 			

		<ul style="list-style-type: none"> • To what extent is the project design (in terms of funding, institutional arrangement, implementation arrangements...) as foreseen in the project document still valid and relevant? 	
		<ul style="list-style-type: none"> • Does it include M&E plan and adequate budget for M&E activities? 	
		<ul style="list-style-type: none"> • Risk management: Are critical risks related to financial, sociopolitical, institutional, environmental and implementation aspects identified with specific risk ratings? Are their mitigation measures identified? Where possible, are the mitigation measures included in project activities/outputs and monitored under the M&E plan? • To what extent does the project design contribute to gender equality, the empowerment of women and the human rights-based approach? 	
2	Logframe	Assessment of the logical framework aimed at planning the intervention.	Yes
		<ul style="list-style-type: none"> • Expected results: Is the expected result-chain (impact, outcomes and outputs) clear and logical? Does impact describe a desired long-term change or benefit to a society or community (not as a mean or process), do outcomes describe change in target group's behaviour/performance or system/institutional performance, do outputs describe deliverables that project will produce to achieve outcomes? Are the expected results realistic, measurable and not a reformulation or summary of lower-level results? Do outputs plus assumptions lead to outcomes, do outcomes plus assumptions lead to impact? Can all outputs be delivered by the project, are outcomes outside UNIDO's control but within its influence? 	
		<ul style="list-style-type: none"> • Indicators: Do indicators describe and specify expected results (impact, outcomes and outputs) in terms of quantity, quality and time? Do indicators change at each level of results and independent from indicators at higher and lower levels? Do indicators not restate expected results and not cause them? Are indicators necessary and sufficient and do they provide enough triangulation (cross-checking)? Are indicators sex-disaggregated, if applicable? Are indicators SMART? 	
		<ul style="list-style-type: none"> • Sources of verification: Are the sources of verification/data able to verify status of indicators, are they cost-effective and reliable? Are the sources of verification/data able to verify status of output and outcome indicators before project completion? 	
		<ul style="list-style-type: none"> • Are key assumptions properly summarized and reflecting the proper level in the results chain in the logframe? 	
C	Project performance	Functioning of a development intervention	Yes
1	Relevance	The extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor.	Yes
		<ul style="list-style-type: none"> • How does the project fulfil the urgent target group needs? 	
		<ul style="list-style-type: none"> • To what extent is the project aligned with the development priorities of the country (national poverty reduction strategy, sector development strategy)? 	
		<ul style="list-style-type: none"> • How does the project reflect donor policies and priorities? 	
		<ul style="list-style-type: none"> • Is the project a technically adequate solution to the development problem? Does it eliminate the cause of the problem? 	
		<ul style="list-style-type: none"> • To what extent does the project correspond to UNIDO's comparative advantages? 	
		<ul style="list-style-type: none"> • Are the original project objectives (expected results) still valid and pertinent to the target groups? If not, have they been revised? Are the revised objectives still valid in today's context? 	

2	Coherence	The compatibility of the intervention with other interventions in a country, sector or institution. The extent to which other interventions (particularly policies) support or undermine the intervention, and vice versa.	Yes
		<ul style="list-style-type: none"> • How consistent is the intervention with the relevant international norms and standards to which the supported institution/government adheres? 	
		<ul style="list-style-type: none"> • How consistent is the intervention with other actors' interventions in the same context? 	
		<ul style="list-style-type: none"> • How does the intervention ensure complementarity, harmonisation and coordination with others? 	
		<ul style="list-style-type: none"> • To what extent is the intervention adding value while avoiding duplication of effort? • To what extent did UNIDO adopt gender-sensitive, human rights-based approaches? 	
3	Effectiveness	The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.	Yes
		<ul style="list-style-type: none"> • What are the main results (mainly outputs and outcomes) of the project? What have been the quantifiable results of the project? 	
		<ul style="list-style-type: none"> • To what extent did the project achieve its objectives (outputs and outcomes), against the original/revised target(s)? 	
		<ul style="list-style-type: none"> • What are the reasons for the achievement/non-achievement of the project objectives? 	
		<ul style="list-style-type: none"> • What is the quality of the results? How do the stakeholders perceive them? What is the feedback of the beneficiaries and the stakeholders on the project effectiveness? 	
		<ul style="list-style-type: none"> • To what extent is the identified progress result of the project attributable to the intervention rather than to external factors? 	
		<ul style="list-style-type: none"> • What can be done to make the project more effective? 	
		<ul style="list-style-type: none"> • Were the right target groups reached? 	
4	Efficiency	A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results.	Yes
		<ul style="list-style-type: none"> • How economically are the project resources/inputs (concerning funding, expertise, time...) being used to produce results? 	
		<ul style="list-style-type: none"> • To what extent were expected results achieved within the original budget and timeframe? If no, please explain why. 	
		<ul style="list-style-type: none"> • Are the results being achieved at an acceptable cost? Would alternative approaches accomplish the same results at less cost? 	
		<ul style="list-style-type: none"> • What measures have been taken during planning and implementation to ensure that resources are efficiently used? Were the project expenditures in line with budgets? 	
		<ul style="list-style-type: none"> • Could more have been achieved with the same input? 	
		<ul style="list-style-type: none"> • Could the same have been achieved with less input? 	
		<ul style="list-style-type: none"> • How timely was the project in producing outputs and outcomes? Comment on the delay or acceleration of the project's implementation period. • To what extent were the project's activities in line with the schedule of activities as defined by the project team and annual work plans? 	

		<ul style="list-style-type: none"> • Have the inputs from the donor, UNIDO and government/counterpart been provided as planned, and were they adequate to meet the requirements? • To what extent did UNIDO promote gender equality, the empowerment of women, and human rights the delivery of project outputs? 	
5	Sustainability of benefits	The continuation of benefits from a development intervention after major development assistance has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time.	Yes
		<ul style="list-style-type: none"> • Will the project results and benefits be sustained after the end of donor funding? 	
		<ul style="list-style-type: none"> • Does the project have an exit strategy? 	
		<ul style="list-style-type: none"> • To what extent have the outputs and results been institutionalized? 	
	<i>Financial risks:</i>	<ul style="list-style-type: none"> • What is the likelihood of financial and economic resources not being available once the project ends? 	
	<i>Socio-political risks:</i>	<ul style="list-style-type: none"> • Are there any social or political risks that may jeopardize the sustainability of project outcomes? 	
		<ul style="list-style-type: none"> • What is the risk that the level of stakeholder ownership (including ownership by governments and other key stakeholders) will be insufficient to allow for the project outcomes/benefits to be sustained? 	
		<ul style="list-style-type: none"> • Do the various key stakeholders see that it is in their interest that project benefits continue to flow? 	
		<ul style="list-style-type: none"> • Is there sufficient public/stakeholder awareness in support of the project's long-term objectives? • To what extent do mechanisms, procedures and policies exist to carry forward the results attained on gender equality, empowerment of women, and human rights by primary stakeholders? 	
D	Cross-cutting performance criteria	Other important criteria that cut across the UNIDO intervention.	
1	Gender mainstreaming	Is the gender marker assigned to this project representative of reality?	Yes
		<ul style="list-style-type: none"> • Was a gender analysis included in a baseline study or needs assessment (if any)? Were there gender-related project indicators? 	
		<ul style="list-style-type: none"> • Are women/gender-focused groups, associations or gender units in partner organizations consulted/ included in the project? 	
		<ul style="list-style-type: none"> • How gender-balanced was the composition of the project management team, the Project Steering Committee (PSC), experts and consultants and the beneficiaries? 	
		<ul style="list-style-type: none"> • Do the results affect women and men differently? If so, why and how? How are the results likely to affect gender relations (e.g., division of labour, decision-making authority)? 	
		<ul style="list-style-type: none"> • To what extent were socioeconomic benefits delivered taking into consideration the gender dimensions? 	
2	M&E	Refers to all the indicators, tools and processes used to measure if a development intervention has been implemented according to the plan (monitoring) and is having the desired result (evaluation).	Yes
	<i>M&E at design</i>	<ul style="list-style-type: none"> ° Was the M&E plan included in the project document? Was it practical and sufficient at the point of project approval? 	
		<ul style="list-style-type: none"> ° Did it include baseline data and specify clear targets and appropriate indicators to track environmental, gender, and socioeconomic results? 	

		<ul style="list-style-type: none"> ◦ Did it include a proper M&E methodological approach; specify practical organization and logistics of the M&E activities including schedule and responsibilities for data collection? 	
		<ul style="list-style-type: none"> ◦ Does the M&E plan specify what, who and how frequent monitoring, review, evaluations and data collection will take place? Is the M&E plan consistent with the logframe (especially indicators and sources of verification)? 	
		<ul style="list-style-type: none"> ◦ Does it allocate adequate budget for M&E activities? 	
	<i>M&E at implementation</i>	<ul style="list-style-type: none"> ◦ How was the information from M&E system used during the project implementation? Was an M&E system in place and did it facilitate timely tracking of progress toward project results by collecting information on selected indicators continually throughout the project implementation period? Did project team and manager make decisions and corrective actions based on analysis from M&E system and based on results achieved? 	
		<ul style="list-style-type: none"> ◦ Are annual/progress project reports complete, accurate and timely? 	
		<ul style="list-style-type: none"> ◦ Was the information provided by the M&E system used to improve performance and adapt to changing needs? Was information on project performance and results achievement being presented to the PSC to make decisions and corrective actions? Do the project team and managers and PSC regularly ask for performance and results information? 	
		<ul style="list-style-type: none"> ◦ Are monitoring and self-evaluation carried out effectively, based on indicators for outputs, outcomes and impact in the logframe? 	
		<ul style="list-style-type: none"> ◦ Do performance monitoring and reviews take place regularly? 	
		<ul style="list-style-type: none"> ◦ Were resources for M&E sufficient? 	
		<ul style="list-style-type: none"> ◦ How has the logframe been used for M&E purposes (developing M&E plan, setting M&E system, determining baseline and targets, annual implementation review by the Project Steering Committee...) to monitor progress towards expected outputs and outcomes? 	
		<ul style="list-style-type: none"> ◦ How well have risks outlined the project document and in the logframe been monitored and managed? How often have risks been reviewed and updated? Has a risk management mechanism been put in place? 	
3	Results-based management (RBM)	Assessment of issues related to results-based work planning, results-based M&E and reporting based on results.	
	<i>Results-Based work planning</i>	<ul style="list-style-type: none"> ◦ Review any delays in project start-up and implementation, identify the causes and examine if they have been resolved. 	
		<ul style="list-style-type: none"> ◦ Are there any annual work plans? Are work-planning processes results-based? Has the logframe been used to determine the annual work plan (including key activities and milestone)? 	
		<ul style="list-style-type: none"> ◦ Examine the use of the project's results framework/ logframe as a management tool and review any changes made to it since project start. 	
		<ul style="list-style-type: none"> ◦ Review the monitoring tools currently being used: Do they provide the necessary information? Do they involve key partners? Are they aligned or mainstreamed with national systems? Do they use existing information? Are they efficient? Are they cost-effective? Are additional tools required? How could they be made more participatory and inclusive? 	

		<ul style="list-style-type: none"> ◦ Do project team and manager make decisions and corrective actions based on analysis from M&E system and based on results achieved? Is information on project performance and results achievement being presented to the PSC to make decisions and corrective actions? Do the project team and managers and PSC regularly ask for performance and results information? 	
	<i>Results-based reporting</i>	<ul style="list-style-type: none"> ◦ Assess how adaptive management changes have been reported by the project management and shared with the PSC. 	
		<ul style="list-style-type: none"> ◦ Assess how well the project team and partners undertake and fulfil donor and UNIDO reporting requirements (i.e. how have they addressed delays or poor performance, if applicable?) 	
		<ul style="list-style-type: none"> ◦ Assess how results and lessons derived from the adaptive management process have been documented, shared with key partners and internalized by partners. 	
4	Environmental impacts	<ul style="list-style-type: none"> • Did the project use an environmental screening and assessment procedure? 	Yes
		<ul style="list-style-type: none"> • To what extent did the project identify and realize opportunities to strengthen the environmental sustainability? 	
		<ul style="list-style-type: none"> • To what extent did the project assess those adverse environmental impacts and risks? 	
		How did the project mitigate adverse environmental impacts and risks?	
5	Social impacts: human rights, including disability	<ul style="list-style-type: none"> • Did the project use a social screening and assessment procedure? 	Yes
		<ul style="list-style-type: none"> • To what extent have poor, indigenous and physically challenged, women, men and other disadvantaged and marginalized groups benefited from the work of UNIDO in the country? 	
		<ul style="list-style-type: none"> • To what extent did the project identify and realize opportunities to strengthen the social sustainability? 	
		<ul style="list-style-type: none"> • To what extent did the project assess those adverse social impacts and risks? 	
		<ul style="list-style-type: none"> • How did the project mitigate adverse social impacts and risks, based on the social safeguards specified in the UNIDO environmental and social safeguards policies and procedures (ESSPP) (which include human rights)? 	
		<ul style="list-style-type: none"> • How did the project address disability inclusion? 	
E	Performance of partners	Assessment of partners' roles and responsibilities engaged in the intervention.	Yes
1	UNIDO	<ul style="list-style-type: none"> ◦ Mobilization of adequate technical expertise for project design 	
		<ul style="list-style-type: none"> ◦ Inclusiveness of project design (with national counterparts) 	
		<ul style="list-style-type: none"> ◦ Previous evaluative evidence shaping project design 	
		<ul style="list-style-type: none"> ◦ Planning for M&E and ensuring sufficient M&E budget 	
		<ul style="list-style-type: none"> ◦ Timely recruitment of project staff 	
		<ul style="list-style-type: none"> ◦ Project modifications following changes in context or after the Mid-Term Review 	
		<ul style="list-style-type: none"> ◦ Follow-up to address implementation bottlenecks 	
		<ul style="list-style-type: none"> ◦ Role of UNIDO country presence (if applicable) supporting the project 	
		<ul style="list-style-type: none"> ◦ Engagement in policy dialogue to ensure up-scaling of innovations 	
		<ul style="list-style-type: none"> ◦ Coordination function 	
		<ul style="list-style-type: none"> ◦ Exit strategy, planned together with the government 	

		<ul style="list-style-type: none"> ◦ Review overall effectiveness of project management as outlined in the project document. Have changes been made and are they effective? Are responsibilities and reporting lines clear? Is decision-making transparent and undertaken in a timely manner? Recommend areas for improvement. 	
		<ul style="list-style-type: none"> ◦ To what extent the project has a proper and operational governance system (e.g. PSC with clear roles and responsibilities)? 	
		<ul style="list-style-type: none"> ◦ Review whether the national management and overall coordination mechanisms have been efficient and effective? Did each partner have assigned roles and responsibilities from the beginning? Did each partner fulfil its role and responsibilities (e.g. providing strategic support, monitoring and reviewing performance, allocating funds, providing technical support, following up agreed/corrective actions)? 	
		<ul style="list-style-type: none"> ◦ The UNIDO HQ-based management, coordination, monitoring, quality control and technical inputs have been efficient, timely and effective (e.g. problems identified timely and accurately; quality support provided timely and effectively; right staffing levels, continuity, skill mix and frequency of field visits)? 	
2	National counterparts	Assessment of roles and responsibilities of national counterparts such as government ministries, NGOs, civil society and the private sector where appropriate.	Yes
	<i>Design</i>	<ul style="list-style-type: none"> ◦ Responsiveness to UNIDO's invitation for engagement in designing the project 	
	<i>Implementation</i>	<ul style="list-style-type: none"> ◦ Ownership of the project 	
		<ul style="list-style-type: none"> ◦ Provide financial contribution as planned (cash or in-kind) 	
		<ul style="list-style-type: none"> ◦ Support to the project, based on actions and policies 	
		<ul style="list-style-type: none"> ◦ Counterpart funding 	
		<ul style="list-style-type: none"> ◦ Internal government coordination 	
		<ul style="list-style-type: none"> ◦ Exit strategy, planned together with UNIDO, or arrangements for continued funding of certain activities 	
		<ul style="list-style-type: none"> ◦ Facilitation of the participation of NGOs, civil society and the private sector where appropriate 	
		<ul style="list-style-type: none"> ◦ Suitable procurement procedures for timely project implementation 	
		<ul style="list-style-type: none"> ◦ Engagement with UNIDO in policy dialogue to promote the up-scaling or replication of innovations 	
3	Donor	<ul style="list-style-type: none"> • Timely disbursement of project funds 	Yes
		<ul style="list-style-type: none"> • Feedback to progress reports, including Mid-Term Evaluation, if applicable 	
		<ul style="list-style-type: none"> • Support by the donor's country presence (if applicable) supporting the project for example through engagement in policy dialogue 	
4	Implementation partner	<ul style="list-style-type: none"> • Timely recruitment of project staff 	Yes
		<ul style="list-style-type: none"> • Project modifications following changes in context or after the mid-term review 	
		<ul style="list-style-type: none"> • Follow-up to address implementation bottlenecks 	
		<ul style="list-style-type: none"> • Review overall effectiveness of project management as outlined in the project document. Have changes been made and are they effective? Are responsibilities and reporting lines clear? Is decision-making transparent and undertaken in a timely manner? Recommend areas for improvement. 	

		<ul style="list-style-type: none"> • To what extent the project has a proper and operational governance system (e.g. PSC with clear roles and responsibilities)? 	
		<ul style="list-style-type: none"> • Review whether the national management and overall coordination mechanisms have been efficient and effective? Did each partner have assigned roles and responsibilities from the beginning? Did each partner fulfil its role and responsibilities (e.g. providing strategic support, monitoring and reviewing performance, allocating funds, providing technical support, following up agreed/ corrective actions)? 	
F	Overall assessment	Overarching assessment of the project, drawing upon the analysis made under Project performance and Progress to Impact criteria above but not an average of ratings.	Yes



Vienna International Centre
Wagramerstr. 5, P.O. Box 300,
A-1400 Vienna, Austria



+43 1 26026-0



www.unido.org



evaluation@unido.org



UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION