

Office of Evaluation and Internal Oversight
Independent Evaluation Division

Evaluation Manual



UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

Office of Evaluation and Internal Oversight
Independent Evaluation Division

Evaluation Manual



UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

Vienna, 2018

FOREWORD

“Follow Your Curiosity! I have no special talent. I am only passionately curious. What piques your curiosity? I am curious as to what causes one person to succeed while another person fails; this is why I’ve spent years studying success. What are you most curious about? The pursuit of your curiosity is the secret to your success” (Albert Einstein)

It is with the above thought in mind that the Evaluation Function in UNIDO aims to contribute to increasing the efficiency, effectiveness and impact of our Organization. Its value derives from the capacity to help UNIDO stakeholders reflecting and learning from what we have done in the past, and then being able to avoid making the same mistakes over again, while not being afraid to make new mistakes.

This evaluation manual presents the core methodology and processes that the UNIDO Independent Evaluation Division applies to conduct its evaluations. It is based on international good evaluation standards and practices, and on the principles established in the UNIDO Evaluation Policy.

UNIDO Evaluation Function is in the context where our member states have put together a global roadmap, a pathway called SDGs or 2030 Agenda and a more specific goal to achieve Inclusive and Sustainable Industrial Development (ISID). UNIDO has the responsibility to be accountable for adding value under this framework, and demonstrate with facts and evidence that the organization is actually contributing to what it is supposed to contribute.

UNIDO has become the custodian of SDG-9, namely to “Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation”. This particular SDG is at the center of UNIDO’s mandate and means significant responsibility. While the SDG-9 would drive UNIDO in the next 15 years, it is also clear that it cannot be isolated from the other SDGs. For example, UNIDO mandate and programs are also addressing other SDGs, such as SDGs 1 (no poverty), 5 (gender equality), 7 (clean energy), 8 (economic growth), 11 (sustainable cities), 13 (climate action) and 17 (partnerships). Without “partnership” in a broader, impact-oriented and most inclusive sense, none of our efforts would be able to contribute effectively to the 2030 agenda.

Hence, UNIDO Independent Evaluation Division has the challenge to provide to all stakeholders with the evaluative evidence on to what extent our organization is really adding the expected value for the 2030 agenda, what works, what does not, why and what needs to be adjusted.

The credibility and effectiveness of UNIDO is therefore directly linked to the strength of its independent, credible and useful evaluation function.

We hope you find this manual useful in your own work with UNIDO,

UNIDO Independent Evaluation Division Team:

Mr. Javier Guarnizo,	Chief
Ms. Muge Dolun,	Evaluation Officer
Ms. Thuy Thu Le,	Evaluation Officer
Ms. Michaela Berndl,	Senior Evaluation Assistant
Ms. Simone La Rosa Monier,	Senior Evaluation Assistant

February 2018

ACKNOWLEDGEMENTS

This evaluation manual has been developed following a consultative process by the staff of the UNIDO Independent Evaluation Division (IED), under the leadership of the IED Chief, and coordinated by Ms. Thuy Thu Le, Evaluation Officer. Mr. Achim Engelhardt, evaluation expert, provided significant support in preparing the evaluation manual.

The evaluation manual benefited from invaluable insight, knowledge, and comments from a broad range of UNIDO Management and staff. Contribution and comments were gratefully acknowledged from independent external evaluators who have been working with IED for the past few years.

The designations employed and the presentation of the material in this document do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations Industrial Development Organization (UNIDO) concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Mention of company names and commercial products does not imply the endorsement of UNIDO.

This document has not been formally edited.

Contents

1 BACKGROUND.....	9
1.1 AUDIENCE AND PURPOSE OF EVALUATION MANUAL.....	10
1.2 WHY ARE EVALUATIONS DONE?.....	10
1.3 MINIMIZING THE MISUSE OF EVALUATION.....	11
2 WHAT IS EVALUATION?.....	12
2.1 DEFINITION.....	12
2.2 EVALUATION AND RELATED CONCEPTS: ROLE WITHIN RESULTS-BASED MANAGEMENT.....	12
2.2.1 RESULTS-BASED MANAGEMENT.....	12
2.2.2 MONITORING.....	14
2.2.3 SELF-EVALUATION.....	14
2.3 UNIDO EVALUATIONS' TYPES AND DIMENSIONS.....	14
2.3.1 INDEPENDENT EVALUATIONS.....	14
2.4 KEY SUCCESS FACTORS OF EVALUATION.....	16
2.5 QUALITY STANDARDS.....	17
2.6 EVALUATION ETHICS.....	17
3 METHODOLOGY.....	18
3.1 EVALUATION CRITERIA.....	18
3.2 FORMULATING EVALUATION QUESTIONS.....	19
3.3 RATING SYSTEMS.....	24
3.3.1 PROJECT DESIGN.....	25
3.3.2 IMPLEMENTATION PERFORMANCE.....	25
3.3.3 PARTNERS' PERFORMANCE.....	25
3.3.4 GENDER MAINSTREAMING.....	27
3.4 RECONSTRUCTING THE INTERVENTION LOGIC: LOGFRAME AND THE THEORY OF CHANGE.....	27
3.4.1 LOGFRAME.....	27
3.4.2 THEORY OF CHANGE.....	28
3.5 DATA COLLECTION TECHNIQUES AND INSTRUMENTS.....	30
3.6 DATA ANALYSIS.....	33
3.7 THE "WHY" AND "HOW" QUESTIONS: UNDERSTANDING CHANGE CAUSES AND PROCESSES.....	34
3.8 THE CHALLENGE OF EVALUATING IMPACT.....	34
3.8.1 IMPACT EVALUATION DOMAINS AND TECHNIQUES.....	35
3.9 BENCHMARKING.....	36
3.10 FORMULATING RECOMMENDATIONS.....	36
3.11 PARTICIPATION.....	37
3.12 USING THE EVALUATION.....	38
3.12.1 USERS OF EVALUATION.....	39
3.12.2 DISSEMINATION OF REPORT.....	39
3.12.3 LESSONS LEARNED AND GOOD PRACTICES.....	39
4 THE PROJECT EVALUATION PROCESS: STEP-BY-STEP.....	41
4.1 PLANNING AN EVALUATION AT THE DESIGN STAGE.....	44
4.1.1 EVALUATION PLANNING AND BUDGET.....	44
4.2 PREPARATION OF EVALUATION TASK.....	44
4.2.1 DEVELOPING TERMS OF REFERENCE.....	45
4.2.2 REVIEWING THE INTERVENTION.....	46

4.2.3	INTERVENTION RATIONALE.....	46
4.2.4	ASSESSING EVALUABILITY.....	47
4.2.5	EVALUATION TEAM COMPOSITION AND RECRUITMENT.....	47
4.3	IMPLEMENTATION AND REPORTING.....	49
4.3.1	INCEPTION PHASE.....	49
4.3.2	QUALITY ASSURANCE OF EVALUATION DELIVERING ON TOR IN INCEPTION PHASE	50
4.3.3	FIELD WORK	51
4.3.4	PRESENTATION OF PRELIMINARY FINDINGS.....	52
4.3.5	REPORTING.....	52
4.4	FOLLOW-UP, LEARNING, AND DISSEMINATION	53
4.4.1	MANAGEMENT RESPONSE TO EVALUATIONS AND INTERNAL FOLLOW-UP	53
5	THE PROCESS FOR STRATEGIC EVALUATIONS: COUNTRY, THEMATIC AND IMPACT EVALUATIONS	55
5.1	BACKGROUND.....	55
5.2	COUNTRY EVALUATIONS.....	55
5.2.1	PURPOSE.....	55
5.2.2	SCOPE.....	56
5.2.3	SELECTION AND TIMING	57
5.2.4	PROCESS AND METHODOLOGY	57
5.2.5	REPORT.....	58
5.3	THEMATIC EVALUATIONS	61
5.3.1	PURPOSE.....	61
5.3.2	SCOPE.....	61
5.3.3	SELECTION AND TIMING	61
5.3.4	PROCESS AND METHODOLOGY	61
5.3.5	REPORT.....	63
5.4	IMPACT EVALUATIONS.....	64
5.4.1	OBJECTIVES	64
5.4.2	PURPOSE.....	64
5.4.3	SCOPE.....	64
5.4.4	SELECTION AND TIMING	64
5.4.5	PROCESS AND METHODOLOGY	64
5.4.6	REPORT.....	66
	ANNEX 1: REFERENCES TO GOOD EVALUATION PRACTICES	67
	ANNEX 2: DEFINITION OF EVALUATION CRITERIA INCLUDING KEY EVALUATION QUESTIONS	68
	ANNEX 3: ENDNOTES.....	74

Figures

Figure 1: The role of UNIDO's Evaluation Manual: bridging the missing middle	9
Figure 2: The purposes of evaluation in UNIDO	10
Figure 3: Categories of misusing evaluation	11
Figure 4: Evaluation and related functions in UNIDO	13
Figure 5: The concepts of monitoring and evaluation in UNIDO	14
Figure 6: Evaluation types in UNIDO under the responsibility of the IED.....	15
Figure 7: Definition of evaluation criteria	19
Figure 8: Good practices for the formulation of evaluation questions.....	20
Figure 9: UNIDO evaluation rating scale	24
Figure 10: Formula transforming UNIDO ratings into GEF ratings	24
Figure 11: Main data collection methods used in UNIDO evaluations	31
Figure 12: Examples of data analysis in UNIDO evaluations.....	33
Figure 13: Flowchart for project terminal evaluations (non-GEF projects)	42
Figure 14: Flowchart for project terminal evaluations (GEF projects)	43
Figure 15: Common challenges in UNIDO country evaluations and possible solutions.....	57
Figure 16: Good practice in identifying and presenting key evaluation findings.....	59
Figure 17: Good practice example in proofing logic between key evaluation findings, conclusions and recommendations.....	60
Figure 18: Common challenges in UNIDO thematic evaluations and possible solutions.....	62

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
EVA	Evaluation Unit UNIDO (prior to IEV)
GEF	Global Environment Facility
IDB	Industrial Development Board
IEV	Independent Evaluation Division
IFAD	International Fund for Agriculture Development
IP	Integrated Programme
LF/Logframe	Logical Framework
MP	Montreal Protocol
MRS	Management Response Sheet
MTPF	Medium-Term Programme Framework
M&E	Monitoring and Evaluation
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
SDGs	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 first section puts UNIDO’s Evaluation Manual in the context of UNIDO’s endeavor to achieving development results and its contribution to the 2030 Agenda. It underscores the acknowledged importance of evaluation in UNIDO and how the Evaluation Manual complements UNIDO’s Evaluation Policy. The purpose of the manual for different audiences and an explanation on why evaluations are carried out within UNIDO follow. This 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 an efficient and comprehensive independent evaluation function that needs to be aligned with international policies, standards and practices for measuring the results and impact of UNIDO’s work at field level. This is in line with the 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 developmentⁱ.

In the light of the 2030 Agenda and the Sustainable Development Goals (SDG), and with the aim to “no-one left behind”, UNIDO’s Programme and Budget Committee, in its forty-fourth session, recommended to the Industrial Development Board the adoption of a draft decision including: “inviting the Director General to (...) continue reporting on the implementation of the UNIDO’s contribution to the realization of the 2030 Agenda and its Sustainable Development Goals in the Annual Report”ⁱⁱ. The role of evaluation in independently assessing progress of UNIDO’s contribution to the SDG’s is an on-going debate. At the same time the Industrial Development Board “encouraged evaluations on results at outcome and impact levels, and the incorporation of information on lessons and performance into management and strategic planning processes.”ⁱⁱⁱ

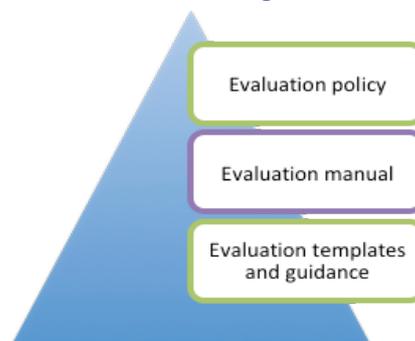
The 2010 Peer review of the evaluation function of UNIDO found that:

“Evaluation is a key component in the organization’s ability to articulate and measure results, demonstrate relevance, and share knowledge about how results are achieved”.

Source: Ministry of Foreign Affairs of Denmark, 2010: Peer review of the evaluation function of UNIDO, page 20

Following the approval and endorsement of UNIDO updated Evaluation Policy in 2015 (following the original version in 2006) and, as part of its continuous efforts to strengthen the evaluation capacity and methodology within UNIDO and among its key stakeholders, the need to develop the UNIDO Evaluation Manual has been acknowledged. The manual aims to be an instrument that would help, on the one hand, to operationalize the Evaluation Policy, and, on the other hand, to link existing evaluation guidance and templates to the Evaluation Policy. Hence, the Evaluation Manual bridges the existing gap between the Evaluation Policy and the Evaluation templates and guidance as shown in **Figure 1**.

Figure 1: The role of UNIDO’s Evaluation Manual: bridging the missing middle



1.1 Audience and purpose of Evaluation Manual

The primary audience of the Evaluation Manual is the staff of the Independent Evaluation Division (IEV), evaluation consultants and UNIDO's Project Managers and their teams. UNIDO Project Managers and their teams benefit from better understanding and guidance in day-to-day operations related to their roles and responsibilities in an evaluation. Evaluation consultants working for UNIDO gain from getting a consistent view of evaluation and how evaluation is practiced in UNIDO.

The secondary audience of the Evaluation Manual comprises UNIDO Senior Management and Member States stakeholders. Both groups get insights on how the Evaluation Policy is operationalized, how evaluations are conducted, and to underscore the robustness of the process, the consistently high quality of evaluation products and its due independence.

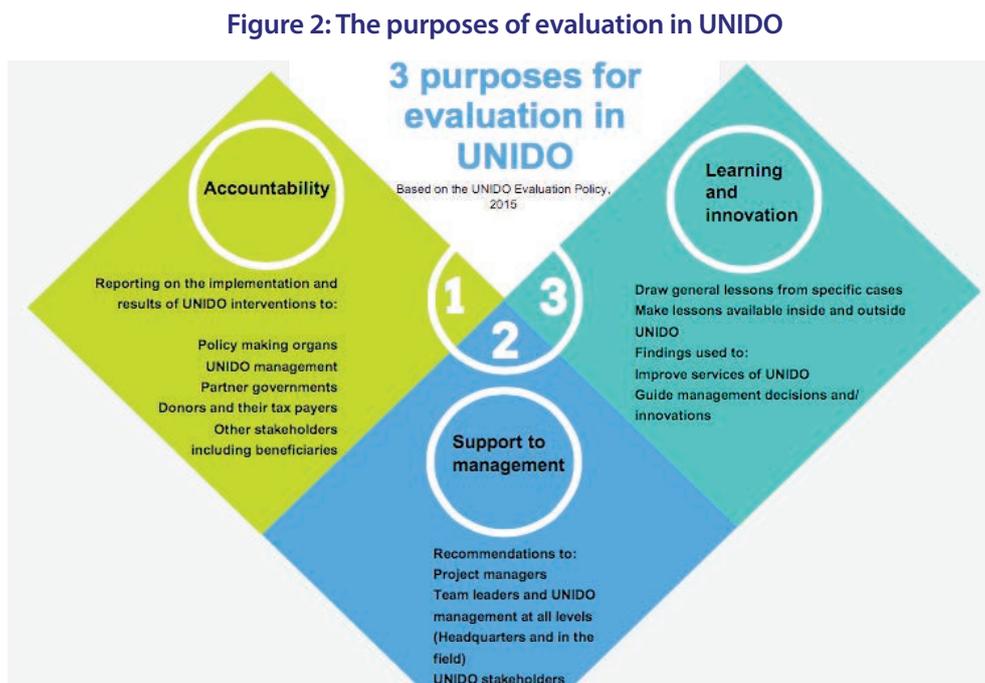
The manual's main purpose is to ensure consistency, rigor, and transparency across independent evaluations and ultimately enhance the effectiveness of the independent evaluation function at UNIDO, and the Independent Evaluation Division's quality of work.

1.2 Why are evaluations done?

Per UNIDO Evaluation Policy, evaluation serves three purposes in UNIDO. It i) assures accountability, ii) supports management and iii) drives learning and innovation.

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

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



1.3 Minimizing the misuse of evaluation

Evaluation is aimed to be used for accountability, learning, and innovation and to support management in UNIDO. However, the occasional intentional or unintentional misuse of evaluation is a reality in evaluating international development in the UN System and beyond.

For UNIDO staff, consultants and stakeholders to identify cases of misuse, the following broad categories¹ are shown in **Figure 3**:

Figure 3: Categories of misusing evaluation

Legitimization

- a) Evaluation is used to mobilise authoritative support for views that are held regardless of the evaluation.
- b) Intended to justify a particular interest, policy, or point of view, rather than to find answers to unresolved questions or provide solutions to outstanding problems.

Tactical use

- a) Evaluation is used to gain time, evade responsibility, or, perhaps, create an opportunity for renewed dialogue among stakeholders.
- b) Intended to convince users that matters are under control.

Ritual use

- a) Evaluations are carried out only as a formality, a box to tick.
- b) Participants remind themselves of the larger meaning of the evaluated activities. The evaluation gives them an opportunity to focus on the 'big picture'.

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 inadequate. Instead, joint workshops with UNIDO and its partners or other engagement processes are recommended for those purposes as part of the intervention roll out.

¹ by Sida, 2007: Looking back, moving forward. Sida Evaluation Manual. 2nd revised version, adapted from Vedung, E., 1997: Public Policy and Program Evaluation. New Brunswick: Transaction Publishers.

2. What is evaluation?

Section 2 explores the definition of evaluation. The role of evaluation in UNIDO's results-based management is clarified, also highlighting the differences between monitoring and evaluation. The section explains the types and dimensions of evaluations in UNIDO. Key success factors for evaluations and quality standards close section 2.

2.1 Definition

UNIDO Evaluation Policy, adopted in 2006 and updated in early 2015, provides a clear definition of evaluation, based on the definition of the Development Assistance Committee of the Organization for Economic Development and Cooperation (OECD/DAC, 1991)^{iv}.

UNIDO builds on the United Nations Evaluation Group (UNEG) Norms and Standards (2016) to further elaborate this definition and includes additional evaluation objects^v: “strategy, policy, topic, theme, sector, operational area or institutional performance”. The 2016 UNEG definition also specifies that evaluation “analyses the level of achievement of expected as well as unexpected results or consequences by examining the results chain, processes, contextual factors and causality”. This complements the definition provided in the UNIDO Evaluation Policy.

“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”.

A glossary of the terminology used in the UNIDO Evaluation Manual can be found in the OECD/DAC, “[Glossary of key terms in evaluation and results-based management](#)” (2002)

2.2 Evaluation and related concepts: Role within Results-Based Management

2.2.1 Results-Based Management

UNIDO defines Results-Based Management (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.”^{vi}

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^{vii}.

This is supported through demonstrating the development impact across all its Technical Cooperation (TC) programmes and projects as well as lessons learned from evaluation

recommendations. 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 through its 2010-2013 Medium-Term Programme Framework (MTPF) (extended until 2015), 2016-2019 MTPF and its recently updated 2018-2021 MTPF, 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)^{viii}.

UNIDO Evaluation Policy clearly distinguishes evaluation from other related functions within the organization, as shown in **Figure 4**.

Figure 4: Evaluation and related functions in UNIDO



2.2.2 Monitoring

Monitoring and evaluation are interconnected but are different functions. Good evaluation depends on good monitoring. **Figure 5** highlights the main differences between these two functions.

Figure 5: The concepts of monitoring and evaluation in UNIDO

Monitoring	Evaluation
Continuous	Periodic
Management function	Independent from management
During implementation; it includes self-evaluations	Possible before (ex-ante), during (e.g. mid-term) and after implementation (final, ex-post)
Intervention objectives are taken as given and used to "keep the ship on course," corrective action to improve performance	Objectives assessed about higher-level goals or to the development problem to be 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

Source: UNEG, 2016, Sida, 2007^{ix}, adapted

2.2.3 Self-evaluation

Self-evaluations in UNIDO are progress reviews at a given moment in time of projects or programmes carried out by those responsible for their implementation (as part of the monitoring tools, e.g. mid-term reviews, final review, final self-evaluation). Self-evaluations and reviews build upon monitoring and reporting and take place according to the rules established in Project Management Guidelines. They are the vehicle for steering corrective action by line management. They are also sources of information for independent evaluations. The main difference between self-evaluations and independent evaluations are the **criteria of independence**, which is ensured by the fact that the Independent Evaluation Division manages independent evaluations. Self-evaluations are conducted by those responsible for their design and implementation, at times supported by external experts. The latter constitutes a good practice of self-evaluation to enhance its objectivity and usefulness.

Independent evaluations are always conducted by external evaluators who must not have been directly involved in the design and implementation of the programmes and projects under evaluation. Independent evaluations are managed and quality assured by the IED.

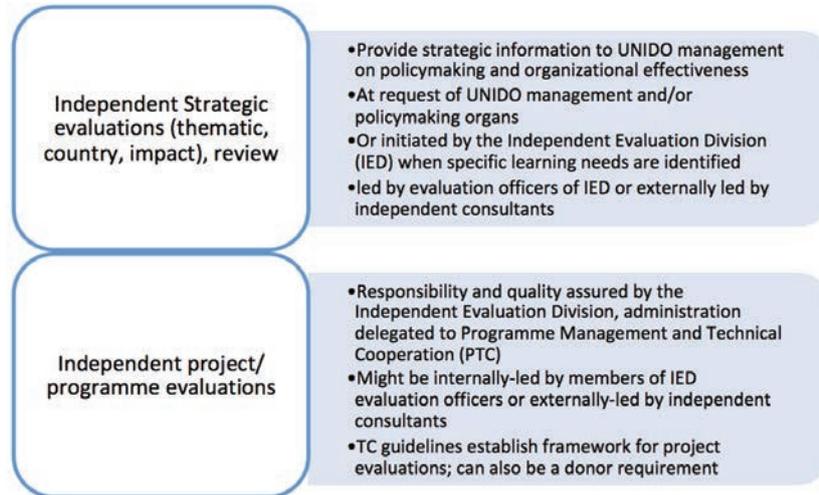
2.3 UNIDO evaluations' types and dimensions

The UNIDO Evaluation Policy distinguishes different evaluation types in the organization. This is further specified by the Work Programme of the Independent Evaluation Division, as shown in the following section.

2.3.1 Independent evaluations

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

Figure 6: Evaluation types in UNIDO under the responsibility of the Independent Evaluation Division



Source: UNIDO, 2015: Director General’s Bulletin: Evaluation Policy, updated

The Independent Evaluation Division is responsible for independent evaluations of all types and at all levels, including “light” evaluations using desk reviews and other data collection methods, while self-evaluations are not under the responsibility of the Independent Evaluation Division but of programme/project management.

Examples of evaluations for the 2016/17 evaluation work programme:

- “Thematic evaluation of UNIDO’s partnerships with donors”
- “Impact evaluation of UNIDO activities in the area of energy efficiency”

Independent Strategic evaluations/reviews (thematic, country, impact)

- ▶ Independent strategic evaluations or reviews such as thematic and impact evaluations are selected following in-house consultations, including requests expressed by the UNIDO Executive Board and Member States. The evaluation focus is on utility and learning from evaluation. Evaluation findings and recommendations aim to feed into results-oriented reporting, strategic planning frameworks and organizational learning, as well as in-house research activities. (A more detailed description of the process for Independent Strategic evaluations/reviews is presented in Section 5.)
- ▶ The Independent Evaluation Division selects country evaluations mainly based on: the size of the TC portfolio, including Integrated Programme (IP), Cleaner Production (CP), Global Environment Facility (GEF) and Montreal Protocol (MP) projects; the presence of a UNIDO field office, regional relevance and balance among regions; or specific requests by Member States.

Independent project/programme evaluations

- ▶ For independent project or programme evaluations, the Independent Evaluation

Division plays an advisory, clearance and Quality Assurance (QA) function. Currently, project evaluation administration is delegated to Project Managers. The Independent Evaluation Division is responsible for identifying and selecting evaluation consultants and clearing evaluation Terms of Reference (ToR), inception report, draft, and final report. Specific donor's (for example GEF) evaluation requirements are considered as needed. The UNIDO Executive Board decided in March 2016 to increase the threshold budget for projects to be subject to independent terminal evaluation, from Euro 1 million to USD 2 million. (A more detailed description of the process for Independent project/programme evaluations is presented in Section 4.)

When required, the Independent Evaluation Division is also involved in joint evaluations for example with sister UN agencies or donor evaluations where the IED provides support.

2.4 Key success factors of evaluation

Experience suggests that a number of factors determine whether evaluations are influential in UNIDO², as shown in the box below.

- (a) The extent of *consultation and inclusiveness* of the evaluation – how many stakeholders were engaged in the evaluation design, implementation and formulation of recommendations, with the aim to ultimately increase the ownership of the evaluation process and evaluation results;
- (b) The extent to which the potential *users and uses* of the report are considered and can participate throughout the evaluation process;
- (c) The extent to which the work being assessed can be *linked to a performance framework, logical framework or theory of change* – can the work be assessed against clear objectives, baselines or benchmarks;
- (d) The existence of baseline information, counterfactuals, targets, etc. to further strengthen the evidence base;
- (e) Length of the report not exceeding 30 pages, excluding annexes;
- (f) Recommendations are SMART (Specific, Measurable, Achievable, Realistic and Time-bound) as well as *Actionable*– do the recommendations make sense in the context of the program/project;
- (g) Management is engaged in the *response* process and commits to taking action;
- (h) The existence of an implementation tracking and reporting process;
- (i) Whether a clear *link* can be drawn between evidence-based findings, conclusions, and recommendations.

2 Based on: World Intellectual Property Organisation, Internal Oversight Division, 2016: Evaluation Manual, amended.

2.5 Quality standards

The UNIDO Evaluation Quality Checklist (QC) aims to ensure the quality of UNIDO evaluations. Evaluation quality standards in UNIDO are guided by the UNEG Norms and Standards for Evaluation (2016) as the main reference point for QA. The UNEG QC for evaluation ToRs and inception reports^x (2010) and the UNEG checklist for evaluation reports^{xi} (2010) serve as concrete tools to transparently and consistently assure the high quality of UNIDO evaluations at different stages in the evaluation process. Emphasis on QA of UNIDO evaluation reports addresses the uneven quality of UNIDO evaluation reports detected by the Peer review of UNIDO's evaluation function (2010).

For evaluation consultants, it is good practice to use the checklists for inception reports and evaluation reports during the respective reporting periods to deliver a product of high quality to the IEV. This should happen as early as at the first draft stages of those reports. The IEV actively encourages the use of those guidelines by systematically annexing the checklists to the evaluation Terms of Reference.

If inception reports deviate from the criteria or questions in the ToR, the evaluation reports should clearly explain such deviations.

Section Info box: Quality standards

[UNIDO evaluation policy and resources](#)



2.6 Evaluation ethics

UNIDO Evaluation function endorses the UNEG Norm on “Ethics”:

Evaluation must be conducted with the highest standards of integrity and respect for the beliefs, manners and customs of the social and cultural environment; for human rights and gender equality; and for the ‘do no harm’ principle for humanitarian assistance. Evaluators must respect the rights of institutions and individuals to provide information in confidence, must ensure that sensitive data is protected and that it cannot be traced to its source and must validate statements made in the report with those who provided the relevant information. Evaluators should obtain informed consent for the use of private information from those who provide it. When evidence of wrongdoing is uncovered, it must be reported discreetly to a competent body (such as the relevant office of audit or investigation).

Source: United Nations Evaluation Group, 2016: Norms and Standards for Evaluation, Norm 6.

3. Methodology

Section 3 addresses methodological issues for planning and undertaking an evaluation in UNIDO. This ranges from evaluation criteria and questions to data collection and analysis as well as rating systems used to assess different types of performance, including cross-cutting issues.

How to address the learning purpose of evaluations, benchmarking, issues of participation and evaluation management in UNIDO follow. The section closes addressing how evaluations are used.

UNIDO aligns with the UNEG Norms and Standards for Evaluation outlining in Standards 4.2 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”³.

3.1 Evaluation criteria

As most UN specialized agencies, UNIDO uses the internationally agreed evaluation criteria based on the OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management^{xii} subsequently aligned to UNIDO’s needs, as shown in the menu of options in **Figure 7**^{xiii}. This is anchored in UNIDO Evaluation Policy⁴. Annex 2 contains a full list of evaluation criteria with related evaluation questions.

It is acknowledged that some issues covered by one criterion might overlap with others. Yet 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 Evaluation and Results-Based Management. The consistent use of the criteria pertinent to the evaluation object allow for comparability of UNIDO’s performance over time^{xiv}. Evaluation questions are formulated around those evaluation criteria in UNIDO, as specified in the following section.

3 UNEG 2016: UNEG Norms and Standards for evaluation, page 23.

4 UNIDO, 2015: Director General’s Bulletin: Evaluation Policy. UNIDO/DGB/(M).98/Rev.1, page 4.

Figure 7: Definition of evaluation criteria

#	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
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
2	Logframe	Assessment of the logical framework aimed at planning 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	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
3	Efficiency	A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results.	Yes
4	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
D	Cross-cutting performance criteria	Other important criteria that cut across the UNIDO intervention.	
1	Gender mainstreaming	The extent to which UNIDO interventions have contributed to better gender equality and gender related dimensions were considered in an intervention.	Yes
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
3	Results-based management (RBM)	Assessment of issues related to results-based work planning, results based M&E and reporting based on results.	Yes
E	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 assessed individually, based on its expected role and responsibilities in the project life cycle.	Yes
2	National counterparts		Yes
3	Donor		Yes
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

3.2 Formulating evaluation questions

The formulation of evaluation questions determines the breadth and depth of an evaluation. An evaluation should seek to answer only the number of questions required to satisfy its practical purpose. Limiting the focus and scope of an evaluation ensures an efficient use of often scarce evaluation resources. The evaluation questions are meant to focus the evaluation work on a limited number of key points, for more targeted data collection, more in-depth analysis and a more useful report.

It is important to keep the evaluation focused and cost-effective and avoid a broad but superficial

evaluation approach. To ensure the usefulness of evaluation questions, it is advisable to double check whether the answer to specific evaluation questions is not already known or whether another on-going study (audit, review, appraisal or evaluation) is to provide the answer. Experience shows that the larger the number of evaluation questions, the less depth can be achieved in an evaluation. Hence a balance needs to be struck and no more than, for example, 15 evaluation questions aimed for, based on a thoughtful selection of the most relevant ones. Evaluators will identify key issues and questions for the evaluation team to focus on in the inception report.

It is good practice to follow a two-step process for the development of evaluation questions. A first set of evaluation questions based on^{xv}:

- a) Analysis of the intervention logic and rationale
- b) Issues justifying the launch of the evaluation
- c) Issues to be studied, as stated in the Terms of Reference
- 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 summarizes good practices for the formulation of evaluation questions⁵.

Figure 8: Good practices for the formulation of evaluation questions

Usefulness	Focus the evaluation work on a limited number of key points, with 15 suggested as a maximum. This allows more targeted data collection, more in-depth analysis and a more useful report.
Origin	Some evaluation questions are inferred directly or indirectly from the intervention logic. Other questions can also concern cross-cutting issues such as gender or the environment.
Selection	About 15 questions are selected (from the lists below) to ensure the potential usefulness of answers and their feasibility. The set of questions has to be discussed with the evaluation team at the inception meeting.
Linkages to evaluation criteria	Evaluation questions are classified by the evaluation criteria used in UNIDO: relevance, efficiency, effectiveness, impact, sustainability, project design, gender mainstreaming, M&E, Results-Based Management and performance of partners. Relating one evaluation question to one specific evaluation criteria is good practice, allowing for aggregation and comparability of results.

Source, EU, 2006, amended

⁵ European Union, Directorate General External Relations, Directorate General Development, EuropeAid Co-operation Office, Joint Evaluation Unit, 2006: Methodological bases for evaluation, page 43.

A selection of sample questions for some of the criteria is presented below⁶. It is good practice to use the most pertinent evaluation questions rather than the entire sample of questions listed below. The quality of the evaluation can suffer if too many evaluation questions in the ToR do not allow for the evaluation to reach sufficient depth. 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 urgent target group needs?
- ▶ To what extent is the project aligned with the development priorities of the country (national poverty reduction strategy, sector development strategy)?
- ▶ How does the project reflect donor policies and priorities?
- ▶ Is the project a technically adequate solution to the development problem? Does it eliminate the cause of the problem?
- ▶ To what extent does the project correspond to UNIDO's comparative advantages?
- ▶ 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?

Efficiency

- ▶ How economically are the project resources/inputs (concerning funding, expertise, time...) being used to produce results?
- ▶ To what extent were expected results achieved within the original budget and timeframe? If no, please explain why.
- ▶ Are the results being achieved at an acceptable cost? Would alternative approaches accomplish the same results at less cost?
- ▶ What measures have been taken during planning and implementation to ensure that resources are efficiently used? Were the project expenditures in line with budgets?
- ▶ Could more have been achieved with the same input?
- ▶ Could the same have been achieved with less input?
- ▶ 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?
- ▶ 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?
- ▶ What is the change attributable to the project? To what extent?
- ▶ What are the social, economic, environmental and other effects, either short-, medium- or long-term, on a micro- or macro-level?
- ▶ What effects are intended or unintended, positive or negative?
- ▶ To what extent is the project contributing to the conditions leading to the long-term transformation (objective of the project/program)? This can be measured by assessing the extent to which contributions have been mainstreamed, replicated or scaled-up.
 - 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 projects?
 - Replication: To what extent are the project's specific results (e.g. methodology, technology or lessons learned) reproduced or adopted?
 - Scaling-up: To what extent are the project's initiatives and results implemented at larger geographical scale?
- ▶ What benefits has the project help bring about (social, economic environmental)?
- ▶ The three UNIDO impact dimensions are:
 - Safeguarding environment: To what extent does the project contribute to changes in the status of environment?
 - Economic performance: To what extent does the project contribute to changes in the economic performance (for example finances, income, costs saving or expenditure) of individuals, groups and entities?
 - Social inclusiveness: To what extent does the project contribute to changes in capacity and capability of individuals, groups and entities in society, including vulnerable groups, and hence generating employment and access to education and training

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?

Section info box



[UNIDO evaluation tools. Guidelines for Terms of Reference for Evaluations.](#)
[Annex 3. Theme Specific evaluation questions.](#)

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, including the GEF. 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 performance and results. The same rating scale is used for all rating areas as shown in **Figure 9**.

Figure 9: UNIDO evaluation rating scale

Score	Definition*	Category
6	Highly satisfactory Level of achievement presents no shortcomings (90% - 100% achievement rate of planned expectations and targets).	SATISFACTORY
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).	UNSATISFACTORY
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).	

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 contains the formula applied to transform the results of UNIDO's six-point rating scale to the GEF's four-point scale for sustainability⁷.

Figure 10: Formula transforming UNIDO ratings into GEF ratings

UNIDO rating	UNIDO rating: sustainability	GEF rating: sustainability
6	Highly likely (HL)	Likely (L)
5	Likely (L)	Moderately Likely (ML)
4	Moderately likely (ML)	Moderately Likely (ML)
3	Moderately Unlikely (MU)	Moderately Unlikely (MU)
2	Unlikely (U)	Moderately Unlikely (MU)
1	Highly unlikely (HU)	Unlikely (U)

⁷ GEF uses a four-point scale for the criterion of sustainability.

This formula underscores the distinction of ratings into “satisfactory” and “unsatisfactory”, both in applying UNIDO’s six-point rating scale and the transformation into the GEF four-point rating scale for sustainability. To ensure coherence in ratings, the rating is defined above^{xvi}. The use of benchmarks like the performance of peers for the same criteria helps to facilitate the interpretation of ratings.

3.3.1 Project design

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

Overall design quality

- Pertinence to country priorities, needs of target groups and UNIDO strategies
- Consideration and use of lessons and evaluative evidence from other projects
- Technical feasibility and validity of project design
- Budgeted M&E plan with clear timelines, roles, and responsibilities
- Adequacy of risk assessment (for example financial, sociopolitical, institutional, environmental and implementation aspects)

Logframe/logframe-like matrix based on the project’s theory of change

- Clarity and logic of results-chain, including impacts, outcomes and outputs
- SMART indicators
- Adequacy of Means of Verification and Assumptions

3.3.2 Implementation performance

Implementation performance criteria correspond broadly to DAC criteria and need to be customized according to the context of the intervention to be evaluated.

- Relevance
- Effectiveness
- Efficiency
- Progress to Impact
- Sustainability of benefits

3.3.3 Partners’ performance

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. Hence, rating the partners is a key part of UNIDO project evaluations⁸. The six-point rating scale applies⁹.

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

9 6 = Highly satisfactory; 5 = Satisfactory; 4 = Moderately satisfactory; 3 = Moderately unsatisfactory; 2 = Unsatisfactory; 1 = Highly unsatisfactory

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

Project design

- Mobilization of adequate technical expertise for project design
- Inclusiveness of project design (with national counterparts)
- Previous evaluative evidence shaping project design
- Planning for M&E and ensuring sufficient M&E budget

Implementation

- Timely recruitment of project staff
- Project modifications following changes in context or after the Mid-Term Review
- Follow-up to address implementation bottlenecks
- Role of UNIDO country presence (if applicable) supporting the project
- Engagement in policy dialogue to ensure up-scaling of innovations
- Coordination function
- Exit strategy, planned together with the government
- Overall effectiveness of project management as outlined in the Project Document
- Project's governance system
- National management and overall coordination mechanisms
- UNIDO HQ-based management, coordination, monitoring, quality control and technical input

To assess the **performance of national counterparts**, the evaluation looks into the following issues:

Project design

- Responsiveness to UNIDO's invitation for engagement in designing the project

Implementation

- Ownership of the project
- Financial contributions (cash or in-kind)
- Support to the project, 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 and the private sector where appropriate
- Suitable procurement procedures for timely project 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 require ratings:

- Timely disbursement of project funds
- Feedback to progress reports, including Mid-Term Evaluation, if applicable
- Support by the donor's country presence (if applicable) supporting the project for example through engagement in policy dialogue

3.3.4 Gender mainstreaming

The UNIDO Policy on gender equality and the empowerment of women, issued initially in April 2009, and revised in March 2015 (UNIDO/DGB/(M).110/Rev.), provides the overall guidelines for establishing a gender mainstreaming strategy and action plans to guide the process of addressing gender issues in the Organization's industrial development interventions. It commits the organization that evaluations will demonstrate effective use of the UNEG guidance on evaluating from a human rights and gender equality perspective, as indicated by the Organization's meta-evaluation scores according to the UNEG Evaluation Scorecard.

In line with the UNIDO Gender Equality and Empowerment of Women Strategy, 2016-2019, all UNIDO technical assistance projects post-2015 are to be assigned a gender marker and should go through a gender mainstreaming check-list before approval. UNIDO's gender marker is in line with UN System-wide action plan (SWAP) requirements, with four categories: 0 — no attention to gender; 1 — some/limited attention to gender; 2a — significant attention to gender; 2b — gender is the principal objective^{xvii}.

Besides, Guides on Gender Mainstreaming for Inclusive and Sustainable Industrial Development (ISID) Projects in different areas of UNIDO's work have been developed and published during 2015^{xviii}, which have specific guidance on suitable outputs/activities/ indicators per technical area.

Considering the above, terminal evaluations of projects that have been approved before 2015 will follow the minimum requirements set out in the UNIDO 2016 TOR template and guidance (Section IV.C and annex 4). If the project design and gender analysis/existing indicators are not sufficient to allow for an accurate appraisal at the final evaluation, specific indicators could be created during the evaluation planning stage (preparing and revising the inception report) and assessed during the evaluation process. Together with the budget, the time required to adequately carry out a gender responsive evaluation will need to be taken into account. The evaluation time depends on the questions the assessment needs to answer, on how deep the analyses are requested to be, and on financial and human resources available as well as other external factors.

For terminal evaluations of projects that have been approved after 2015, evaluations should assess if the rating was correctly done at entry, if appropriate outputs/activities/indicators and monitoring were put in place during implementation and what results can be actually observed at the time of terminal evaluation (in line with UNIDO's organizational results reporting to SWAP). The Gender Mainstreaming six-point rating scale should then be used accordingly.

For projects that have 2a or 2b ratings at project design/entry at least one evaluation team member should have demonstrated/significant experience in evaluating GEEW projects. For other projects, evaluators are encouraged to further familiarize themselves with the key gender aspects and impacts of UNIDO projects, both through the foundation modules of "I know Gender" online course of UN Women and the UNIDO's Guides on Gender Mainstreaming ISID Projects.

3.4 Reconstructing the intervention logic: Logframe and the Theory of Change

3.4.1 Logframe

As in the case of the intervention rationale, the intervention logic should be part and parcel of the Project Document. This is in line with UNIDO's quality control for UNIDO Project Documents^{xix}. The Logframe (LF) used in UNIDO is one of the tools to express the intervention logic and contains indicators, baselines and targets. Again, for thematic or country evaluations, the intervention logic could be contained in the related Strategy Document.

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

3.4.2 Theory of Change

The Theory of Change for evaluations has been used increasingly by the evaluation office of many international development organizations such as the GEF and IFAD.

"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: GEF 2017: Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects.

What is a theory of change? A **theory of change** specifies the intervention logic showing how an intervention leads to results. It shows pathways to change in a narrative and graphic form. In this context pathways refer to the "sequence or hierarchy of changes and events that map out how things will change"^{xx}.

The theory of change is built on a set of assumptions and on how the intervention designers think a change will happen. This is to be tested in the evaluation process. The added value of a theory of change 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 behavior of a target group to benefit from industrial development.

"Getting the theory of change (TOC) right is one of the basic conditions for a good evaluation and therefore sufficient time, effort and resources must be allocated for this".

Source: IFAD 2015: Evaluation Guidelines. Second edition.

Linkages between outputs, outcomes, and impact are articulated in the theory of change. The theory of change identifies barriers to achieving the objectives and the ultimate impact of the intervention, as well as drivers of change as the result of a situation analysis, followed by the intervention's response. In this respect, the TOC complements the LF approach, as the context regarding barriers and drivers of change is often missing from the LF. Yet understanding the why, how, when and under what conditions specific actions lead to certain results is important for a correct interpretation of the evaluation results.

The theory of change is to be complemented by a Logframe to ensure that indicators and targets are specified and, if possible, quantified. Experience tells that "using 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 an intervention is implemented"^{xxi}.

To pose the right evaluation questions, the intervention logic should be clarified and goals and objectives (impact and outcomes), indicators and assumptions known before the evaluation starts. However, in reality, the evaluation often serves the purpose to reconstruct the intervention logic. This puts the evaluation manager in a situation where difficult choices have to be made. The evaluation manager needs to balance the time required to verify or reconstruct the intervention logic before the actual start of an evaluation, on the one hand, and the time required for the same task during an evaluation, with the risk to widen the evaluation scope through new evaluation questions, on the other hand.

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 program" (White and Phillips, 2012^{xxii}).

Evaluators can use the TOC at various levels, for example by: i) validating the intervention's assumptions and main problems; ii) assessing the causal pathway of the results chain from outputs to outcomes and impact; and iii) reviewing to what extent the intervention addressed the barriers identified.

Steps for undertaking an evaluation based on a theory of change 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 program 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./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":

1. <http://www.learningforsustainability.net/evaluation/theoryofchange.php>
2. http://devinfohive.info/impact_evaluation/ie/img/downloads/Theory_of_Change_ENG.pdf
3. http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/James_ToC.pdf
4. http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/DFID_ToC_Review_VogelV7.pdf
5. http://www.3ieimpact.org/media/filer_public/2012/05/07/Working_Paper_3.pdf



3.5 Data collection techniques and instruments

This section provides an overview of data collection techniques and instruments 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 through consulting major stakeholders, as stipulated in UNIDO's Evaluation Policy¹⁰.

The data collection tools used in UNIDO evaluations depend on variables such as the evaluation type, availability of data, local context and resources and time available. Frequent problems in data collection in UNIDO and possible solutions are identified below¹¹.

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.
- 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 using a mix of different methods, a “mixed-methods approach” appropriate for the issues that it addresses 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.

Figure 11 summarizes the main data collection methods used in UNIDO.

10 UNIDO, 2015: Director General's Bulletin: Evaluation Policy. UNIDO/DGB/(M).98/Rev.1, page 9

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

Figure 11: Main data collection methods used in UNIDO evaluations

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 predefined 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 labor-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 behaviors and body language; observer can obtain from the observation information that people would be reluctant to discuss (sensitive issues)	Can be labor-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

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

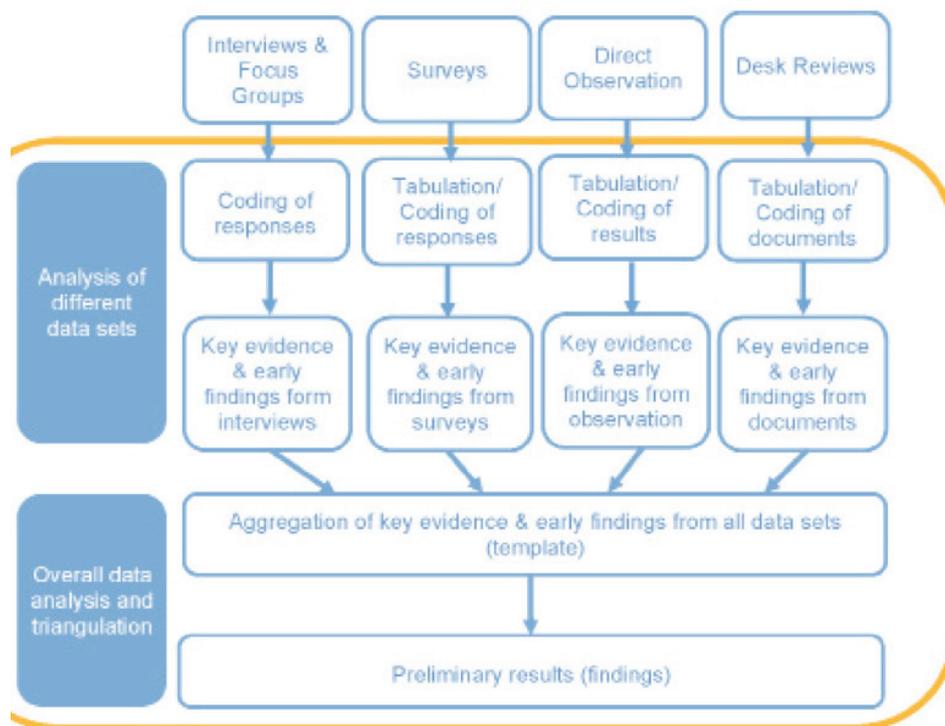
3.6 Data analysis

Data analysis follows data collection. At this step, the evaluation processes information gathered during data collection with the aim to answer 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.”¹²

A sound and methodologically rigorous data analysis is a cornerstone to fulfilling the expectations expressed for quality evaluations in UNIDO Evaluation Policy. This is the basis for soundly grounding evaluation findings, conclusions and recommendations on evidence¹³.

Once the validity, accuracy, and completeness of data has been verified, different data sets can be analyzed. In case the evaluation identifies shortcomings in the validity, accuracy, and completeness of data, this needs to be stated in the methodology section of the evaluation report under “limitations” accompanied by measures taken to mitigate those limitations. **Figure 12** summarizes generic steps in data analysis for selected data collection methods¹⁴.

Figure 12: Examples of data analysis in UNIDO evaluations



12 UNIDO, 2015: Director General’s Bulletin: Evaluation Policy. UNIDO/DGB/(M).98/Rev.1, page 4.
 13 See also: OIOS, 2014: Inspection and Evaluation Manual, page 84.
 14 OIOS, 2014: Inspection and Evaluation Manual, page 89.

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

To address the learning purpose of UNIDO evaluations, efforts are required to understand the proximate causes of good performance or limitations in achieving UNIDO development objectives. As a result, evaluations should ensure sufficient coverage of the “why” question in complementing “what” the performance was. This is required to both understand the rationale behind performance ratings and 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>)

3.8 The challenge of evaluating impact

It is often too early to assess the long-term impacts of the project at completion point. Yet it may be feasible to assess on the progress toward long-term impacts as some evidence on progress and the validity of the theory of change may be available. Where the economic, social and environmental changes are being achieved at scales beyond the immediate area of intervention, the evaluations should provide account of the process such as sustaining, mainstreaming, replication and scaling up through which these changes have taken place.

UNIDO evaluations need to assess impact as much as feasible when the timeline of the intervention allows for long-term change to have taken place. Impact measures the effects of an intervention on a target group, positive or negative, intended or unintended. Impact examines the longer-term consequences of achieving or not achieving the objectives of UNIDO interventions, and the issue of wider socioeconomic change.

While the evaluation of impact was controversially discussed in the development community for decades, no consensus has emerged for the rigorous attribution of impact of development interventions. In fact, the topic is nearly completely absent from the 2016 UNEG Norms and Standards given the lack of consensus.

UNIDO made the following methodological experiences in its first impact evaluation in 2010, to be considered for any future impact evaluation in the Organization¹⁵:

o **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;

o **Time-lag:** It has been more difficult than planned 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;

o **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:

- i) Socio-economic 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 EU due to Sri Lanka accessing to the privileged status of a GSP+ country;
- iv) Turbulences of the world textile market due to the end of the Multi-fibre Agreement affecting heavily the textile sector, Sri Lanka’s number one foreign currency earner.

o **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.

Lessons learned from UNIDO and from peers about the challenging balance between adopting a rigorous and credible approach and remaining proportionate to the level of resources available. Impact evaluations would require considerably larger budgets than other evaluation types due to the more complex methodology, as well as earlier preparation during the project lifecycle, and to ensure adequate and reliable data is collected for that purpose.

3.8.1 Impact evaluation domains and techniques

To credibly assess the impact of UNIDO’s work, planning for impact assessment **must start as early as at the project or programme development stage**. To be able to robustly answer questions about impact, it is recommended to use baselines and, if feasible, comparator groups.

The following impact domains apply to UNIDO's work:

- Institutions and policies
- Environment management
- Economic performance of enterprises and institutions
- Human and social capital and empowerment

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 the situation or scenario that would hypothetically prevail 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 learned from its peers that the identification of comparator groups at the outset of a development intervention requires upmost 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 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.

Feasibility of using control groups for impact evaluation in UNIDO:

In many developing countries and least developed countries where UNIDO works, there are often only a handful of companies of comparable capacity/size in each sector. If UNIDO already works with 80% of the main manufacturers/companies in a sector, how do we identify control groups?

To the extent possible, the Independent Evaluation Division will keep identifying and testing other options and approaches for assessing impact.

3.9 Benchmarking

Benchmarking allows comparing the performance of a specific project, sector, geographic area, or UNIDO as a whole, with other comparable units. Efforts are required to compare like with like, despite two units of assessment unlikely to be identical. 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 rather limited. Despite the value of benchmarking, comparators have learned that benchmarking results should always be interpreted carefully^{xiii}.

3.10 Formulating recommendations

UNIDO uses UNEG guidance on the formulation of evaluation recommendations. UNEG's Standard

4.10 advises that “recommendations should be firmly based on evidence and analysis, clear, results-oriented and realistic concerning implementation”¹⁶. It is of particular importance to ensure that recommendations are not opinion-based, as endorsed by UNEG and practiced by UNIDO.

Evaluators will 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&E”¹⁷.

UNIDO using UNEG standard 4.10: Recommendations

Utilization-focus:

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

Recommendations are firmly based on evidence and conclusions.

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

Recommendations clearly identify the target group for each recommendation.

Recommendations are clearly stated with priorities for action made clear.

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

The UNEG Quality Checklist for Evaluation Reports (2010)^{xxiv} outlines the following quality criteria for evaluation recommendations, going beyond the UNEG standard 4.10:

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; and iv) recommendations. Participation also increases the credibility and quality of evaluations. This is anchored in UNIDO’s Evaluation Policy¹⁸ 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)¹⁹.

16 UNEG 2016: UNEG Norms and Standards for evaluation, page 26.

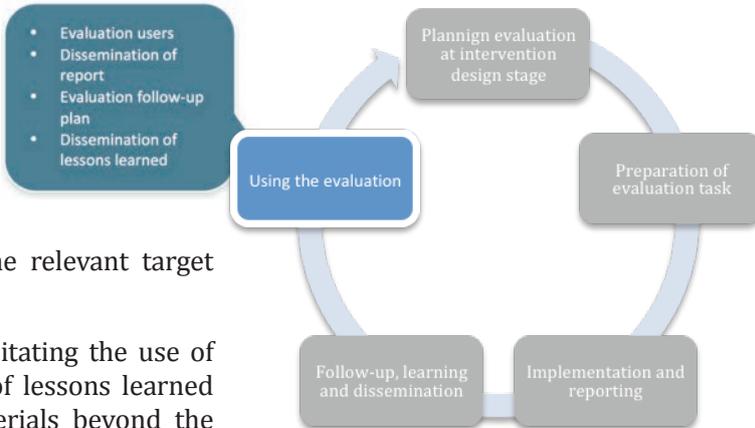
17 GEF, 2008: Guidelines for GEF Agencies in conducting terminal evaluations. Page 14.

18 UNIDO, 2015: Director General’s Bulletin: Evaluation Policy. UNIDO/DGB/(M).98/Rev.1, page 9.

19 UNEG, 2016: Norms and Standards for evaluation, page 24.

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.



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

- Policy briefs for UNIDO’s Member States or evaluation briefs for all stakeholders;
- Short videos summarizing the evaluation results and posted on the UNIDO website;
- Use of social media highlighting key evaluation results;
- Press releases; or
- 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 Evaluation Policy, page 12.

A range of evaluation products beyond the actual evaluation report can help to enhance the use of evaluations, as shown in the box below^{xxv}.

Options for evaluation products:

Evaluation insights: Insights are two-sided brochures of approximately 800 words and focus on one learning issue emerging from an evaluation.

Evaluation profiles: Profiles are also two-sided brochures (also around 800 words) and contain a user-friendly overview of the main findings and recommendations of the evaluation (evaluation types beyond project evaluations).

Overview booklets: This short format publication is intended for those who do not have time to read the whole report. It comprises an overview or summary and Management’s response, depending on the type of evaluation.

Annual report on results and impact: An annual synthesis of the organization’s results and performance based on evaluative evidence from the Organization’s evaluations.

As a complementary means, UNIDO’s Evaluation Policy foresees peer learning sessions to facilitate

the dissemination and use of the evaluations. The Independent Evaluation Division is responsible for such events.

3.12.1 Users of evaluation

The use of any evaluation should be planned for from the outset of the exercise. When considering the purpose of the evaluation, the evaluation manager should also think about who the main users of the evaluation results and lessons learned may be. The interest of those potential users can also be strengthened through involving them as part of an evaluation reference group with inputs into the evaluations' Terms of Reference and through a presentation of preliminary findings, as specified in the relevant sections above.

3.12.2 Dissemination of report

The dissemination of the evaluation and its lessons learned is the responsibility of the IEV, as outlined in UNIDO's Evaluation Policy²⁰. This underscores the independence of the Independent Evaluation Division with its right to disclosure. With this right comes the duty of proactive and effective communication, calling for communication and dissemination strategies, 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.

Evaluation Team Leaders (TL) should provide the Independent Evaluation Division with an electronic version of the report, and the Independent Evaluation Division might decide to print evaluation reports. All evaluation reports are publicly available on the UNIDO website and Intranet.

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**"²¹*

20 UNIDO, 2015: Director General's Bulletin: Evaluation Policy. UNIDO/DGB/(M).98/Rev.1, page 12.
21 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 is not mistaken for a recommendation or conclusion.

Lessons learned and **good practices** are part of a continuum with the aim of reusing lessons learned 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 determined to be specifically useful for replication or up-scaling

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

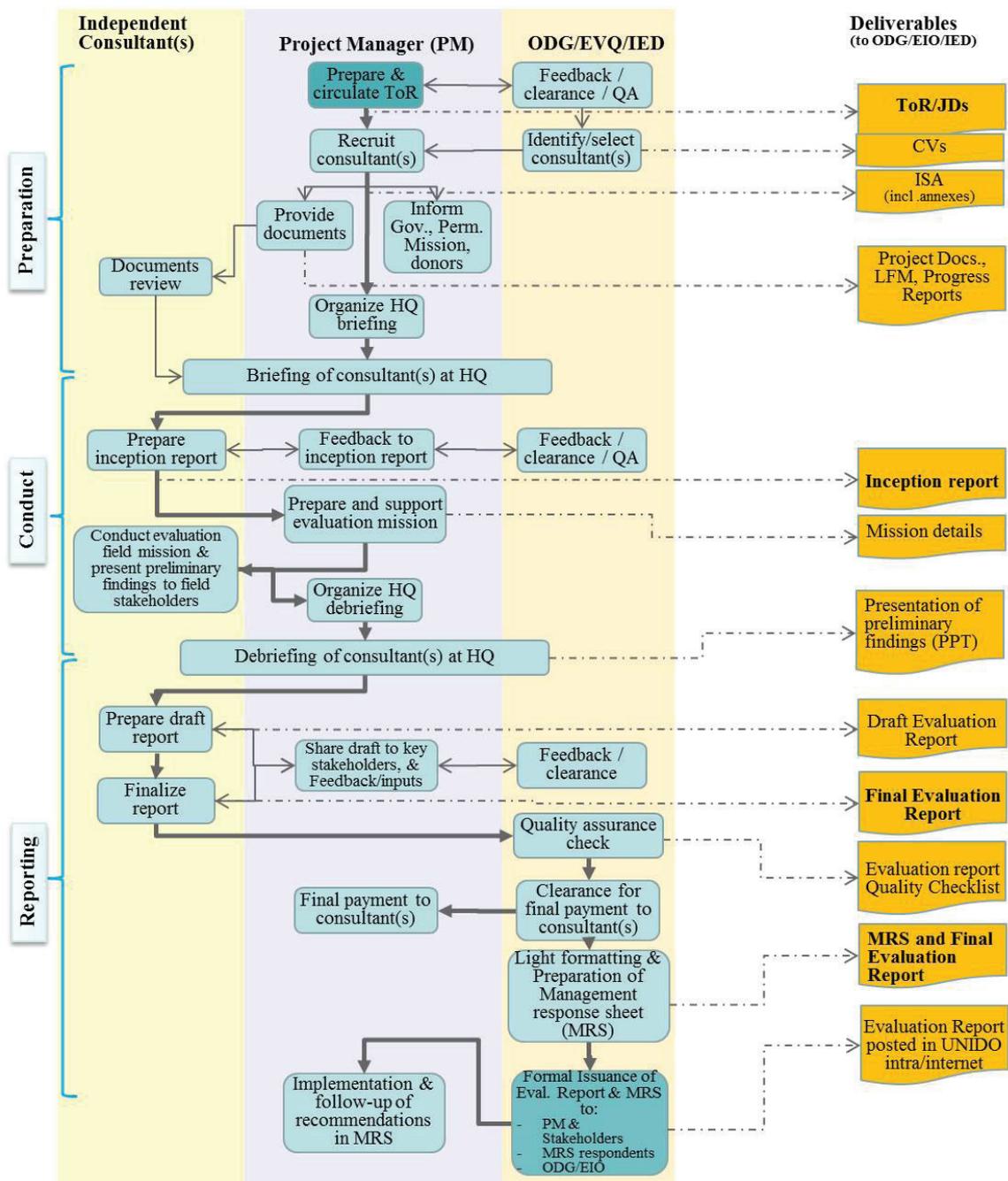
- ▶ Potential for replication
- ▶ Link to UNIDO policy goals

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

This section leads the readers through a standardized project terminal evaluation process in UNIDO: Planning and budgeting for the evaluation at intervention design, preparing the evaluation task, implementing the evaluation and the reporting phase, follow-up, learning and dissemination and using the evaluation. Strategic evaluations are addressed in Section 5.

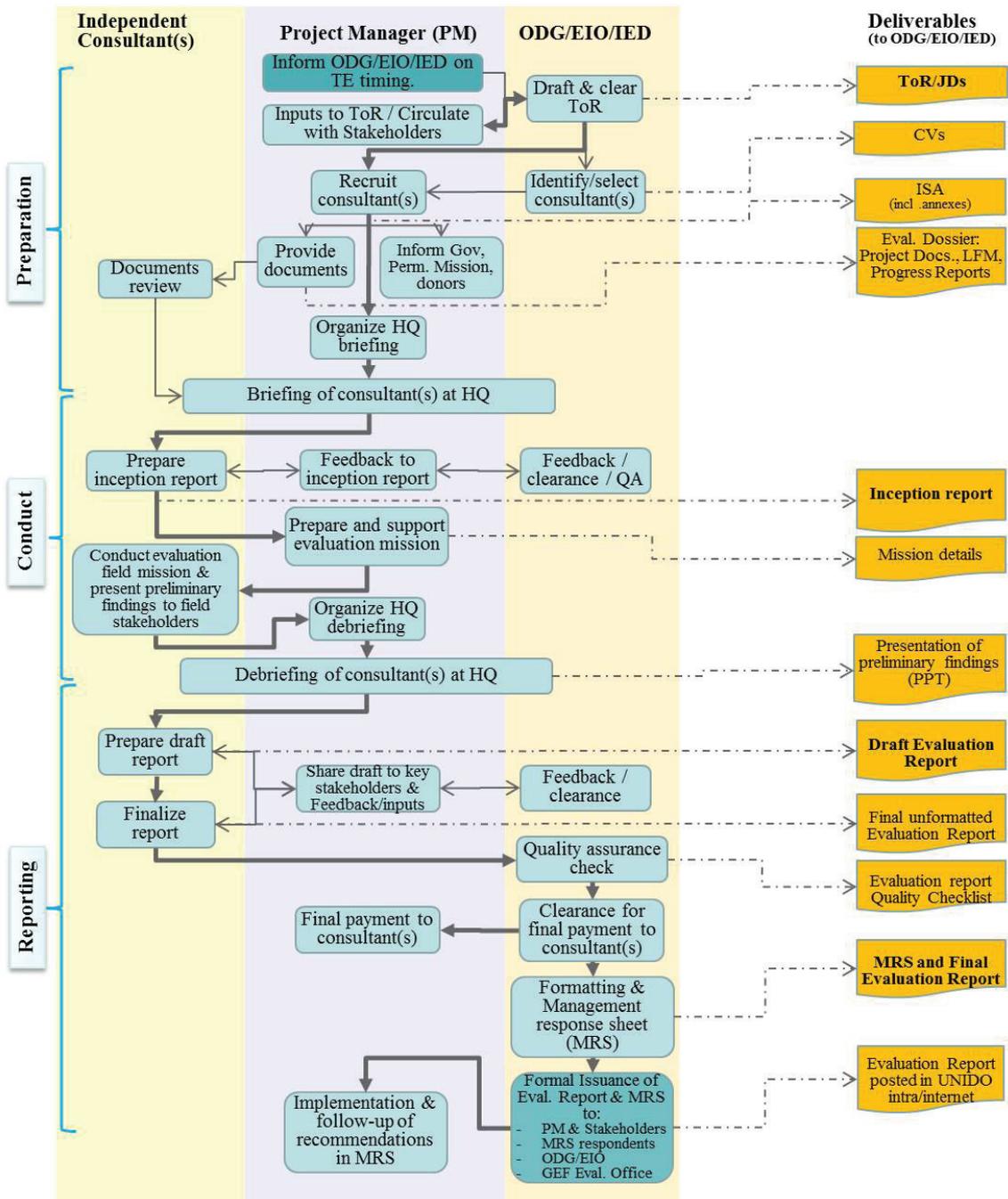
The Independent Evaluation Division distinguished between non-GEF project evaluations and GEF project evaluations. Flowcharts in **Figure 13** and **Figure 14** provide an overview of specific steps, roles, and responsibilities during the two evaluation processes for project terminal evaluations.

Figure 13: Flowchart for project terminal evaluations (non-GEF projects)



Source: UNIDO, 2016: UNIDO Evaluation tools. Guidelines for Terms of Reference for Evaluations, page 13

Figure 14: Flowchart for project terminal evaluations (GEF projects)



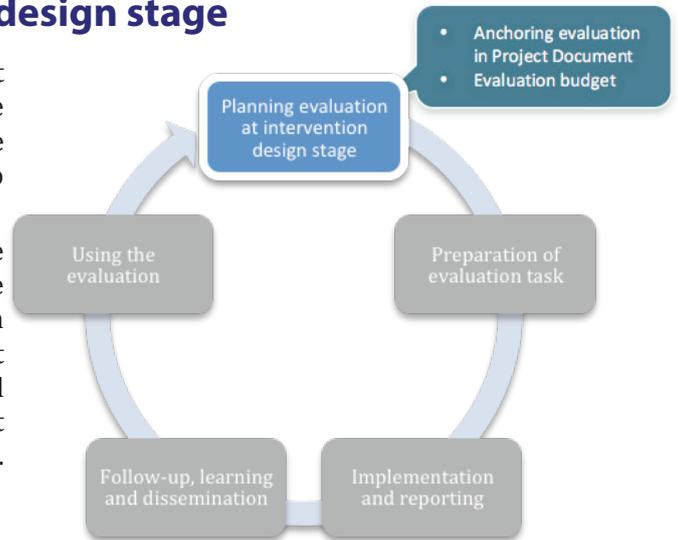
Source: UNIDO, 2016: UNIDO Evaluation tools. Guidelines for Terms of Reference for Evaluations, page 4

The following sub-sections describe the process of a non-GEF project terminal evaluation.

4.1 Planning an evaluation at the design stage

To assess development results and impact, a project terminal evaluation needs to be planned for from the design stage of the intervention. This is good practice in UN specialized agencies²² and also applicable to UNIDO.

The Project Document is central to anchor the evaluation for 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.



4.1.1 Evaluation planning and budget

For projects requiring a mandatory terminal evaluation (by UNIDO threshold of EURO 2 million, Donor requirement, project requirement), UNIDO Project Managers (PMs) should make sure that the project design and budget include the respective Monitoring and Evaluation output. The M&E plan should outline key M&E activities/tasks (e.g. annual implementation reviews, mid-term review and independent terminal evaluation) and timeframe during the implementation. Sufficient funds are to be budgeted for those M&E activities in the Project Document. The Project Managers need to issue the first notification to the Independent Evaluation Division with the indicative timing of an evaluation at the project design stage to avoid ad-hoc requests to the Independent Evaluation Division. In the case of modifications to the original timeframe, the Independent Evaluation Division needs to be informed at least 6 months in advance an evaluation taking place.

PM to-do list:
 o Design M&E plan
 o M&E budget line

The average cost of an independent Terminal Project Evaluation is around EUR 40,000-70,000 (around 2% to 3.5% of the total project budget), and also depending on the project size, complexity, and geographic coverage. In case of regional projects, the terminal evaluation budget would be around EUR 70,000 to 100,000.

4.2 Preparation of evaluation task

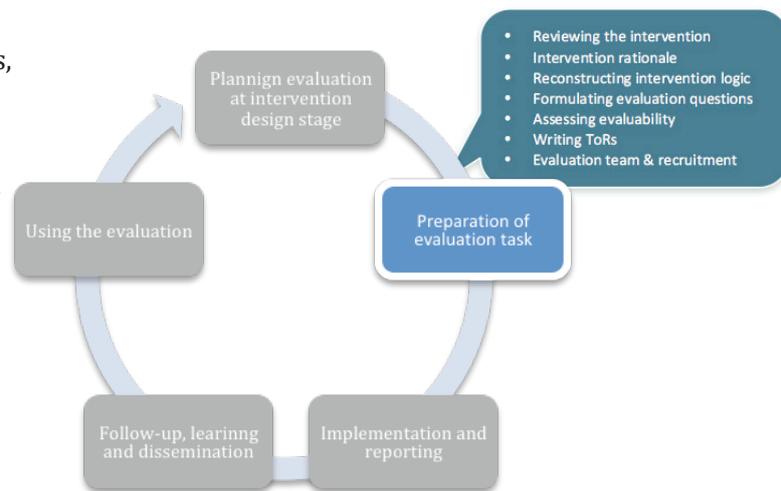
While keeping the principle of independence, UNIDO evaluations need to be participatory and inclusive (see section 3.9). In this spirit, stakeholders should be involved as early as possible in the preparation of evaluations. The extent of stakeholder engagement should be balanced with considerations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

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

22 See for example UNODC, 2012: Evaluation handbook.

As good practice in the UN System, the timeliness, and intentionality of the evaluation is to be addressed from the outset of the evaluation planning. Timely information meeting the needs of intended users should enable decision-making processes.

Every independent evaluation in UNIDO is assigned with a responsible **Evaluation Officer (EO)** from the Independent Evaluation Division, who will manage, facilitate, quality assure and monitor all the respective evaluation process.



4.2.1 Developing Terms of Reference

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

Before applying for as a service provider for an evaluation in UNIDO, evaluation consultants are referred to the ToR as the first step to learning about evaluation requirements.

The key components for any evaluation Terms of Reference are:

- The evaluation context and purpose;
- 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, which will be validated and confirmed by the inception report;
- Evaluation methodology;
- Management arrangements;
- Expected deliverables; and
- The evaluation process and timetable.

UNIDO Guidelines for Terms of Reference specify that writing ToR is a consultative process led by the Evaluation Officer. For project evaluations (other than GEF project evaluations) the drafting of the ToR is delegated to the Project Manager. For GEF project evaluations, and thematic and country evaluations, the assigned Evaluation Officer of the Independent Evaluation Division is responsible for writing the Terms of Reference.

The importance of the consultative nature of writing ToR is emphasized in the box below.

PM to-do list:

- Prepare & circulate ToR

EO to-do list:

- Feedback, clearance & QA

For GEF evaluations:

PM to-do list:

- Inform about timing of evaluation
- Inputs to ToR and circulation to stakeholders

EO to-do list:

- Draft and clear ToR

“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’s Standard 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”²³.

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”²⁴. Those rating criteria are described in the relevant section above.

Complying with UNIDO guidelines ensures high-quality ToR. The latter are the cornerstone for a quality evaluation. Evaluation managers also benefit from both an UNIDO and UNEG checklist for evaluation ToR to ensure the consistently high quality of ToR in UNIDO^{xxvi}.

Section info box: Intranet

[UNIDO Guidelines for Terms of Reference](#)
[UNIDO evaluation tools. Guidelines for Terms of Reference for Evaluations.](#)



4.2.2 Reviewing the intervention

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.

4.2.3 Intervention rationale

UNIDO’s project documents should contain the rationale for the intervention, which should be summarized in the project Logframe. 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”²⁵.

23 UNEG, 2016: UNEG Norms and Standards, UNEG Standard 4.1: Timeliness and intentionality, page 21.

24 UNIDO, 2016: UNIDO evaluation tools. Guidelines for Terms of Reference for Evaluations. Page 2.

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

For thematic or country evaluations, such a “Project Document” might not be available. However, thematic or country strategies 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.

4.2.4 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 UNIDO²⁶.

The evaluation manager should verify: i) the specificity of the intervention logic; ii) the existence and quality of baseline, monitoring, and implementation data; iii) the availability of key informants; iv) the timing of the evaluation in relation to the intervention cycle; and v) opportunities to make use of planned stakeholder meetings for evaluation interviews.

The evaluability assessment helps to find out to what extent the proposed evaluation questions can be answered, potentially leading to modifications of the evaluation design, the reformulation of evaluation questions or the timing of the evaluation.

4.2.5 Evaluation team composition and recruitment

The UNIDO Guidelines for Terms of Reference elaborate on the composition of the evaluation team. The specific Evaluation Team requirements are laid out in the respective ToR for the evaluation, based on the evaluation subject, focus, methods, and analyses required. A good skills mix is necessary and interdisciplinary teams prove to be of added value. It is important to keep in mind that Independent Terminal Project evaluation^{xxvii} is more about assessing actual results, changes and effects of the interventions and learning out of it, and should not be mixed or considered as a “technical” study (which can always be conducted by the PM as part of the project). The evaluation team should have an adequate background 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 with thematic expertise 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 the project design and implementation and free of any conflict of interest. The latter is particularly important for highly specialized technical experts engaged in global discourse in sub-sectors with very limited human resources available for evaluations. Evaluations might be misused to underpin their arguments or school of thought in the global discourse.

EO to-do list:

- Identification of consultants
- Invite PM to propose evaluators, particularly national consultants
- Address donor needs as applicable
- Selection

PM to-do list:

- Recruitment of evaluator(s)

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
- 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

Source: UNIDO, 2016: UNIDO evaluation tools. Guidelines for Terms of Reference for Evaluations. Evaluation team composition, Page 11.

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 by an evaluation. UNEG calls for strong evaluation expertise in evaluation teams: "***the core members of the evaluation team must be experienced evaluators with appropriate methodological expertise***"²⁷. This pledge is backed up by evidence²⁸, as shown in the box below.

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

"United Kingdom-based Center for Development Impact, of the Institute of Development Studies at the University of Sussex, 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 Center for Development Impact found that this is supported by other research in this area. United States Agency for International Development (USAID) found that USAID evaluations with 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: Center 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.

27 UNEG, 2016: Norms and Standards for evaluation, page 25.

28 Center 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.

For UNIDO evaluations, evaluation consultants or the evaluation manager are the team leaders and technical experts support the 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, as the availability of good evaluators is limited. Well-qualified evaluators must often be contacted months in advance, from at least three to six months. “Recruiting evaluators at the last minute may considerably reduce prospects for obtaining a good evaluation”²⁹.

Section info box

UNIDO Employment Opportunities - online platform
<http://www.unido.org/employment.html>



- PM to-do list:**
- Provide documents
 - Inform Gov., Perm. Mission, donor(s), stakeholders in HQ and in the field and respective area office(s)
 - Organize HQ briefing of consultants

The recruitment process itself may also add to the time required before the evaluation can start. UNIDO’s evaluation manager can accelerate the recruitment process by indicating the requirements for contract signature such as a completed Personal History form, creating or updating a candidate profile on the “UNIDO Employment Opportunities” online platform, passing the basic or advanced security in the field test and, as required, supplying a medical certificate or security clearance. The evaluation manager has the responsibility for selecting the evaluator(s). This might be done following prior consultation

with the donor or other partners.

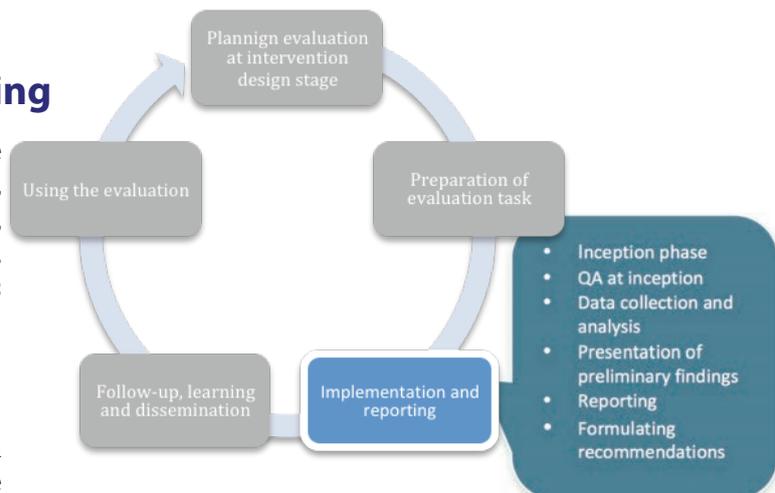
Before the implementation of the evaluation begins, some processes need to happen, shown in the adjunct post-it note.

4.3 Implementation and reporting

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

4.3.1 Inception phase

UNIDO’s guidance on evaluation inception reports^{xxviii} describes the inception phase as the point during the evaluation process where “the evaluation team reviews project/programme



29 Sida, 2007: Looking back, moving forward. Sida Evaluation Manual. 2nd revised version. Page 79.

documents, analyses the ToR, and develops a detailed proposal for the implementation of the evaluation. The results are laid down in an inception report that is submitted to UNIDO for review and approval. The continuation of the evaluation process is conditional on UNIDO Independent Evaluation Divisions approval of the report”.

During the inception phase, the evaluation team can further operationalize the ToR. The evaluation methodology can be finalized, key questions and focus defined, and sampling strategies and criteria clarified. The intervention logic or a theory of change can be reconstructed during the inception phase.

PM to-do list:

- Feedback to inception report

EO to-do list:

- QA, feedback, clearance

The inception report contains, among other elements, the evaluation work plan with a concrete timetable, 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 page: Intranet

[UNIDO evaluation inception report: standard format](#)



4.3.2 Quality assurance of evaluation delivering on ToR in inception phase

Experience shows that the inception stage can be used to refine the evaluation to implement the Terms of Reference but also to deviate from the Terms of Reference. The evaluation manager’s close oversight of the evaluators is required to ensure that the scope and objectives of the Terms of Reference are cherished. Any deviations should be clearly explained. Once the inception report is approved by the evaluation manager, the evaluators move to the main phase of data collection and the approach agreed in the inception report cannot be rectified any more.

Example of how an inception report can deviate from the original Terms of Reference

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 was to be expected after carefully reading the inception report. Quality assurance of the evaluation manager failed at inception.

4.3.3 Field work

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

In UNIDO, strategic evaluations may benefit from a preparatory mission and the main evaluation mission. The preparatory mission helps the evaluation manager to further familiarize with the country and intervention context. At the same time, the evaluation manager can seek last inputs from partners into the evaluation ToR such as an updated list of intervention stakeholders and brief partners about the evaluation methodology. The preparatory mission is also an opportunity to identify suitable national consultants. Assessing the availability of data and identifying data sources is another benefit from the preparatory mission. About one week is required for a preparatory mission, including time in intervention areas. Depending on budget availability, the preparatory mission is sometimes carried out over phone, emails or through the UNIDO field representation.

PM to-do list:

- Prepare & support evaluation mission

The actual Evaluation field mission follows the preparatory mission after about month. Project evaluations normally contain the main evaluation mission only. Among the core objectives of the main mission are:

- i) Data collection to build the evidence trail
- ii) Validation and supplementing desk review results
- iii) Interaction with project stakeholders, including beneficiaries
- iv) Direct observation of stakeholder behavior and project results

Country 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 and meeting the right stakeholders. Visits to more remote projects sites allow for insights into the reality on the ground beyond the capital cities.

In this context, it is important to recall that field visits aim to fill important information gaps and are not expected to assess all activities undertaken by the project or programme or to undertake large scale household surveys. Complementing existing information and the validation role of field visits are valuable.

To manage the challenges of time and budget constraints of field visits, it is recommendable to undertake case studies. Also, it is good practice to concentrate on components in need of 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 on-site 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 stakeholders comprise the project management team, governmental counterparts, Chief Technical Advisor and representatives of the donor, among others.

4.3.4 Presentation of preliminary findings

After the evaluation field mission, it is good practice to have an internal debriefing at HQ to present preliminary findings. The audience for this debriefing can be the Project Manager, evaluation manager, managers of the division and departments or a reference group at the headquarters in Vienna.

PM to-do list:

- Organize HQ debriefing
- Debriefing of consultants

Sharing preliminary findings 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 evaluating results^{xxix}.

“The debriefing session will be instrumental in ensuring ownership”

Source: Vienna-based specialised UN agency

The debriefing can take place either before the reporting is started (for example at the end of a field visit) or at the stage when the draft report is written and even recommendations ready for discussion.

4.3.5 Reporting

Once all data is captured and analyzed, reporting can start based on the agreed reporting outline. The reporting outline is specified either in UNIDO’s evaluation Terms of Reference 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^{xxx}. The report must comply with the UNIDO Quality Checklist for Evaluation Reports, which is attached to the evaluation Terms of Reference^{xxxi}.

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

The responsibility for delivering a draft report on time according to the timeframe agreed in the inception report lies with the evaluation TL. The EO 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 also submits its own feedback on the draft report to the evaluation team leader on time.

As a next step, the evaluation TL is responsible for addressing all the feedback received. Due to the independent nature of UNIDO independent evaluations (with exception of self-evaluations), the feedback is not binding for the evaluation TL but for addressing factual errors. To uphold transparency, it is good practice to keep a log of how the feedback was considered in the final report, again 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. He or she might consider

using the services of a professional editor for this purpose at his/her expense.

The final report (in one document as MS Word, including annexes) is submitted by the evaluation team leader to the evaluation officer, who will conduct the final QA and proceed with the final document preparation (cover page, registry) and preparation of the e-book for distribution and dissemination. The executive summary is required in English, regardless of the language of the main report.

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

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

4.4 Follow-up, learning, and dissemination

4.4.1 Management response to evaluations and internal follow-up

UNIDO's Evaluation Policy defines management response and follow-up processes, roles, and responsibilities by evaluation type³⁰.

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

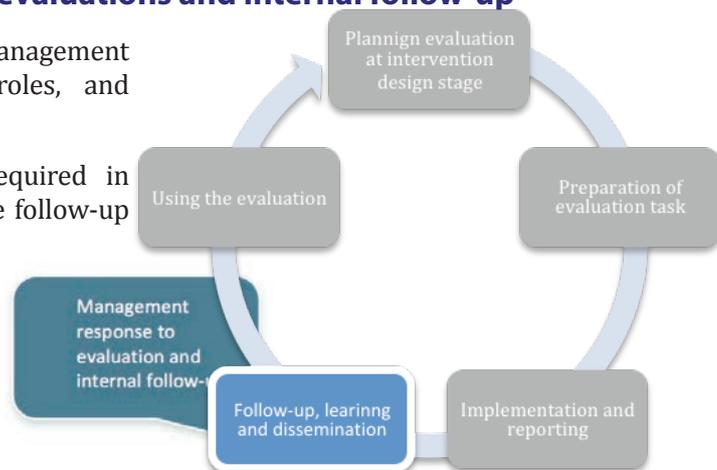
For independent evaluations, UNIDO Evaluation Policy outlines that³¹:

The Evaluation Report together with a management response sheet (MRS) is disseminated to relevant project stakeholders. This sheet enables tracking for each recommendation that covers the comments of acceptance or non-acceptance of evaluation recommendations, the deadlines, and action taken by those responsible for follow-up.

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

The Independent Evaluation Division monitors the information in the MRSs and compiles periodic reports on the status of those recommendations and follow-up their implementation; and draws the attention of UNIDO management as needed.

The Independent Evaluation Division reports to the Member States, among other things, on the



30 UNIDO, 2015: Director General's Bulletin: Evaluation Policy. UNIDO/DGB/(M).98/Rev.1, page 11.

31 UNIDO, 2015: Director General's Bulletin: Evaluation Policy. UNIDO/DGB/(M).98/Rev.1, page 12.

status of follow-up to recommendations on a biannual basis.

Besides, ad-hoc status requests on this subject can be received by the Independent Evaluation Division, for example from the External Auditor. In the case of staff leaving, a prompt handover and information/confirmation of successors who will assume the responsibility to follow-up on evaluation recommendations is required to ensure an uninterrupted follow-up process. The Independent Evaluation Division maintains an electronic management response system available on the Independent Evaluation Division's site on the UNIDO Intranet through which the response status can be tracked.

EO to-do list:

- Final formatting & Management Response Sheet
- Submission of evaluation report & MRS to: PM & stakeholders, MRS respondents, ODG/EVQ

PM to-do list:

- Implement & follow-up recommendations in MRS

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

In the last section of this Evaluation Manual, background information is provided about strategic evaluations. The section clarifies objectives and methodology, describes the process and specifies reporting requirements.

5.1 Background

The evaluation process in strategic evaluations (thematic, country or impact evaluations) is in principle equivalent to the independent project evaluation process.

Main differences are, for example, the evaluation planning, which is rooted exclusively in the biannual evaluation work plan of the Independent Evaluation Division.

The governance structure for those strategic evaluations also shows differences, with the Independent Evaluation Division being fully in charge of and responsible for the evaluations.

The evaluation management, under the responsibility of the assigned Evaluation Officer (EO), can benefit from a Reference Group to enhance in-house understanding and ownership particularly of complex thematic or country evaluations.



5.2 Country evaluations

5.2.1 Purpose

UNIDO’s country evaluations aim to provide an assessment of UNIDO’s country engagement, specific results and impact at country level 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 concerned country, and any other country partners.

5.2.2 Scope

While projects can be the main source of direct country engagement, the country evaluation also focuses on the country level specific results of: i) UNIDO’s analytical and policy advisory services; ii) standard-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.

It is good practice to cover UNIDO’s cooperation in the country over the previous six to ten years, if possible, to build on a critical mass of evidence. This timeframe also enhances the probability to track change stipulated by an UNIDO intervention.

Country evaluations also represent a unique opportunity for assessing the sustainability of previous UNIDO interventions and verify to what extent ownership and project results are still producing or consolidating the expected change and impact.

Historically UNIDO country evaluations used to be a summary of project evaluations only. Recent good practice in UNIDO uses a cluster approach to group projects thematically and includes the performance of the UNIDO Office and country specific efforts such as the UN Delivering as One approach. The cluster approach seems highly relevant, given that the UNIDO project portfolio to be evaluated consists of an average of 10 projects (see box above)^{xxxii}.

The added value of country evaluations for UNIDO beyond summarizing project level results shows for example in the case of the Country Evaluation Pakistan (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 Country Evaluation Pakistan, page 10.

5.2.3 Selection and timing

UNIDO’s biannual work plan determines the selection of country evaluations. To maximize the utility of country evaluations, the selection is guided by the country’s policy cycle, the rotation cycle of UNIDO country representatives and pending decisions about the future degree of supporting

the country. UNIDO’s overall investment of the project portfolio and the total number of projects can also guide the country selection. It is good practice to state why the evaluation was needed at that specific point in time, beyond the fact that the evaluation was included in the Independent Evaluation Division’s work plan. Clients of the evaluation should be listed as well as the likely use of evaluation results. Demand-driven country evaluations seem particularly useful, as in the case of the Country Evaluation of the Russian Federation (2014), following a demand from the Russian Federation Permanent Mission to the Vienna-based UN agencies.

5.2.4 Process and methodology

The main reference points for the evaluation are UNIDO’s Country Strategies and the countries’ industrial development strategies. Past country evaluation often mentioned the lack of a UNIDO Country Strategy but would disregard the countries’ industrial development strategies to assess UNIDO’s performance. Evaluating against expected results is important, and project targets, with baselines and indicators, are one means to assess UNIDO’s contribution to the countries’ industrial development strategies.

The evaluation process for country evaluations is theory-based, following the good practices of UNIDO’s peers in the UN System. This allows building on a strong context analysis and on how the country engagement aimed to lead to results and impact, based on explicit assumptions. Section 4 of this Evaluation Manual explains the theory of change approach in more detail.

The evaluation should be guided by the respective ToR and the inception report, including a table indicating evaluation questions, information sources and methods for data collection and analysis. This is the basis for the evaluation questionnaire. Roles of the evaluation team can be defined through the evaluation framework, for example for reporting and field visits.

Project evaluation results for the specific country are aggregated to the extent possible, including ratings to identify patterns and trends by clusters. Evaluating against expected results is important and one of the less strong areas of previous country evaluations.

The country evaluation establishes linkages to other UNIDO core functions at the country level where possible. To keep the evaluation focused and affordable, case studies might be elaborated for the processes and results of UNIDO’s analytical and policy advisory services, standard-setting and compliance and UNIDO’s convening and partnership role.

Figure 15 summarizes common challenges faced in UNIDO country evaluations and lists possible solutions.

Figure 15: Common challenges in UNIDO country evaluations and possible solutions

Lack of documentation, including baseline data and monitoring data	UNIDO Independent Evaluation Division to coordinate selection of project and country evaluations to increase the evaluative evidence for the given country
Insufficient availability of evaluative evidence	
Large number of projects to be evaluated (average of 10) during 3-week field visit if no cluster approach is taken	i) Using good practice of clustering projects thematically; ii) sampling within a cluster; iii) focus sampling on projects with no evaluative evidence
Difficulties in creating a single theory of change for the country programme	Unless a country programme exists, it might not be possible to create a theory of change for the entire project portfolio. However, creating theories of change by cluster is an option to still benefit from using a theory-based evaluation approach.

5.2.5 Report

For all strategic evaluations, the main report should contain about 30 pages (excluding annexes), with a 2 to 3-pages Executive Summary and a one-page briefing sheet.

Executive summary

A meta-evaluation of UNIDO country evaluation undertaken between 2014 and 2016 using the UN Evaluation Group's quality checklist for evaluation reports showed that 6 to 9-pages executive summaries tend to be the rule rather than the exception. The Country Evaluation Uruguay (2015) structured the executive summary along the evaluation criteria, which helped to create a well-structured, precise and short executive summary. This good practice example contrasts with lengthy executive summaries struggling to summarize results by project or intervention area.

Presenting the key recommendations in the executive summary as one-line bullet points also helps to shorten executive summaries, as in the case of the Country Evaluation Tanzania (2016). At the same time, key recommendations become clearer and more concise, which is also the purpose of an executive summary.

Main report

UNIDO's standard reporting template is the starting point of a well-structured, logical, clear and complete report. To truly reflect the scope of the country evaluation, the report should be structured along the selected evaluation criteria rather than the results of individual projects. Project summary sheets containing the most important data and findings of the projects within the clusters evaluated should be annexed. The Country Evaluation of the Russian Federation (2014) excelled with annexing those useful project summary sheets to shorten the main report while strengthening its analytical nature. UNIDO country evaluations are particularly visual regarding the use of country maps, ideally also showing the location of the main project sites.

The use of *graphics, tables, and figures* also helps to visualize evaluation findings and to show performance patterns. This can be achieved by i) comparing expected project, cluster or country programme results (outputs and outcomes, by indicators) with the actual results at the time of the evaluation; ii) capturing quantitative data by systematically using an evaluation questionnaire.

In this respect, consideration should be given to the application of *evaluation questionnaires* as a means to substantially and systematically underpin the country evaluations' evidence base. This is further explained in section 5.3.4 for thematic evaluations.

Figure 16: Good practice in identifying and presenting key evaluation findings

Criteria	Findings	Evaluators' rating
Effectiveness	<p>Finding 4: Projects in the 1st cluster depict varying degrees of achievement of results. The most notable successful intervention is the strengthening of the analytical technical capacity of the National Bureau of Statistics (NBS) and staff of the Ministry of Industry, Trade and Investment (MITI).</p> <p>Finding 5: It appears from the available data that all projects in the 2nd cluster had either achieved their intended results or were on track to achieve them. However, data is incomplete. With some exceptions, for the projects funded by the UN One Fund there are no project documents available which clearly stated expected results for UNIDO and it is therefore unclear if all results have or will be achieved.</p> <p>Finding 6: Results in GEF-funded projects were clearly identified at the design stage and it can be concluded that they had been achieved or were likely to be achieved. The results of projects funded from the One Fund were identified in a more generic manner, some achievements could be identified, though.</p> <p>Finding 7: Several examples demonstrate that there is potential for synergies between different UNIDO projects. However, overall there are limited synergies between the different UNIDO projects. Reasons are the conception of the projects as individual unrelated interventions, funding uncertainties or structural issues. This led to rather isolated projects, fragmented geographically, thematically and institutionally, lending themselves to limited synergies.</p>	satisfactory
Long-term benefits and changes (impact and sustainability)	<p>Finding 8: Overall, there is a reasonably strong ownership among stakeholders of the UNIDO supported activities. In some cases, however, the ownership is limited because of perceived rather rigid, top-down implementation.</p> <p>Finding 9: Long term benefits of UNIDO projects can be found at macro, meso and micro level. There are also examples of potential multiplier effects. The outreach at the micro level to end beneficiaries is generally rather small. At times, the socio-economic context limits the impact of UNIDO interventions.</p> <p>Finding 10: There are no explicit theories of change used in the UNIDO Country Programme. Instead, the Country Programme is using logframes, which are also used at the project level. The absence of theories of change might have contributed to weakening the coherence of the interventions, the limited synergies and limited anticipation of socio-economic challenges.</p>	moderately satisfactory
Gender and youth	<p>Finding 11: While there are many references to gender and youth in the UNDP/UNIDO Country Programme, at the project level gender and youth are not fully mainstreamed. Youth receives even less attention compared to the gender dimension. However, it can be assumed that the UNIDO projects benefit men and women equally and to a limited extent also the young population.</p>	moderately unsatisfactory

Source: Independent UNIDO Country Evaluation United Republic of Tanzania, 2016, page 70

The use of *key evaluation findings* for each evaluation criterion helps to underpin the logic of the evaluators to come to conclusions and recommendations. The Tanzania CPE (2016) sets the pace regarding a strong evaluation logic by identifying key evaluation findings by evaluation criteria. **Figure 16** shows the successful attempt to transform a country evaluation from a mangle-mangle of project evaluations to a truly strategic institutional evaluation.

The evaluation matrix and interview protocol/questionnaire used for stakeholder consultations should also be annexed.

Where feasible, using rating results in the executive summary and throughout the report is a good practice that deserves systematic application across country evaluations. Ratings by evaluation criterion can be the result of aggregating the respective project level ratings. The evaluation report should state whether the ratings are based on the evidence collected and aggregated evaluators' judgment or on a particular source of feedback such as the responses of interviewees.

The Country Evaluation of the Russian Federation (2014) serves as a good practice example to further strengthen the logic between key evaluation findings, conclusions, and recommendations. In this evaluation, conclusions and recommendations are listed in tabular format side by side. This is the strongest proof of logically linking evaluation conclusions and recommendations. Good practices in UNIDO's peers even suggest combining key findings, conclusions, and recommendations in a table. This can be presented either as part of the executive summary, at the beginning of a dedicated section of or in an annex. **Figure 17** shows a good practice example of an UNIDO peer to combine key findings, conclusions, and recommendations in one table.

Figure 17: Good practice example in proofing logic between key evaluation findings, conclusions and recommendations

	Key findings	Conclusions	Recommendations
Donor policy priorities	Fully in line with the Government of country A's trade policy plan	High relevance is given	No recommendation
Skills needs	High relevance for skills and business development needs for MSMEs in country A and B (85% high to very high ratings); key need of access to finance, particularly in country A still to be met	Programmes do mainly address skills and business development needs in both countries	Implementer: Step up efforts to facilitate access to finance for MSMEs, given the high relevance of both programmes. High priority: action next 3 months
Country level relevance	Intervention operates outside the United Nations Development Assistance Plan in country B	Value added of the donors' multilateral investment in country B is questionable	The Donor is advised to rethink whether the country focus on country B is adding any value to its bilateral cooperation portfolio. Medium priority: action next 6-9 months

Color code:

Highly satisfactory	Satisfactory	Moderately satisfactory	Moderately unsatisfactory	Unsatisfactory	Highly unsatisfactory
---------------------	--------------	-------------------------	---------------------------	----------------	-----------------------

A good mix of qualitative and quantitative data, including the quantification of qualitative results, enhances the strengths of the evidence base. Clearly presented and robustly backed-up evaluation results help to inform evidence-based decision making.

5.3 Thematic evaluations

5.3.1 Purpose

Thematic evaluations intend to provide timely, credible and useful information that can contribute to inform UNIDO's thematic, programmatic and strategic dimensions, and to enable evidence-based decision making. UNIDO's peers have experienced the strong influence of thematic evaluations on providing guidance at the strategic and policy level.

5.3.2 Scope

UNIDO's thematic evaluations address systemic issues and cut across TC projects, UNIDO core functions, countries and regions.

As for country evaluations, the thematic evaluation assesses the results of: i) UNIDO's analytical and policy advisory services; ii) standard-setting and compliance; and iii) UNIDO's convening and partnership role on a specific theme. This is complemented by relevant evaluative evidence from TC projects.

5.3.3 Selection and timing

As for country evaluations, UNIDO's biannual work plan determines the selection of thematic evaluations in UNIDO. The selection of thematic evaluations should be guided by topics that are relevant on the agenda of UNIDO, partner governments and donors, or topics where evaluative evidence is lacking for decision making about future strategic investments. The latter also influences the timing of thematic evaluations. As for country evaluations, it is good practice to state why the evaluation was needed at that point in time. 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 exercise was commissioned "to feed into the development and implementation of the Inclusive and Sustainable Industrial Development vision and strategy"^{xxxiii}. Clients of the evaluation should be listed as well as the likely use of evaluation results to provide a complete overview of the evaluation purpose.

5.3.4 Process and methodology

For thematic evaluations, the main reference points are UNIDO's policies, thematic strategies or the thematic strategies of regional bodies or partners of UNIDO.

Again, as for country evaluations, UNIDO's thematic evaluations are theory-based. The evaluation should be based on a logic model, theory of change, and/or the expected results chain. If unavailable, those should be reconstructed based on the strategic documentation available and interviews with key stakeholders. Evaluators aggregate to the extent possible 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 the TC projects.

Thematic evaluations in UNIDO usually rely primarily on past project evaluations. Primary data collections related to TC projects through field visits is the exception, given the budget and time constraints.

The box below highlights the consequences on relying on past project evaluations, pointing towards issues of representativeness, time-lag and limitations of data on specific evaluation criteria:

“This has implications on the sample selection as only projects with available evaluations could be included and not necessarily projects that would best represent the different types of (thematically relevant) initiatives. Moreover, relying on past evaluations entails an important time lag, as some of the findings may be based on past, rather than current practice. Further, evaluation reports often provided only limited information on impact in terms of contributions to employment or poverty reduction”.

Source: UNIDO Evaluation Group, 2013: Independent Thematic Evaluation. Thematic Evaluation of UNIDO projects related to “Industrial Upgrading”, page x.

Figure 18 summarizes common challenges encountered in UNIDO thematic evaluations^{xxxiv}. Suggestions on how those challenges can be addressed are also listed.

Figure 18: Common challenges in UNIDO thematic evaluations and possible solutions

<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 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 Division 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 on-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"> SAP system did not allow a clear-cut allocation of resources to 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

For the verification of secondary data, the use of evaluation questionnaires in interviews or a survey are recommendable. In fact, 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 as a means of 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 strategically for data collection, the evaluation's evidence base would be insufficiently strengthened.

Using comparators from other similar organizations to benchmark UNIDO's strategic approach and performance for specific thematic issues proved 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.

5.3.5 Report

For all strategic evaluations, the main report should contain about 30 pages (excluding annexes), with a 2 to 3-pages executive summary and a one-page briefing sheet. The report should include a self-standing section on methodology and limitations encountered^{xxxv}.

Findings

Presenting key findings at the beginning of subsections is a good practice example. This presentational approach helps for systematic and appropriate data analysis, as successfully practiced in recent thematic evaluations^{xxxvi}. As stipulated for country evaluations, the use of ratings is useful for the specification of evaluation results beyond a narrative. The tables in **Figure 16** and **Figure 17** in section 5.2.5 are also applicable for the systematic presentation of thematic evaluation results.

UNIDO's thematic evaluations are particularly strong in the graphical presentation of secondary data to summarize trends, patterns, strengths and shortcomings. A practice to be continued.

Conclusions

All evaluation reports should have a section on conclusions, including thematic evaluations. It is good practice to structure conclusions along the lines of evaluation criteria, following key findings structured in the same way. Evaluators can provide headline messages in the conclusions section for example "positive trend" or "no clear strategy", as practices in the Independent Thematic Evaluation: "UNIDO's Public Private Partnerships" (2014). Conclusions logically lead to evaluation recommendations.

Recommendations

UNIDO's thematic evaluations, in general, shy away from targeting specific units inside or outside the Organization with evaluation recommendations. This is the case in the last seven thematic evaluations (2014-2015). However, it is important to target evaluation recommendations to enhance the likelihood that the correct stakeholders take action and also to indicate a prioritization of those recommendations.

5.4 Impact evaluations

Multilateral organizations are facing an increasing demand for results and particularly for evidence-based impact from development interventions. UNIDO responds to this demand with the evaluation type of impact evaluations. The methodology section of this Evaluation manual describes challenges to evaluate impact and impact evaluation domains and techniques.

5.4.1 Objectives

The objective of an impact evaluation is to see whether a particular UNIDO intervention or a group of interventions or a programme have actually produced or clearly contributed to the broader and sustainable change, and had a lasting effect on its target population. The impact evaluation wants to examine to what extent shared prosperity was actually created, economic competitiveness advanced and/or the environment safeguarded as a contribution of UNIDO.

5.4.2 Purpose

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

5.4.3 Scope

The unit of assessment is typically a programme (rarely would it be a single project, unless it went through several phases, i.e. equivalent to a programme). The programme should be of a mature nature, as the time-lag for change to happen is significant. In other words, projects or programmes in a second or third phase are particularly interesting for evaluating the impact. Alternatively, impact evaluation can be ex-post, up to 10 years after the end of an UNIDO intervention. In this case, however, tracing stakeholders and records/data might be challenging.

5.4.4 Selection and timing

Impact evaluations are in general more expensive than other strategic evaluations. Hence projects or programmes should be selected based on whether they have a particular strategic importance for UNIDO and its partners.

5.4.5 Process and methodology

There is no recipe or one-only methodology for conducting impact evaluation. The approach to impact evaluation is being permanently discussed in the international development and evaluation community over the last years.

The use of randomized control trials and other experimental and quasi-experimental approaches are considered as the “gold standard” of impact evaluation but face in practice significant limitations in complex social and economic settings. Methodological limitations 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 have faced several implementation challenges, combined with the burden of exploding evaluation budgets and the demand for vast amount of data and long timeframes for this type of impact evaluations. 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 into account the overall level of available resources of the Independent Evaluation Division.

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 in other specialized UN agencies.

Theory-based impact evaluations

A particularly cost-effective way for impact evaluations compared to experimental and quasi-experimental approaches is the use of a theory-based approach, as successfully applied by one of UNIDO’s main partners, the GEF^{xxxvii}. The approach entails collecting evidence towards 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) 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 document and analyze change processes along the lines 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.

Like for 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 of UNIDO are incorporated into broader stakeholder mandates and initiatives such as laws, policies, regulations, and programs. 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.

An important dimension of assessing progress towards impact is behavior 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 to 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.

5.4.6 Report

For all strategic evaluations, the main report should contain about 30 pages (excluding annexes), with a 2 to 3-pages executive summary and a one-page briefing sheet. Key annexes should include the tools used, case studies, and further details as needed.

Annex 1: References to good evaluation practices

[UNIDO Evaluation Policy](#)

Intranet: [UNIDO evaluation templates and guidelines](#)

[DAC RBM glossary](#)

Quality checklists

UNEG quality checklist for evaluation ToR and inception reports (2010)

www.uneval.org/document/detail/608

UNEG checklist for evaluation reports

www.uneval.org/document/detail/607

Theory of change

Aspen Institute:

http://www.dochas.ie/Shared/Files/4/TOC_fac_guide.pdf

UNICEF:

https://www.unicef-irc.org/publications/pdf/brief_2_theoryofchange_eng.pdf

London School of Economics/The Asia Foundation

<http://www.lse.ac.uk/internationalDevelopment/research/JSRP/downloads/JSRP17.Valters.pdf>

Data collection techniques:

OIOS, 2014: Inspection and Evaluation manual, pages 53-80

https://oios.un.org/resources/2015/01/OIOS-IED_Manual.pdf

Data analysis techniques

OIOS, 2014: Inspection and Evaluation manual, pages 84-91

https://oios.un.org/resources/2015/01/OIOS-IED_Manual.pdf

Checklist for evaluation terms of reference

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

https://www.iom.int/jahia/webdav/site/myjahiasite/shared/shared/mainsite/about_iom/eva_techref/UNEG_TOR.pdf

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.	Yes
		✓ <u>Mainstreaming</u> : 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?	
		✓ <u>Replication</u> : To what extent are the project's specific results (for example methodology, technology or lessons learned) reproduced or adopted?	
		✓ <u>Scaling-up</u> : To what extent are the project's initiatives and results implemented at larger geographical scale?	
		✓ What difference has the project made to the beneficiaries?	
		✓ What is the change attributable to the project? To what extent?	
		✓ What are the social, economic, environmental and other effects, either short-, medium- or long-term, on a micro- or macro-level?	
		✓ What effects are intended or unintended, positive or negative? The three UNIDO impact dimensions are:	
		✓ <u>Safeguarding environment</u> : Biophysical changes in reduction of threats emanating from action of humans and changes in the status of the environment.	
		✓ <u>Economic performance</u> : 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	
	✓ <u>Social inclusiveness</u> : 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
		✓ Is the problem, need or gap to be addressed by the project clearly identified, with clear target beneficiaries?	
		✓ Was the project design adequate to address the problems at hand?	
		✓ 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?	
		✓ 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?	
		✓ 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?	
		✓ Does it include M&E plan and adequate budget for M&E activities?	
	✓ 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?		

2	Logframe	<p>Assessment of the logical framework aimed at planning the intervention</p> <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? ✓ 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? ✓ 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? ✓ Are key assumptions properly summarized and reflecting the proper level in the results chain in the logframe? 	Yes
C	Project performance	Functioning of a development intervention.	Yes
1	Relevance	<p>The extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor</p> <ul style="list-style-type: none"> ✓ How does the project fulfil the urgent target group needs? ✓ To what extent is the project aligned with the development priorities of the country (national poverty reduction strategy, sector development strategy)? ✓ How does the project reflect donor policies and priorities? ✓ Is the project a technically adequate solution to the development problem? Does it eliminate the cause of the problem? ✓ To what extent does the project correspond to UNIDO's comparative advantages? ✓ 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? 	Yes
2	Effectiveness	<p>The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance</p> <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? ✓ 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? 	Yes

3	Efficiency	A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results	Yes
		✓ How economically are the project resources/inputs (concerning funding, expertise, time...) being used to produce results?	
		✓ To what extent were expected results achieved within the original budget and timeframe? If no, please explain why.	
		✓ Are the results being achieved at an acceptable cost? Would alternative approaches accomplish the same results at less cost?	
		✓ What measures have been taken during planning and implementation to ensure that resources are efficiently used? Were the project expenditures in line with budgets?	
		✓ Could more have been achieved with the same input?	
		✓ Could the same have been achieved with less input?	
		✓ 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?	
		✓ Have the inputs from the donor, UNIDO and Government/counterpart been provided as planned, and were they adequate to meet the requirements?	
4	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
		✓ Will the project results and benefits be sustained after the end of donor funding?	
		✓ Does the project have an exit strategy?	
		✓ To what extent have the outputs and results been institutionalized?	
		<i>Financial risks:</i> ✓ What is the likelihood of financial and economic resources not being available once the project ends?	
		<i>Socio-political risks:</i> ✓ 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?			
D	Cross-cutting performance criteria	Other important criteria that cut across the UNIDO intervention.	

1	Gender mainstreaming	✓	Did the project design adequately consider the gender dimensions in its interventions? Was the gender marker assigned correctly at entry?	Yes
		✓	Was a gender analysis included in a baseline study or needs assessment (if any)? Were there gender-related project indicators?	
		✓	Are women/gender-focused groups, associations or gender units in partner organizations consulted/ included in the project?	
		✓	How gender-balanced was the composition of the project management team, the Steering Committee, experts and consultants and the beneficiaries?	
		✓	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)?	
		✓	To what extent were socioeconomic benefits delivered by the project at the national and local levels, including consideration of 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>	○	Was the M&E plan included in the project document? Was it practical and sufficient at the point of project approval?	
		○	Did it include baseline data and specify clear targets and appropriate indicators to track environmental, gender, and socio-economic results?	
		○	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?	
		○	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)?	
		○	Does it allocate adequate budget for M&E activities?	
	<i>M&E at implementation</i>	○	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?	
		○	Are annual/progress project reports complete, accurate and timely?	
		○	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 Project Steering Committee to make decisions and corrective actions? Do the project team and managers and PSC regularly ask for performance and results information?	
		○	Are monitoring and self-evaluation carried out effectively, based on indicators for outputs, outcomes and impact in the logframe? Do performance monitoring and reviews take place regularly?	
		○	Were resources for M&E sufficient?	
		○	How has the logframe been used for Monitoring and Evaluation 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?	
		○	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)	<p>Assessment of issues related to results-based work planning, results based M&E and reporting based on results.</p> <p><i>Results-Based work planning</i></p> <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. ○ 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)? ○ Examine the use of the project's results framework/ logframe as a management tool and review any changes made to it since project start. <p><i>Results-based monitoring and evaluation</i></p> <ul style="list-style-type: none"> ○ Verify whether an M&E system is in place and facilitated timely tracking of progress toward project objectives by collecting information on selected indicators continually throughout the project implementation period. ○ 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? ○ 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 Project Steering Committee to make decisions and corrective actions? Do the Project team and managers and PSC regularly ask for performance and results information? <p><i>Results-based reporting</i></p> <ul style="list-style-type: none"> ○ Assess how adaptive management changes have been reported by the project management and shared with the PSC. ○ 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?) ○ Assess how results and lessons derived from the adaptive management process have been documented, shared with key partners and internalized by partners. 	
E	Performance of partners	Assessment of partners' roles and responsibilities engaged in the intervention.	Yes

1	UNIDO	○ Mobilization of adequate technical expertise for project design	Yes
		○ Inclusiveness of project design (with national counterparts)	
		○ Previous evaluative evidence shaping project design	
		○ Planning for M&E and ensuring sufficient M&E budget	
		○ Timely recruitment of project staff	
		○ Project modifications following changes in context or after the Mid-Term Review	
		○ Follow-up to address implementation bottlenecks	
		○ Role of UNIDO country presence (if applicable) supporting the project	
		○ Engagement in policy dialogue to ensure up-scaling of innovations	
		○ Coordination function	
		○ Exit strategy, planned together with the government	
		○ 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.	
		○ To what extent the project has a proper and operational governance system (e.g. PSC with clear roles and responsibilities)?	
○ 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)?			
○ 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, Non-Governmental Organizations (NGOs), civil society and the private sector where appropriate.	Yes
	<i>Design</i>	○ Responsiveness to UNIDO's invitation for engagement in designing the project	
	<i>Implementation</i>	○ Ownership of the project	
		○ Provide financial contribution as planned (cash or in-kind)	
		○ Support to the project, 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 and the private sector where appropriate	
		○ Suitable procurement procedures for timely project implementation	
		○ Engagement with UNIDO in policy dialogue to promote the up-scaling or replication of innovations	
3	Donor	✓ Timely disbursement of project funds	Yes
		✓ Feedback to progress reports, including Mid-Term Evaluation, if applicable	
		✓ Support by the donor's country presence (if applicable) supporting the project for example through engagement in policy dialogue	
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

Annex 3: Endnotes

- i UNIDO, 2015: Director General's Bulletin: Evaluation Policy. UNIDO/DGB/(M).98/Rev.1.
- ii UNIDO, 2016: Industrial Development Board. Forty-fourth session. Vienna, 22 – 24 November 2016. Agenda item 4. Report of the Programme and Budget Committee. IDB.44/8.
- iii IDB.44/Dec.12.
- iv OECD Development Assistance Committee, 1991: Principles for evaluation of development assistance.
- v United Nations Evaluation Group, 2016: Norms and standards for evaluation.
- vi UNIDO, 2007: Results-based Management and Logframe approach. An introduction to the concept and to the way they are used in UNIDO (Draft).
- vii UNIDO, 2014: Forty-second session of UNIDO's of Industrial Development Board on the implementation of the Lima declaration towards inclusive and sustainable industrial development. (https://www.unido.org/fileadmin/user_media/PMO/IDB/IDB42/idb42_16e.pdf)
- viii UNIDO, 2016: Industrial Development Board. Forty-fourth session. Vienna, 22 – 24 November 2016. Agenda item 4. Report of the Programme and Budget Committee. IDB.44/8.
- ix Sida, 2007: Looking back, moving forward. Sida Evaluation Manual. 2nd revised version.
- x United Nations Evaluation Group, 2010: UNEG checklist for evaluation terms of reference and inception reports. Guidance document.
- xi United Nations Evaluation Group, 2010: UNEG checklist for evaluation reports. Guidance document.
- xii OECD/DAC Working Party on Aid Evaluation, 2002: Evaluation and Aid Effectiveness.
- xiii OECD/DAC, 2002: Glossary of key terms in evaluation and results-based management; IFAD 2009: Evaluation Manual; UN Women (<http://www.unwomen.org/en/how-we-work/un-system-coordination/gender-mainstreaming>); Collins Dictionary; Toold4Dev.org.
- xiv As practiced in the Annual Results Report and Impact of IFAD or synthesis reports of other specialized UN agencies.
- xv *ibid*, page 45, amended
- xvi See IFAD, 2015: Evaluation Manual. Second edition, page 45
- xvii http://intranet.unido.org/intra/Gender_Mainstreaming_Tools_and_Guides
- xviii www.unido.org/en/what-we-do/cross-cutting-issues/gender/publications.html

- xix UNIDO, 2007: Results-based management and logical framework approach. An introduction to the concepts and to the way they are used in UNIDO, Annex 1.
- xx Vogel, I./DFID, 2012: Review of the use of “Theory of Change” in international development. Review report, page 44.
- xxi Ibid., page 43
- xxii 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.
- xxiii IFAD, 2009: Evaluation Manual. Methodology and processes.
- xxiv United Nations Evaluation Group (UNEG), 2010: UNEG Quality Checklist for Evaluation Reports. Guidance Document.
- xxv IFAD, 2015: Evaluation manual, second edition.
- xxvi United Nations Evaluation Group, 2010: UNEG checklist for evaluation terms of reference and inception reports. Guidance document
- xxvii or even more impact evaluations
- xxviii UNIDO, undated: UNIDO evaluation inception report: standard format
- xxix UN Office on Drugs and Crime, 2012: Evaluation handbook
- xxx A good practice stipulated by World Intellectual Property Organization (WIPO), 2010: WIPO Independent Evaluation Guidelines
- xxxi UNEG, 2010: UNEG Quality Checklist for Evaluation Reports
- xxxii Median for country evaluations undertaken in 2015 and 2016
- xxxiii UNIDO Interventions in the Area of Enterprise Development for Job Creation, including for Women and Youth (2015), page 1
- xxxiv Between 2012 and 2016
- xxxv Which was at times omitted in recent thematic evaluation reports: seven most recent thematic evaluation reports commissioned by UNIDO in 2015 and 2014, as published in the UNIDO website (no thematic evaluations were available for 2016 at the time of preparing this Evaluation Manual)
- xxxvi 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).
- xxxvii Global Environment Facility, 2017: Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects.



UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

Further information:
Office of Evaluation and Internal Oversight
Independent Evaluation Division
evaluation@unido.org
<http://www.unido.org/evaluation>