ILO Guidelines to Results-Based Evaluation

Principles, Rationale, Planning and Managing for Evaluations

Version 1 - January 2010

EVALUATION UNIT
(EVAL)

Acronyms

BCPR	Biennial Country Programme Review
PARDEV	ILO Partnerships and Development Cooperation Department
CPE	Country Programme Evaluation
Danida	Danish International Development Assistance
DeGEval	Deutsche GesellschaftfürEvaluierung (*German Evaluation Society*)
DFID	UK's Department for International Development
DWCP	Decent Work Country Programme
ED	Executive Director
FO	Field Office
EVAL	ILO Evaluation Unit
FAO	Food and Agriculture Organization of the United Nations
HQ	Headquarters
IFAD	UN International Fund for Agriculture Development
ILO	International Labour Organization
NGO	Non-governmental Organization
OECD/DAC	Organization for Economic Cooperation and Development/ Development Assistance Committee
PRODOC	Project document
RBTC	Regular Budget Technical Cooperation resources
RBM	Results Based Management
RBSA	Regular Budget Supplementary Account
RD	Regional Director
Sida	Swedish International Development Cooperation Agency
ToR	Terms of Reference
UN	United Nations
UNDAF	United Nations Development Assistance Framework
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNICEF	United Nations Children's Fund
US	United States of America
XBTC	Extra-Budgetary Technical Cooperation resources

Table of Contents

Tables

Figures

Boxes

Annexes

Preamble

1. These ILO evaluation guidelines are written for any ILO staff tasked to plan, manage, oversee and/or follow up on an evaluation. The guidelines are based on the ILO evaluation policy[1] and are meant to elaborate on the Director-General's 2009 Announcement[2] on Evaluation in the ILO (2009) and the Office Directive[3] on the ILO Evaluation Unit (2009). The latest developments in evaluation in international development such as the Norms and Standards of the UN Evaluation Group are accommodated.[4]

2. The ILO evaluation guidelines have been produced as an internal governance document rather than a set of recommended good practices. The guidelines will be stored on the ILO intranet as well as the internet. They consist of two parts: Part 1 entails an introduction to the concepts of evaluation and aims to support the understanding of its added value, as well as indicate how these concepts are to be applied when planning and managing an evaluation. Part 2, presented in a separate document, explains specific evaluation types and topics as they are applied in the ILO.

3. Section 1 of the present document provides an overview of generic evaluation guidelines in the ILO, and aims to clarify basic concepts. This section applies to all evaluation types and approaches used in the ILO and serves as an introduction to explain the added value of evaluation to the organization. Section 2 guides the reader through the generic processes of planning and managing an evaluation. For further reading, references to existing ILO evaluation guidance notes or other background material are provided. A glossary of evaluation terms, and thus of the terms used herein, is available at http://www.oecd.org/dataoecd/29/21/2754804.pdf.

[1] GB.294/PFA/8/4: A new policy and strategic framework for evaluation at the ILO. Geneva, 2005.
[2] Director-General's announcement, IGDS Number 75.
[3] Office Directive, IGDS Number 74.
[4] The evaluation guidelines were prepared by EVAL with the support of the independent consultant Dr Achim Engelhardt, associate of the Lotus M&E Group (Achim@lotus-group.org).

1 Principles and rationale for evaluation in the ILO

1. Evaluation in the ILO is to be used as a management and organizational learning tool with the aim to support ILO constituents in forwarding decent work and social justice. It is a critical means to generate knowledge in the organization and aims to improve programmes and policies by improving decision-making and providing verifiable evidence of effectiveness. Evaluation answers key questions such as “are we doing the right things?”, “are we doing them correctly?”, “what works?”, “what doesn’t work and why?”Being an integral part of results-based management (RBM) in the ILO, evaluation strengthens the coherence between results, impact and resource allocation in the programming process. Ultimately, evaluation is only useful if it guides decision-making processes and is used for managing the ILO’s work.

1.1 Principles of ILO evaluation

2. The ILO adheres to the principles of evaluation in the UN system: usefulness, impartiality, independence, quality, competence, and transparency and consultation as listed in Box 1.

Box 1 Principles of evaluation in the UN system

Usefulness: The selection, design and follow-up of evaluations aim for usefulness, particularly to support decision-making

Impartiality: Evaluation processes are established to minimize bias and protect impartiality at all stages of the evaluation, thereby supporting the credibility of the evaluation function and evaluation results. Reports must present in a complete and balanced way the evidence, findings, conclusions and recommendations.

Independence: There will be a clear separation of evaluation responsibility from line management functions. Evaluators are selected with due regard to their independence and professionalism to avoid potential conflicts of interest.

Quality: Each evaluation should employ design, planning and implementation processes that are inherently quality oriented, covering appropriate methodologies for data collection, analysis and interpretation.

Competence: Those engaged in designing, conducting and managing evaluation activities shall have all necessary skills to conduct high-quality and ethical work as defined in the UN Evaluation Group’s professional standards.

Transparency and consultation: Transparency and consultation with the major stakeholders are essential features in all stages of the evaluation process. This improves the credibility and quality of the evaluation. It can facilitate consensus building and ownership of the findings, conclusions and recommendations.

3. The evaluation function in the ILO is designed to ensure transparency and independence of evaluations, which in turn reinforces their credibility and usefulness. Line managers are called upon by the Director-General to safeguard the integrity of the evaluation process by ensuring adherence to the ILO’s evaluation policies and guidelines.

1.2 Definition of evaluation

4. In plain English, evaluation is essentially a reality test, a process for determining the significance or worth of something. A more formal ILO definition of evaluation as a learning and accountability tool is the UNEG definition provided in Box 2. This definition is also used in the ILO's evaluation policy.[5]

Box 2 Evaluation: A definition

An evaluation is an assessment, as systematic and impartial as possible, of an activity, project, program, strategy, policy, topic, theme, sector, operational area, institutional performance, etc. It focuses on expected and achieved accomplishments, examining the results chain, processes, contextual factors and causality, in order to understand achievements or the lack thereof.
It aims at determining the relevance, impact, effectiveness, efficiency and sustainability of the interventions and contributions of the organizations of the UN system. An evaluation should provide evidence-based information that is credible, reliable and useful, enabling the timely incorporation of findings, recommendations and lessons into the decision-making processes of the organizations of the UN system and its members.

Source: UN Evaluation Group: Norms for Evaluation in the UN System, April 2005.

5. Evaluation focuses on achievement of development results and, in this context, considerations of design, implementation and management processes. It should not be confused with implementation monitoring and reporting, audit, inspection, investigation or assessment of individual performance. Audits assess internal practices with regard to effectiveness of risk management, control and governance processes. Investigations respond to complaints related to conduct or security.

6. Although it takes the form of data-based analysis, evaluation is not academic research. The ILO does support assessment of impact or progress towards impact, as a core part of its evaluation methodologies. This may involve some application of research methods.

1.3 Purpose of evaluation

7. The aim of evaluation is to promote both accountability and learning. These two purposes often appear in opposition since participation and dialogue are necessary for learning, whilst independent, objective, impartial and independent evaluation is usually considered a precondition for accountability, as shown in Box 3.

Box 3 Definition of accountability

Obligation to demonstrate that work has been conducted in compliance with agreed rules and standards or to report fairly and accurately on performance results vis a vis mandated roles and/or plans. This may require a careful, even legally defensible, demonstration that the work is consistent with the contract terms.

Source: OECD/DAC, 2002: Glossary of key terms in evaluation and results based management.

8. UN agencies agreed in their Norms for the UN System (2005) that the "purposes of evaluation include understanding why and the extent to which intended and unintended [positive and negative] results are achieved, as well as their impact on stakeholders.

[5] GB.294/PFA/8/4: A new policy and strategic framework for evaluation at the ILO. Geneva, 2005, p. 2.

Evaluation is an important source of evidence about the achievement of results and institutional performance. Evaluation is also an important contributor to building knowledge and to organizational learning. Evaluation is an important agent of change and plays a critical and credible role in supporting accountability". The ILO is using the definition of lessons learned presented in Box 4.

Box 4 Definition of lessons learned

"A lesson learned is a knowledge [artefact] or understanding gained by experience. The experience may be positive, as in a successful test or mission, or negative, as in a mishap or failure. Successes are also considered sources of lessons learned. A lesson must be significant in that it has a real or assumed impact on operations; valid in that it is factually and technically correct; and applicable in that it identifies a specific design, process, or decision that reduces or eliminates the potential for failures and mishaps, or reinforces a positive result."

Source: ILO guidance on evaluation lessons learned, page 2.

9. With respect to **accountability and learning** in the ILO, evaluation aims to ensure the following:

Accountability

- To inform managerial decisions involving line management and constituents;
- To inform future planning in a particular sector or country context;
- To inform policy in the ILO or amongst partner country governments in line with the Paris Declaration on donor harmonization and the UN reform's focus on transparency.[6]

Learning

- To learn lessons from the work evaluated that can be applied elsewhere;
- For implementing partners to learn during the evaluation process.

1.4 Evaluation in a changing development environment

10. The importance of evaluation is highlighted in the context of the UN reform process. The Secretary-General's High Level Panel recommended in its report "Delivering as One" a common UN evaluation system to promote transparency and accountability.[7]

11. The Declarations of Monterrey 2002[8], Rome 2003[9], Marrakech 2003[10] and the 2005 Paris Declaration on aid effectiveness show a growing trend to integrate all external development support into the national development processes. Consequently, the UN is increasingly challenged to document its role and contribution within the broader

[6] The Paris Declaration sets the agenda for far-reaching and monitorable actions to reform the ways of delivering and managing aid. The signatory thereby commit to fostering "ownership, harmonization, alignment, results and mutual accountability," i.e. among other things, to strengthening partner countries' national development strategies and operational frameworks, increasing their capacities and in particular to addressing weaknesses in partner countries' institutional capacities to develop and implement results-driven development strategies.
[7] United Nations: Delivering as One. Report of the General Secretary's High-level panel, 2006, New York.
[8] http://www.un.org/esa/ffd/monterrey/MonterreyConsensus.pdf.
[9] http://www.oecd.org/dataoecd/54/50/31451637.pdf.
[10] http://www.g77.org/doc/docs/Marrakech%20Final%20Docs%20(E).pdf.

context.[11] Joint UN evaluations and UNDAF evaluations, along with other initiatives focusing on coherent, comprehensive, multipartner programmes or frameworks, play an important role in this process.

12. Evaluations with UNDAF partners, i.e. the UNCT, the government of the programme country and donors who support the programmes, are growing in importance, as is demand for evaluation of DWCPs and projects by individual agencies such as the ILO. In this context, UN interventions, including those supported by the ILO, are also to be evaluated as to how they support the achievement of the MDG-related national priorities through advocacy, capacity development and programming.[12]

1.5 Evaluation in the ILO's results-based management framework

13. Evaluation and results-based management (RBM) are interlinked. The ILO's evaluation policy raises the expectation that "evaluation will strengthen the coherence between results, impact and resource allocation in the programming process[13]" and that it guides decision-making processes.

14. RBM tends to be used to assess "whether" results were achieved while evaluation is used to enquire about "why" and "how" results were achieved.

15. According to the ILO's recent RBM guidebook the evaluation process provides a distinct, essential and complementary function to performance measurement and RBM. The evaluation function provides information not readily available from performance monitoring systems, in particular in-depth consideration of attribution, relevance, effectiveness and sustainability. Evaluation also brings to the performance system elements of independence of judgment. It addresses why results were or were not achieved and provides recommendations for appropriate management action. For these reasons evaluation is an essential component of RBM.

1.6 Types of evaluation in the ILO

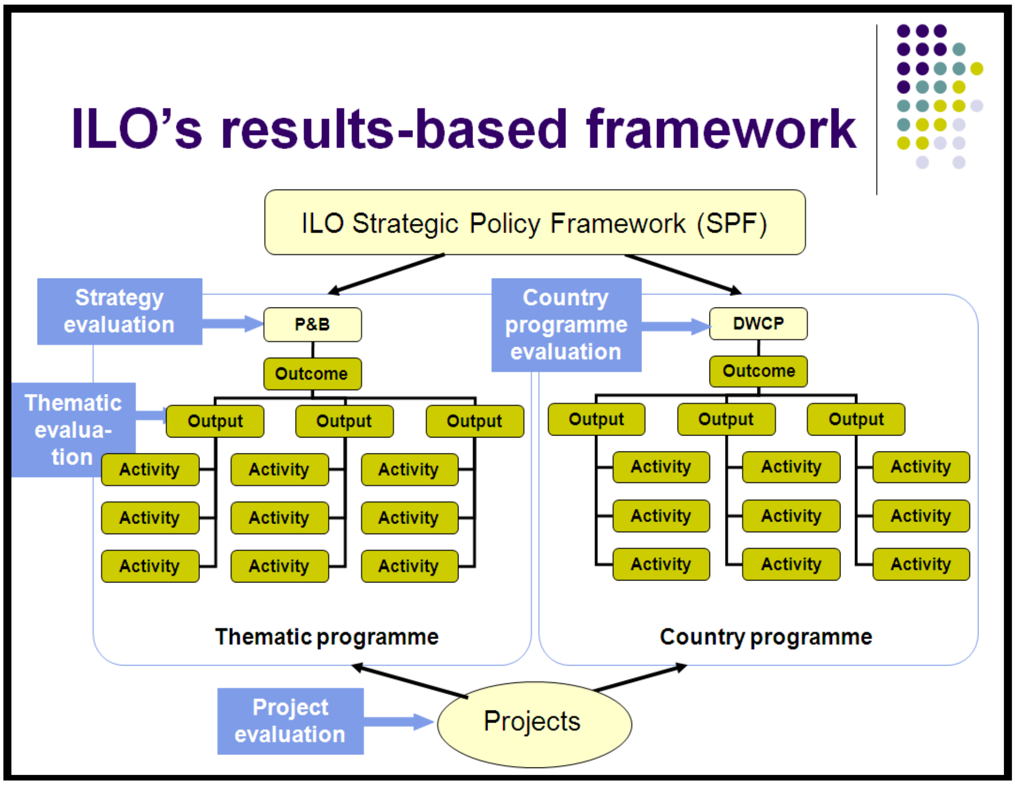
16. According to the ILO evaluation policy, the organization benefits from five evaluation types:

- Strategy and policy evaluation
- Evaluation of Decent Work Country Programmes
- Thematic evaluation
- Project evaluation
- Organizational review (self-evaluation)

These evaluations align with different levels of programming within the Office and are designed to generate knowledge and inform decisions at different programming levels (see figure 1 below).

[11] UN: UNDAF evaluation – guidelines for Terms of Reference, April 2005. http://cfapp1-docs-public.undp.org/eo/evaldocs1/workshop/uneg/UNDAF%20Evaluation%20Final%20draft%20for%20PG%20-%20feb.1%2005.doc.
[12] UN: Common Country Assessment and UNDAF guidelines, February 2007.
[13] GB.294/PFA/8/4: A new policy and strategic framework for evaluation at the ILO. Geneva, 2005, p. 5.

Figure 1 The ILO results-based framework



17. Table 1 provides an overview of the evaluation types used in the ILO. The five evaluation types are explained in more detail in Part 2 of the ILO evaluation guidelines.

18. Strategy, policy and DWCP evaluations are high-level evaluations, managed by the ILO Evaluation Unit; all other types (thematic, self-evaluations and project evaluations) are decentralized.

Table 1 Overview of evaluation types

		Type of evaluation	Main purpose	Responsibility	Timing
High-level evaluation	Independent	Strategy, policy	- Review major policies or institutional issues. - Assess impact, effectiveness and benefits of ILO core strategies as described in P&B. - Improve strategies and policies, and the functioning of the Office .	- EVAL to plan and manage. - Governing Body confirming topics. - Evaluation Advisory Committee reviewing follow-up.	At least one a year; additionally as mandated and resourced.

High-level evaluation	Independent	Decent Work Country Programme	- Assess the extent to which significant impacts are being made towards decent work through overall country-level activities. - Feed into country tripartite dialogue on impact, effectiveness and relevance of ILO action at the country level.	- Planned and managed by Evaluation Unit. - Implemented in coordination with the regions regional evaluation officers. - Regional Offices responsible for financing these evaluations.	At least one each year. All regions to be covered over a four-year period.
Decentralized	Independent/internal	Thematic evaluation	- Assess effectiveness and impact of specific means of actions and interventions. - Develop cross-cutting lessons, including success stories to innovate and feed into organizational learning on technical interventions and strategies.	- Technical sectors and other technical groups to plan and manage. - Evaluation Unit to coordinate and support as requested. - Technical programmes to resource.	At least one every year; sectors to submit annually summaries of thematic evaluation work planned and completed.
Decentralized	Internal	Organizational review/self-evaluation	- Assess priority, relevance of the programme activities in relation to actual performance against planned outcomes. - Self-assess achievement and results aimed at improving effectiveness and efficiency. - Get timely information and management decision in achieving planned outcomes against target and indicators. - Applies also to Country Programme Reviews.	- Line managers to ensure compliance with ILO policies. - Organizational group leaders to conduct. - Executive and regional directors(ED/RD) responsible for ensuring appropriate resources are available.	Self-evaluation is biennial. All subject to internal review, scopes aligning with P&B and DWCP outcomes.

Decentralized	Independent for budgets aboveUS$500,000; otherwise internal	Project (independent for budgets above US$500,000)	- Assess projects for relevance, efficiency, effectiveness, sustainability and contribution to broader impact. - Appropriateness of design to ILO's strategic and national decent work programme frameworks.	- Planning of evaluation is the responsibility of the person to whom the project manager reports but management of evaluation supervised by regional or sector-level evaluation officers. - EDs and RDs responsible for ensuring their line managers apply ILO policy. - PARDEV to support communication on evaluation as part of donor liaison responsibilities management good practice. - Evaluation Unit provides oversight of quality and compliance. - Cost of evaluation to be included in project budget.	Mid-term or final or as set in the project evaluation plan.

19. The following are common to all evaluation types: the principles of orientation by national and ILO longer term priorities and objectives; the focus on decent work, results-based, on the assumption that all interventions in the ILO are managed for these results; and the principle of reinforcing a sense of joint ownership by the stakeholders of ILO interventions.

1.7 Key roles and responsibilities

20. As illustrated in table 1 and previously mentioned, ILO evaluations can be either decentralized or high-level ones. High-level evaluations are mandated by the ILO's Governing Body or the Director-General. If independent, they are planned and managed directly by the ILO Evaluation Unit. The evaluation summary report and subsequent status reports on implementation of the recommendations are presented to the ILO's Governing Body. Follow-up to high-level evaluations is reviewed by the ILO Evaluation Advisory Committee[14], who in turn report to the Director-General on adequacy of follow-up.

21. Decentralized evaluations focus on programmatic areas more directly under the control of ILO management, such as technical cooperation projects and implementation of country programmes, including interventions funded from all sources of funds. The planning and management of these evaluations are primarily with the regions and sectors assigned responsibility for their implementation. Specific roles and responsibilities are listed in Table 2.

[14] See Circular number 245.

Table 2 Key roles and responsibilities in the evaluation process in decentralized evaluation

Actors	Key roles and responsibilities
Evaluation Manager	- Managing all independent and internal evaluations - From sector or region in which the intervention is implemented - No links to decision-making of the intervention to be evaluated -Finalizes evaluation terms of reference and submits for approval -Selects consultant and negotiates contractual terms, and submits consultant choice for approval.
Evaluation Focal Person/ Regional Evaluation Officer	-Based in Regional Office or technical sector - Approves final version of theToR and selection of evaluation consultants - Eventually carries out evaluability study and/or scoping mission - Up-loads evaluation process documents into i-track
Project/Programme Manager and staff	- Facilitation and support of the evaluation (e.g. provides information, assists in data gathering and coordinates exchanges of evaluation team with partners) -Provides inputs to ToR - In the case of mid-term evaluations, elaborates follow-up plans
ILO responsible official and intervention backstopppers	- Ensure that sufficient funds are secured for evaluations at the intervention design stage - Administrative and technical support for the evaluation, including help in preparing ToR - Follow-up of the evaluation and dissemination of lessons
Regional Directors and Executive Directors	- Ensure that principles supporting the evaluation function are applied for all evaluations falling within their domain
Lead Evaluator	- Is always an external person for all independent evaluations but may be aided by an independent ILO evaluator - Undertakes the evaluation within the agreed ToR for the evaluation - Evaluators shall be independent and have sole responsibility for the content of the final evaluation report - Prepares and submits draft and final evaluation reports
EVAL	- Assures quality in line with international standards and monitors compliance with the ILO evaluation policy - Provides standards and guidance on procedures - Approves the final evaluation report prior to submission to donor - Stores and publishes ToRs and evaluation reports

Further reading about specific roles and responsibilities in the project evaluation process is provided in the EVAL guidance note "Concept and policies of project evaluation" available on the intranet.

1.8 Evaluation modes in the ILO

22. The ILO uses a series of evaluation modes for the abovementioned evaluation types. Timing and the degree of independence define those modes.

23. Modes defined by timing are annual reviews, mid-term and final evaluation and (final or ex-post) impact assessments. Modes defined by actors are self-evaluation, internal

evaluation, independent evaluation and external evaluation. Evaluation modes defined by time are further explained below:

- **Annual reviews** focus on outputs and outcomes of projects, programmes, strategies or policies. They are a form of self-evaluation during which the stakeholders reflect upon how well the intervention is progressing towards achieving its objectives taking into account available monitoring and evaluation data. Reviews with this type of focus may also be organized to look at specific issues. ILO managers and CTAs are responsible for annually reviewing and reporting their progress. PROGRAM and PARDEV oversee the processes at an organizational level and report performance to the Governing Body;
- **Mid-term and final evaluations** focus on the outcomes of projects, programmes, strategies or policies and the likelihood that they will achieve impact. Evaluations provide an opportunity for in-depth reflection on the strategy and assumptions guiding the intervention. They assess progress made towards the achievement of the intervention's objectives and may recommend adjustments to its strategy. They are also a means by which to assess how well intervention-level actions link to and support higher level ILO strategies and objectives, as articulated in Decent Work Country Programmes (DWCPs) and the ILO's Programme and Budget. The Evaluation Unit has oversight responsibility for all independent mid-term and final evaluations not considered to be at a high level. Regional and sector-level evaluation officers have responsibility for hands-on supervision;
- **Impact assessments** determine whether interventions have contributed to longer term impact. They can be part of ex-post evaluations of projects or they can be part of thematic or country programme evaluations that also consider linkages between different interventions and longer term development outcomes.

24. The following rules apply for project evaluations (XBTC, RBTC and RBSA):

Table 3 Rules for project evaluation in the ILO

BUDGET-BASED REQUIREMENTS FOR INDEPENDENT PROJECT EVALUATIONS
• All projects with budget >US$500,000 must undergo at least one *independent* evaluation; • Multi-phase projects with combined budgets > US$500,000 must undergo at least one *independent* evaluation; • For projects with budget < US$500,000 evaluations are not required to be *independent*.

SCHEDULE FOR PROJECT-LEVEL EVALUATIONS BASED ON PROJECT DURATION
• For projects duration< 18 months, a final evaluation is required; • For projects duration> 18 but < 30 months, annual reviews/progress reports and a final evaluation are required; • For projects duration>30 months, annual reviews, a mid-term evaluation and a final evaluation are required. • A single evaluation may be conducted to cover several projects, which are clustered by theme or geographic focus, provided that the evaluation of multiple projects 1) applies a scope, purpose, and methodologies comparable to what would be used for an individual evaluation, 2) has donor consent, and 3) is approved by EVAL or regional evaluation officers for projects with budget > US$500,000.

25. Evaluation modes defined by the degree of independence are further explained below and summarized in Table 4:

- **Self-evaluation** – managed and conducted by ILO staff members, including project management, technical specialists and backstoppers;
- **Internal evaluation** – managed and conducted by independent ILO officials, i.e. staff members who have not been involved in the design, management or backstopping of the project that they are evaluating (e.g. the regional or sectoral evaluation focal person);
- **Independent evaluation** – managed by independent ILO officials. Conducted and led by external evaluators who have no previous links to the project. Other independent ILO officials may participate as team members in the evaluation;
- **External evaluation** – managed from outside the ILO and conducted by external evaluators who have no previous links to the project being evaluated. External evaluations are usually initiated, led and financed by a donor agency.

Table 4 ILO evaluation approaches by actor (for decentralized evaluations only)

	Management	Evaluators	Degree of Impartiality	Costs to the ILO
Self-evaluation	ILO (incl. project management)	ILO (incl. project management)	Low	Low
Internal evaluation	ILO (excl. project management)	ILO (excl. project management)	Medium	Medium
Independent evaluation	ILO (excl. project management)	External (leadership) possibly plus ILO (excl. project management)	Medium to high	High
External evaluation	External	External	High	High

1.9 ILO regional and sector evaluation networks

26. Regional evaluation networks support the implementation of evaluation activities, particularly with regard to decentralized technical cooperation projects. A regional network will comprise a designated evaluation officer at the Regional Office and evaluation coordinators in sub-regional Offices. EVAL will provide technical guidance and assistance to the network in undertaking their functions.

27. The regional evaluation officer is the regional focal person on internal country programme reviews and TC project evaluations. The regional evaluation officer oversees and advises the process of planning, managing, and following up of DWCP reviews and project evaluations.

28. Sector level evaluation officers and focal persons fulfill the same role and functions (i.e. oversee and advise the process of planning, managing and following up) for thematic reviews and evaluations of centralized projects.

2 Planning and managing an evaluation

29. This section aims to explain the specific steps and approaches required for planning and managing an independent evaluation. Planning an evaluation, in particular, comprises seven key steps ranging from defining the purpose of the evaluation to the recruitment of the evaluation team.

30. Guidelines are also provided for undertaking the evaluation (including data collection, data analysis) and reporting requirements. This section closes with highlighting the procedures for using evaluation results in future work, and reporting on follow-up action taken.

31. In addition to guidelines, EVAL uses the i-track system to support work flow for planning and managing evaluations as shown in Box 5.

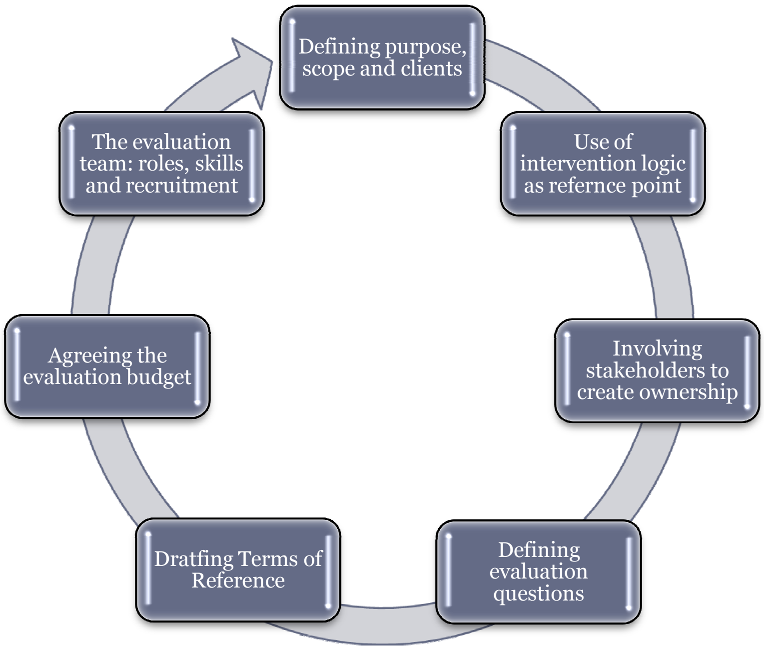
Box 5 Eval-track: information for planning and managing evaluations

What information does Eval-track contain for planning?	**What information does Eval-track contain for managing evaluations?**
• Related projects and budgets • Evaluation type, timing and nature • Responsible ILO office • Evaluation Manager • Planned date of evaluation • Amount of funds	Evaluation details page contains related information including: Suggested method, consultant and-- Workflow sequence • Final ToR: attached when completed Start date of evaluation • Draft report: attached when completed • Consolidated comments: attached when completed • Final report approved by REO: attached when completed • Final report approved by EVAL: attached when completed Report sent to stakeholders • Management response and follow-up plan: attached when completed • Executive summary: attached when completed • Follow-up report: 6 months, 12 months • Appraisal report

2.1 Evaluation planning

32. For the planning of each evaluation, a number of key steps need to be followed, as summarized in Figure 1. These steps are followed to ensure that the evaluation design process is participatory, transparent and independent of any one stakeholder's specific interests. These initial planning steps also help to establish the credibility and usefulness of the evaluation exercise, and therefore should not be greatly modified or skipped.

Figure 1 Key steps in evaluation planning



2.1.1 Step 1: Defining the purpose, scope and clients of an evaluation

33. Initially, the stakeholders should determine the objectives, coverage and key clients of the evaluation. The evaluation manager joins in once the initial plan is agreed. Consultation with the key stakeholders to determine the scope of the evaluation is a good way of identifying some key parameters for the evaluation, and awakening interest in the evaluation's findings. The scope can be defined in terms of time and space (project start/end and geographic areas of implementation) or by project phase or elements of a project. This consultation process helps the evaluation manager to accommodate the key stakeholders' priorities when drafting the ToR and avoids major revision of the draft after circulation. When determining the purpose and scope the evaluation manager should also keep in mind that the evaluation itself should be effective in delivering its purpose and efficient in its use of time and financial resources.

2.1.2 Step 2: Applying the intervention logic as a reference point for evaluation

34. For projects the reference point for each evaluation is the project document (PRODOC) which provides the background and rationale of the project, its planned objectives, outcomes, outputs and activities, and corresponding outcome indicators and assumptions. Good project design and a well-written project or programme document represent a strong foundation for an evaluation.

35. Programmes and strategies are also based on an intervention logic, a theory of change often expressed in logic models and at times in logframes for projects and programmes, and articulated in the P&B and SPF for themes, sectors, regions and the overall organization.

36. Evaluation assesses the relevance of the intervention objectives and approach, establishes how far the intervention has achieved its planned outcomes and objectives, the extent to which its strategy has proven efficient and effective, and whether it is likely to have a sustainable impact. Evaluation tests the propositions upon which the intervention logic is based to see if it is valid. Country programme evaluations are based on DWCP documents agreed with constituents.

37. Project, programme, theme, strategy or policy documents are not all of the same quality, nor do they always reflect the real situation at implementation start-up. In some cases, the strategy and design of the project is modified during the course of implementation to adapt to changing conditions.

38. The evaluator needs to get a clear understanding of the intervention design, logic and strategy. Therefore, prior to writing the ToR, the evaluation manager, in close consultation with the project/programme manager and other key stakeholders, should determine if the project or programme document continues to accurately describe the situation or if not, why it has changed.

39. Implementation planning, progress reports and project/programme/strategy revision documents are key sources of information on modifications to the original design during the implementation process. The project/programme manager should update the documentation on the intervention logic and add supplementary documentation and explanation. The evaluation manager draws on this information to prepare the ToR. This documentation is later submitted to the evaluator who should consider the appropriateness of any changes in the strategy.

For further reading about questions for reviewing the intervention logic, refer to the ILO Technical Cooperation Manual and the evaluation guidance note "M&E planning for projects" available on the intranet.

2.1.2.1 Assessing feasibility of a credible evaluation (evaluability/scoping)

40. The process of reviewing the intervention logic helps to decide whether it is actually possible to evaluate the project, programme, theme, strategy or policy. This process should be conducted during the planning stage and may require an additional scoping mission or scoping activities. Scoping is defined as "analyzing alternative ways for conducting an evaluation. It clarifies the validity of issues, the complexity of the assignment, the users of

final reports, and the selection of team members to meet the needs of an evaluation. Scoping ends when a major go/no-go decision is made about whether to do the evaluation”.[15]

41. Any evaluability exercise would normally be undertaken in coordination with the regional evaluation officer or an evaluation expert who is capable of conducting a technical assessment of the basic parameters for a comprehensive evaluation. The final report should address the following factors:

- clarity of intent of the project or programme to be evaluated (relevance and design of the expected outcome statements and result matrices)
- quality of the design for the achievement of results (e.g. the existence of clear and measurable indicators and eventually baselines and milestones for reliable analysis)
- overall quality of monitoring systems
- initial appraisal of processes for optimal involvement of the relevant national and international stakeholders
- potential degree of national ownership and leadership in the evaluation process; and
- external factors that have influenced or would be expected to influence the realization of the expected outcomes.

42. A selection of evaluability questions is listed in Box 6.

Box 6	**Sample evaluability questions**

- Does the project or programme have a documented results framework by which to assess progress against outcomes?
- Has the project implemented against this results-based plan?
- Is baseline data available to track change processes?
- Have monitoring reports been produced and stored?
- Are key informants available for interviews?
- Is the timing of the evaluation fitting within the intervention cycle?
- Does the security situation in country allow for field missions?
- Is the political context in country conducive to the consultation/participation of key stakeholders?
- Is the budget available for the scale of evaluation envisaged (resource adequacy)?

43. Lack of information and negative answers to most of the evaluability questions are not necessarily grounds for canceling the evaluation, but can help narrow down the set of key evaluation questions, and chose methodologies and timing.

2.1.3 Step 3: Involving stakeholders to facilitate ownership through a consultative process

44. Once the scoping questions are answered, it should be possible to agree with the relevant stakeholders on the main focus and the key questions that the evaluation should address. These evaluation questions are of significant importance to check the assumptions of the intervention logic of projects or programmes, and to assess risks taken.

45. The ILO’s primary stakeholders are the tripartite constituents, which constitute its organizational membership. If the key stakeholders feel involved in getting answers to the

[15] US Environmental Protection agency’s Programme evaluation glossary: www.epa.gov/evaluate/glossary/s-esd.htm.

questions they are interested in, they are more likely to be involved in implementing the recommendations. Relevant stakeholders apart from the ILO's constituents are likely to include relevant HQ and field staff as well as partners in country, such as UN officials from partnering agencies, government officials in collaborating ministries, implementing agencies and representatives of other donors as listed in Box 7. The specific primary stakeholders to involve in an evaluation will be determined by their involvement in design and delivery of the programme, strategy or project being evaluated.

Box 7 Evaluation stakeholders of the ILO

Primary stakeholders

- Representatives of governments (e.g. ministries of labour)
- Representatives of employers' organizations
- Representatives of workers' organizations

Other key stakeholders (for projects and DWCP)

- ILO HQ staff of cooperating departments
- ILO field staff
- UN agencies in country
- NGOs
- Other partners in country (for example donor agencies)

46. Participation is one of the ILO's guiding principles through its tripartite approach and one of its comparative strengths. The core stakeholders to an evaluation should participate as early as the planning stage to create a common understanding about the purpose and use of the evaluation and the approach to be taken. Stakeholders also participate as key informants, being interviewed individually or in groups at workshops or consulted through questionnaires. Maximizing participation in the planning helps to ensure that the focus and methodology are appropriate and that interest in the results has been aroused. Maximizing participation in the data collection phase should ensure that the evaluation team registers all points of view.

2.1.4 Step 4: Defining evaluation questions and measurement standards

47. In line with international good practices for evaluations, the ILO expects that each evaluation will assess the **key evaluation criteria** explained in Table below.

Table 5 Definitions of key evaluation criteria

Evaluation criteria	Description
Relevance and strategic fit of the intervention	The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities and partners' and donors' policies. The extent to which the approach is strategic and the ILO uses its comparative advantage.
Validity of intervention design	The extent to which the design is logical and coherent.
Intervention progress and effectiveness	The extent to which the intervention's immediate objectives were achieved, or are expected to be achieved, taking into

	account their relative importance.
Efficiency of resource use	A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results.
Effectiveness of management arrangements	The extent to which management capacities and arrangements put in place support the achievement of results.
Impact orientation and sustainability of the intervention	The strategic orientation of the project towards making a significant contribution to broader, long-term, sustainable development changes. The likelihood that the results of the intervention are durable and can be maintained or even scaled up and replicated by intervention partners after major assistance has been completed.

48. These terms are a reconfiguration of the internationally agreed standard performance criteria – relevance, effectiveness, efficiency, impact and sustainability – into more operational categories.

49. For each criterion the evaluation manager should include two or three specific evaluation questions to guide the evaluator(s) and the evaluation process on key aspects and issues to consider. While evaluation criteria are fairly standard, evaluation questions should be tailored to the specifics of the project with the answers to these leading to recommendations for guiding key decisions on next steps.

50. The evaluation manager, who should take the priorities of the main stakeholders into consideration, develops the analytical framework in advance. This is part of the ToR which the evaluation manager circulates as a draft for comment to all the main stakeholders of the evaluation.

Further reading about the model analytical framework with a detailed list of typical evaluation questions of a mid-term project evaluation is provided in the ILO Technical Cooperation Manual.

2.1.5 Terms of Reference

51. The Terms of Reference (ToR) document forms the contractual basis for undertaking an evaluation. Writing the ToR with sufficient clarity and in sufficient detail will improve the basis for joint understanding with the evaluators about what they are expected to deliver. Well-considered and well-written ToR are the foundation of a good evaluation. ToR specify the reason for the evaluation and summarize the different stakeholders' expectations of the evaluation. They describe the project or programme to be evaluated and its context. The content of the ToR should follow the outline indicated in Box 8 and include the cover page.

Box 8 Outline of Terms of Reference

1. Introduction and rationale for evaluation
2. Brief background on project and context
3. Purpose, scope and clients of evaluation
4. Suggested analytical framework
5. Main outputs
6. Methodology to be followed
7. Management arrangements, work plan and time frame

Further reading about the structure and content of ToR is provided in Annex 2: "A ToR quality checklist".

2.1.5.1 Procedure for drafting and circulating ToR

52. The evaluation manager works with the project manager and ILO officials to prepare the first draft of the evaluation ToR. In the drafting process, the evaluation manager consults with, and receives inputs from, the project manager, the ILO Office Director and the line manager of the technical unit backstopping the project.

53. The evaluation manager then circulates the draft ToR to the following key stakeholders who provide comments within a specified time span:

- Project or programme manager and key staff,
- Main national partners,
- ILO Field Office Director,
- Technical backstopper at headquarters,
- Field technical specialist,
- Responsible evaluation focal point, and
- Donor, if required.
- Constituents as appropriate

54. The evaluation manager integrates the comments into the draft ToR, as appropriate, and passes the ToR to the responsible evaluation focal point for approval. Copies of the final ToR are then sent to the same group of stakeholders who provided comments on the draft.

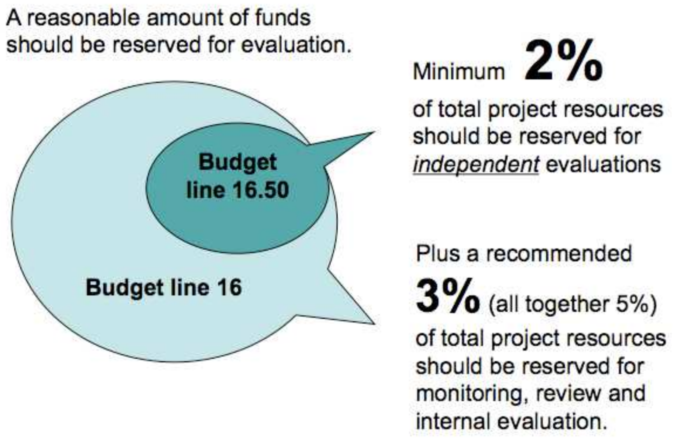
2.1.6 Evaluation budget

55. Most evaluations in the ILO are financed from programme or project budgets[16]. As per ILO evaluation policy, a minimum of 2 per cent of the total project funds should be reserved for independent evaluations, which should be assigned to budget line 16.50 for projects. In contrast to other budget lines, use of the resources under budget line 16.50 requires the approval of the evaluation manager and an ILO evaluation official and not the project manager. All contracts for hiring evaluators should be approved by an ILO evaluation officer prior to finalization. Figure 2 provides an overview of financing for project evaluations in the ILO.

56. The project budget line 16 should reserve adequate resources to cover monitoring and evaluation activities for all the phases of the project or programme e.g. gathering of baseline data and development of monitoring and evaluation plan, end of phase evaluations, and end of programme evaluation to assess end results and impact. Although self-evaluations may not require extra staff costs, they also need to be scheduled and budgeted for as they may involve additional travel costs or workshop costs for consulting partners.

[16] The regions, through extra budgetary and/or regular budget funds, should absorb the costs for the evaluation of DWCPs.

Figure 2 Financing for project evaluations



57. In order to calculate the evaluation budget for financing review missions, these five steps should be followed:

- Step 1: Calculation of number of consultants working days
- Step 2: Determination of consultants level (P3, P4, P5 or professional equivalency)
- Step 3: Calculation of travel costs, including travel days, vehicle use for field trips…
- Step 4: Calculation of accommodation and DSA costs
- Step 5: Calculation of any other costs (interpretation services, workshop facilities for focus group and stakeholder meetings, etc.)

58. Delays in the evaluation process can add to the overall evaluation costs. Therefore it is recommended to plan the timeframe of the evaluation realistically, taking note of national holidays when key offices may be closed. The evaluation manager should take into account the time needed to complete the tasks specified in the terms of reference, including travel days, time for stakeholders to comment on the draft report and for the consultants to accommodate these comments, and where necessary a final conference for clearance of the final report by EVAL.

59. The ILO responsible officials should ensure that an adequate budget exists to implement the evaluation activity while also financing the broader monitoring and evaluation plan, and that it is indicated in the original budget proposal.

2.1.7 Evaluation teams for independent evaluations: roles and skills

60. Experienced external technical experts and professional evaluators can be found qualified to undertake independent evaluations if they are deemed independent of the project or programme being evaluated, or any other conflict of interest. External consultants hired for technical knowledge but with limited evaluation expertise should undergo an orientation on ILO's guidelines and quality standards for evaluation.

61. Both external and internal evaluators should adhere to the highest technical and ethical standards. They should fulfil the criteria of professionalism, impartiality and credibility. If an evaluation team is composed, it is recommended that a local expert be part of the team. The Office also supports gender balance in evaluation consultancies.

62. For an evaluator to be considered independent the following requirements need to be fulfilled:

- No current or previous involvement with the ILO in the project or programme being evaluated;
- No plans to work for the project or organization in the near future.

63. One of the first steps in selecting an evaluator is to think through the evaluator's role. Is the evaluator's role narrowly defined to judging the merit or worth of a project or programme or is he/she expected to act as a facilitator and idea giver and provide options as to the way forward?

64. It is also crucial for the evaluation manager together with project staff and stakeholders to determine the purpose and scope of the evaluation, the questions to be addressed as well as the degree of participation of stakeholders and identified challenges requiring special evaluation skills. The methodology (design and data collection methods) should be agreed and the evaluator's familiarity with the intended methodology should constitute one of the primary criteria in the selection of the consultant.

65. Although the consultant may subsequently adapt the methodology, any changes should be agreed between the evaluation manager and the evaluator. This agreement very often takes the form of an inception report, which is contractually specified and approved by the ILO.

66. Although the lead evaluator of an independent evaluation is an external consultant, the rest of team may be composed of evaluation specialists including members of the Evaluation Unit and/or technical specialists experienced with evaluation but independent of the evaluation focus. The roles and responsibilities of the recruited consultant will ultimately be delineated by the type of ILO evaluation, but may also be significantly influenced by the overall composition of the team.

67. The search for a consultant should be open, with several candidates being placed on a shortlist, and their relative strengths considered against objective criteria. EVAL is in the process of putting together a database of evaluation consultants. One of the most effective, efficient and transparent ways of searching for an evaluator still is posting a public advertisement or call for tenders on the relevant listservs and networks.

68. Calls for tenders should include: who is tendering, what the assignment is about, the starting date and duration of the assignment, the core requirements, the language of the report and a contact email. The evaluation terms of reference may be attached or should be available upon request.

69. The procedure for hiring an external consultant is similar to the procedure for drafting and finalizing the ToR. The evaluation manager is responsible for proposing the external evaluation consultant. She/he can receive suggestions from any side including the project management. Before contracting the consultant the evaluation focal point must approve the choice. If required by the donor, the evaluation manager can send the Curriculum Vitae of the proposed evaluator and a brief explanation of why the consultant was selected to the donor and other key stakeholders. As a last step the finalized ToR and the consultants CVs must be entered in the i-track.

70. Once the consultant has been selected, one must ensure that there is a clear mutual understanding of the expected outputs. In particular, clear deliverables must be referenced on the final version of the ToR. Those deliverables should include:
- an inception report;
- a comprehensive analytical draft report including a set of actionable recommendations and lessons learned regarding how to move forward, followed by a final report; and

- a summary of the report using the ILO template format.

A tentative timeline and explicit deadlines as well as reporting lines should also be contracted.

For a non-exhaustive list of listservs and dedicated networks as well as any further reading about searching and finding an evaluator, refer to "Searching an evaluator through evaluation listservs, networks and databases of evaluators" available on EVAL's intranet.

3 Undertaking the evaluation

71. ILO evaluation managers play a critical role in ensuring that evaluations are carried out credibly, particularly with regard to sound methods. The first step for undertaking an evaluation is to acknowledge or, if required, amend the evaluation methodology proposed in the ToR. For independent evaluations, evaluators are expected to specify the evaluation methodology and/or evaluation instruments to be used in a short inception report prior to conducting the evaluation. The timing and approval of the inception report should be listed as an output in the ToR. Sources and methods for data collection, data analysis and reporting are required. In addition, choice of any site-visits within a country by the evaluator, who is entitled to select locations randomly or based on sound selection criteria, should be specified in the report.

72. Approval by the evaluation manager of the inception report constitutes an acceptance by the ILO of the results generated through the methodology proposed. It is therefore important for the evaluation manager to check the interview lists and guides, questionnaires and sampling, etc. for any aspect that could bias and distort results. Those reviewing should also check that methods draw on both subjective as well as objective sources of data to provide a balanced but insightful report.

3.1 Data collection process: tools and methods

73. For the purpose of data collection, most evaluations use a mix of primary and secondary data. For the collection of primary data, evaluators need to select specific techniques and instruments. Depending on the type of evaluation, data availability, local context, availability of resources and time those techniques are selected to reply to the evaluation questions. Examples of techniques and instruments for primary data collection are: questionnaires, focus group discussions, statistical surveys or case studies.

74. Primary data consists of information evaluators observe or collect directly from stakeholders about their first-hand experience with the intervention. This data is collected through the use of surveys, meetings, focus group discussions, interviews or other methods that involve direct contact with the respondents. It can facilitate deeper understanding of observed changes and the factors that contributed to change.

75. In most cases, collection of data through interviews and focus groups should be carried out in a confidential manner without involvement of those whose work is being evaluated. Evaluators requiring assistance should be provided with the help they need (e.g. non-ILO translators).

76. Secondary data, by contrast, is existing data that has been, or will be, collected by the ILO or others for another purpose. Secondary data can take many forms but usually consists of documentary evidence that has direct relevance for the purposes of the evaluation: nationally and internationally published reports; economic indicators; project or programme plans; monitoring reports; previous reviews, evaluations and other records; country strategic

plans; and research reports. The advantages of using primary or secondary data are described in Box 9below[17].

Box 9	**Advantages of using primary or secondary data**

"The use of secondary data represents tremendous cost and time savings to the evaluation and every effort should be made to establish what secondary data exist (…)
Primary data is often collected unnecessarily and at great expense simply because monitors or evaluators had not been aware that the data were already available. It is critical to invest the initial time and resources to investigate what data exist, what data collection exercises are planned for the future, and how relevant the existing data are (for evaluation).

However, primary data collection is sometimes warranted. Although a review of secondary data sources should precede any primary data collection, existing data do not always provide the appropriate indicators or the appropriate disaggregation of indicators needed (…) to evaluate (…) effectively. Even secondary data that provides the appropriate indicators and disaggregation of indicators may not be useful if the data is out of date and the situation is likely to have changed since they were collected".

Source: WFP, undated: M&E guidelines. Choosing methods and tools for data collection, pp. 5-6.

77. When applying those techniques, consideration of socio-economic characteristics among respondents and sex-disaggregation of data is required. The evaluation manager is encouraged to explore other potential lines of disaggregation such as grouping based on national livelihoods and poverty outcomes.

Further reading about how to treat gender in evaluations is provided in "Considering gender in monitoring and evaluation of projects" available on EVAL's intranet.

3.2 Quantitative and qualitative techniques

78. For decades different schools of thought have argued for and against the use of quantitative vs. qualitative techniques for data collection. Mixing both approaches seems useful where applicable. Quantitative techniques often "force" responses or people into categories that might not "fit" in order to make meaning. Qualitative techniques, on the other hand, sometimes focus too closely on individual results and fail to make connections to larger situations or possible causes of the results.

79. A detailed but non-exhaustive list of data collection tools and methods including a short description and analysis of advantages and disadvantages is presented in Table 6 and Table 7.

80. Most of the presented data collection tools and methods will help generate qualitative data: desk reviews, interviews, including focus group or key informant interviews, on-site observation, memory recall, most significant change technique, expert panels and case studies. Each tool and method has advantages and disadvantages in terms of time, usefulness and resources requirements. The following checklist can facilitate an informed choice for the selection of data collection methods[18]:

- Determine which data-collection methods best answer key evaluation questions.
- Tie method selection to available resources. This may mean revising the

[17] http://www.wfp.org/content/monitoring-and-evalutation-guidelines.
[18] UNFPA, 2004: The Programme Manager planning, monitoring and evaluation toolkit.

evaluation design and methods, or determining other options to stay within budget. It may also mean finding additional resources to fund the most effective and useful evaluation design.

- Choose methods, which will facilitate the participation of key programme stakeholders in the evaluation.
- Strengthen the credibility and usefulness of evaluation results by mixing evaluation methods where appropriate.

Table 6 Data collection tools and methods

	Method	Description	Advantages	Challenges
Qualitative				
	Desk reviews	Systematic analysis of existing documentation, including quantitative and descriptive information about the initiative, its outputs and outcomes, such as documentation from capacity development activities, donor reports, and other evidence.	Cost efficient.	Documentary evidence can be difficult to code and analyze. Difficult to verify reliability and validity of data.
	Interviews (face-to-face, telephone or computer-assisted)	Solicit responses to questions designed to obtain in-depth information about a person's impressions or experiences. Can be fully structured, semi, or unstructured	Facilitates fuller coverage, range and depth of information on a topic.	Can be time consuming. Can be difficult to analyze. Can be costly. Potential for interviewer to bias interviewee's responses.
	Direct On-Site Observation	Entails use of a detailed observation form to record accurate information on-site about how a programme operates (ongoing activities, processes, discussions, social interactions and observable results as directly observed during the course of an initiative).	Can see operations of a programme as they are occurring. Can adapt to events as they occur.	Can be difficult to categorize or interpret observed behaviours. Can be expensive. Subject to (site) selection bias.
	Focus Group Interviews	A small group (6 to 8 people) are interviewed together to explore in-depth stakeholder opinions, similar or divergent points of view, or judgments about a development initiative or policy, as well as gather information about their behaviours, understanding and perceptions of an initiative or to collect information around tangible and non-tangible changes resulting from an initiative.	Quick, reliable way to obtain common impressions from diverse stakeholders. Efficient way to obtain a high degree of range and depth of information in a short time.	Can be hard to analyze responses. Requires trained facilitator. May be difficult to schedule.
	Memory Recall	Entails interviews with beneficiaries and other stakeholders, individually or in groups, who reconstruct their situation before the project.	Useful especially when baseline surveys are not available.	Though often applied, this method is subject to significant bias
	Historical narration/ most significant change technique	Collection of change stories from the field and selection of the most significant of these stories by panels of designated stakeholders or staff. Once changes have been captured, selected people read the stories aloud and have regular in-depth discussions about the value of the reported changes.	Content is likely to be rich in insight	Content may be subjective, especially if the selection of significant changes is done by external agents. Cross-checking of results with other techniques is recommended. Generalization of findings can be an issue.
	Key Informant Interviews	Qualitative in-depth interviews, often one-on-one, with a wide range of stakeholders who have first-hand knowledge about the initiative's operations and context. These experts can provide particular knowledge and understanding of problems	Can provide insight on the nature of problems and give recommendations for solutions. Can provide different perspectives on a single issue or on	Subject to sampling bias. Must have some means to verify or corroborate information.

		and recommend solutions. The majority of questions are open-ended and meant to stimulate discussion	several issues.	

Table 7 Data collection tools and methods (continued)

	Method	Description	Advantages	Challenges
Qualitative	**Expert Panels**	A peer review, or reference group composed of external experts to provide input on technical or other substance topics covered by the evaluation.	Adds credibility. Can serve as additional (expert) source of information that can provide greater depth. Can verify or substantiate information and results in topic area.	Cost of consultancy and related expenses if any. Must ensure impartiality and that there are no conflicts of interest.
	Case Studies	Involves comprehensive examination through cross comparison of cases to obtain in-depth information with the goal to fully understand the operational dynamics, activities, outputs, outcomes and interactions of a development project or programme.	Useful to fully explore factors that contribute to outputs and outcomes.	Requires considerable time and resources not usually available for commissioned evaluations. Can be difficult to analyze.
Quantitative	**Surveys (samples of respondents, including project/programme and control observations)**	A sample of the project/programme population(and possibly of a control group) is extracted. Questionnaires are usually administered face-to-face by enumerators on the basis of a pre-written and pre-coded questionnaire. Entries are recorded on electronic support media and analysed using computer software on the basis of standard descriptive, inferential and econometric techniques.	The sampling procedure should aim to select a statistically representative sub-set of the population . Large samples allow for more refined analysis and are representative of more sub-categories of the population (sub-region, province, etc.)	Trained specialists are required for survey design planning and data analysis Larger surveys can be costly and time-consuming to implement

Sources: UNDP, 2009 and IFAD, 2009, amended.

81. Options in terms of evaluation design are often limited by the absence of baseline data, i.e. the description and documentation of the specific situation in the country/area targeted for change prior to the ILO intervention. Likewise, the absence of data relative to the evolution of outcomes for a comparison group makes accounting systematically for interventions' effects and causal links particularly difficult. Box 10 explains how to deal with the lack of baseline data when undertaking an evaluation.

Box 10 What to do when baseline data is not available?

Where baseline surveys have not been undertaken or are not of the required quality, the approach paper should identify how data collection ought to proceed in order to secure a plausible proxy for the assessment of initial conditions.
For instance, evaluators may conduct in-depth interviews with project beneficiaries and have them reconstruct – using memory recall, structured interviews and/or focus groups discussions – the logical chain of behavioral, productive or organizational changes generated or supported by the project. Evaluators should exercise caution and triangulate the information secured from diverse sources.

Source: IFAD, 2009: Evaluation Manual: Methodology and processes, page 19

3.3 Triangulation

82. For the process of data collection and analysis, it is recommended to use a combination of methods to ensure data accuracy and facilitate its interpretation. This use of mixed methods and data from mixed sources is called triangulation and defined by the OECD/DAC as "the use of three or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment."[19]

83. The technique of triangulation allows evaluators to overcome the bias that comes from single information sources, the use of single methods or single observations. Hence triangulation strengthens the accuracy, robustness and reliability of evaluation results.

3.4 Reporting

84. For each evaluation report, a draft and a final version must be prepared. The draft version provides stakeholders with an opportunity to give feedback prior to the preparation of the final evaluation report.

85. Certain standard elements should be addressed in every evaluation report. Annex 3 contains a checklist aimed at ensuring the quality of evaluation reports. The precise structure of a particular evaluation report depends on the specific focus, needs and circumstances of the project or programme and its evaluation.

86. Box 11 presents a ***sample*** structure and table of contents of an evaluation report and identifies the standard elements that should be addressed in each evaluation. The single elements are explained below.

87. The body of the evaluation report should not usually exceed 20 to 30 pages. Annex 1 contains a sample title page that must be used for all ILO evaluation reports.

Box 11 Sample structure and table of contents of an evaluation report

Cover page with key intervention and evaluation data

1. Abstract
2. Brief background on the project and its logic
3. Purpose, scope and clients of evaluation
4. Methodology
5. Review of implementation
6. Presentation of findings regarding project performance
7. Conclusions
8. Recommendations
9. Lessons learned

Annexes

3.4.1 Use the right language style, be constructive

88. Evaluation users are often busy and need easily readable reports in plain English. An evaluation report that stimulates the readers' interest, matches their decision-making and learning requirements, and economizes with their time, enhances the overall value of the evaluation. The use of non-defensive language and a constructive writing style stimulates

[19] OECD/DAC, 2002: Glossary of key terms in evaluation and results based management.

acceptance of evaluation reports. Experience shows that even the most controversial evaluation findings can find acceptance when they are backed with evidence and presented in a sensitive and fair-minded way. On the contrary, investments in evaluations have sometimes failed to add value when presented in a language and style that alienated key stakeholders.

3.4.2 Structure the report, stick to the reporting logic

89. Written reports need to be well-structured, clear and concisely written in relevant languages. Evaluators should **not express their own opinions** in the section on findings but report the evidence presented to them, and then draw their conclusions. Interested parties may try to influence the content of an evaluation report. While the evaluation team should always be open to input from stakeholders, it is important the team hold their ground where no clear evidence can be found for changing their findings. Robust discussions of findings can be expected, but intimidation or other unethical behaviour from third parties to influence the independent evaluators should be reported to the evaluation manager.

Further reading about the structure of an evaluation report is provided in the ILO Technical Cooperation Manual.
For further reading about the reporting process, including feedback also refer to the ILO Technical Cooperation Manual.

3.4.3 Commenting on the draft evaluation report

90. Evaluators are expected to submit only complete and readable draft reports. Evaluators send the completed draft report to the evaluation manager who, after reviewing the draft for adequacy and readability, circulates it to all stakeholders all at once. This avoids any opportunity for one stakeholder group to edit the draft prior to wider circulation. Stakeholders are encouraged to make written comments but to not directly edit the document. Comments may be sent individually to the evaluation manager by stakeholders on a confidential basis, and/or collectively. The evaluation manager is expected to respect the confidentiality of those commenting and to forward all comments in a single communication to the evaluator.

3.4.4 Process for approving the final evaluation report

91. The checklist for ensuring the quality of evaluation reports is a useful support tool in the process of approving the final evaluation report.[20] Specific attention should be paid to two key outputs of the report: the lessons learned captured during the evaluation and recommendations on follow-up to the report.

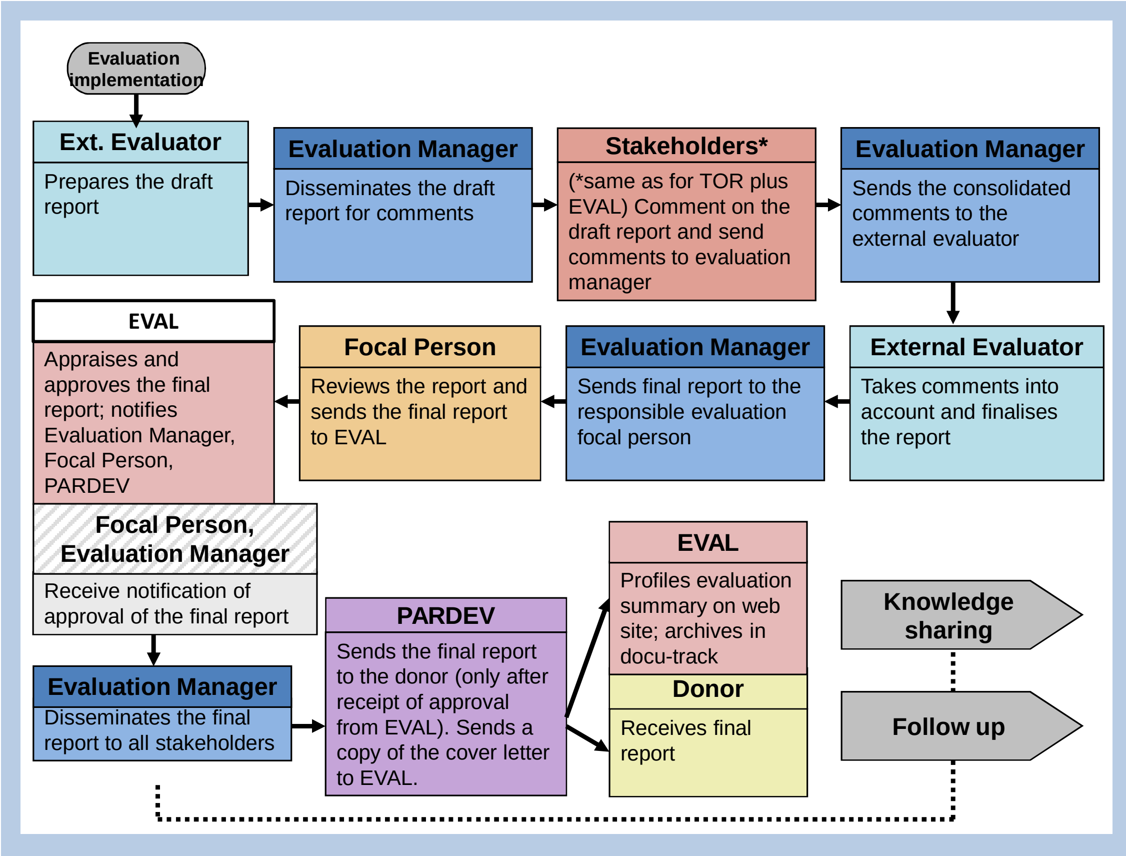
92. Evaluation recommendations that are specific, measurable, achievable, realistic and time-bound are essential for meaningful management responses. The following questions emerge:

- Are the recommendations clearly based on findings and conclusions?
- Are recommendations clear, concise, constructive, realistic and of relevance to the intended user(s)?
- Are the recommendations specific on what should be done and when to implement the recommendation?

93. Key steps in the preparation of the evaluation report are recapped in Figure 3.

[20] See Annex 3. The checklist can of course also be used to ensure the quality of drafts.

Figure 3 Workflow and preparation of the evaluation report



3.5 Learning and generating knowledge from evaluation

94. A lesson learned is a knowledge [artefact] or understanding gained by experience. The experience may be positive, as in a successful test or mission, or negative, as in a mishap or failure. Successes are also considered sources of lessons learned. A lesson must be significant in that it has a real or assumed impact on operations; valid in that it is factually and technically correct; and applicable in that it identifies a specific design, process, or decision that reduces or eliminates the potential for failures and mishaps, or reinforces a positive result[21].

95. Generally, lessons learned serve the purpose of summarizing knowledge and/or understanding gained from experience (including good practices). More specifically, evaluation lessons learned:

- highlight interventions' observed strengths and weaknesses;
- allow practitioners to learn from previous experience and avoid "reinventing the wheel" as well as repeating unsuccessful outcomes;
- contribute to learning and knowledge-sharing amongst stakeholders by focusing the debate on a set of actionable principles and thereby contribute to improving quality of delivery; and
- help stakeholders (e.g. national constituents, projects partners, donors, project management team, responsible ILO field office, field technical specialist(s) and the ILO technical unit at headquarters which backstops the project) to better understand the design,

[21] ILO guidance on evaluation lessons learned.

monitoring and evaluation of a given intervention, and to identify where collaboration and coordination need to be strengthened.

96. Lessons learned are designed to be used, wherever appropriate. When they are applied, they should impact positively on programme or project performance, outcome, or impact.

97. There are many ways in which lessons learned can be categorized. For the purpose of categorizing lessons learned, the focus may be on organizational processes to which they are correlated (procedural lessons learned). Focus may also be based on the programme or project activities evaluated (substantive lessons learned), or on the corresponding characteristics (general and abstract versus targeted and specific).

98. Evaluation lessons learned can address both the internal and external logic of interventions. Accounting for the internal logic can help project and/or programme managers to know whether they are "doing things right." i.e. account for how to improve a current way of working or managing activities by, for instance, identifying shortcomings and mishaps and correcting them.

99. By contrast, focusing on the external logic helps to know whether the "right thing is being done," i.e. question the assumptions about how the project, programme or policy works, including the manner in which it fits into the broader context and environment. The intervention's concept/theory of change, design, development objective and strategy are in that case subject to critical scrutiny.

3.5.1 Evaluation lessons learned

100. Each ILO evaluation report has to contain a section on lessons learned. These lessons have to be captured, validated, stored, disseminated and reused if they are to fulfill their purpose.

101. The capturing of lessons to be learned describes the process of (i) gathering, (ii) documenting and (iii) analyzing feedback (e.g. evaluation findings) on projects or programmes. The capturing may occur at the end of the project/programme or at the end of each phase of the intervention cycle (e.g. design, implementation, or evaluation). In this exercise, evaluation findings are scrutinized for lessons to be learned having conceptual use. The following questions might be of use in identifying evaluations findings that can be translated into lessons:

- What went well – either for the project/programme or for the management team?
- What activities or tasks were concerned?
- What is the supporting evidence?
- What went wrong or what did not go well or had unintended outcomes?
- What were the drawbacks?
- What activities or tasks were concerned?
- What is the supporting evidence?
- What unintended outcomes did the intervention dealt with?
- Were the project/programme outcomes achieved?
- If key stakeholders had to do things, over again, what would they likely do differently?
- Did key stakeholders have a clear description of what has been learned?
- Is this description general or specific and targeted?
- What recommendations could be made, so that the same tasks or activities in future

interventions will achieve better results?
- What were the underlying factors that contributed to successful outcomes – either for the project/programme or the management team?
- What were the underlying factors that contributed to drawbacks of implementing activities?
- How well did the intervention objectives suit the priorities of the targeted groups or other clients?
- How well were analyses, formulation and implementation planning of the project/programme coordinated?
- Did project/programme manager and staff formulate appropriate assumptions with respect to effectiveness and sustainability; development hypothesis; implementation and management?
- How well were the activities and progress kept on track?
- How well did the project/programme integrate monitoring and evaluation arrangements?

102. Lessons generated by evaluation and captured from evaluation reports are normally submitted to a validation process. Criteria for validation can include (but are not limited to) the following:

- **Conformability (and objectivity)**

Are the lessons to be learned extracted from the available evidence (e.g. evaluation findings)? What is the evidence supporting the lessons?

- **Reliability (and dependability)**

Are the lessons to be learned consistent and coherent? Can they be replicated over time and across other settings?

- **Internal Validity (credibility)**

Are the lessons to be learned credible to the key stakeholders (project management team, project partners, funders and others clients)? Do the lessons provide them with a clear and valuable description of what must be learned?

- **External Validity (transferability and fittingness)**

Do the lessons to be learned fit other contexts and how widely can they be generalized? Are the lessons to be learned generic enough to be replicated in other settings?

3.5.2 Disclosure and dissemination of the evaluation report

103. As stated in a recent Director-General announcement[22], officials are expected to fully respect the confidential nature of draft evaluation reports and to strictly follow the guidelines set down for handling such documents.

104. To carry out their functions effectively, evaluators may need to interview staff, and have access to all relevant Office documents. Staff at all levels is expected to cooperate fully with evaluators and to take all necessary steps to ensure timely access to requested information.

105. Final evaluation reports are disseminated in accordance with ILO policy on public information disclosure.[23] In addition, all evaluation information is stored in a central

[22] IGDS Number 75, 31 March 2009.
[23] http://www.ilo.org/intranet/edmsp1/igds/groups/dirdocs/documents/igds/igds_002063.pdf.

repository of evaluation documentation to ensure transparency and accessibility (see i-track database).

106. For independent project evaluations, all key project stakeholders – i.e. the donor, the national constituents and key national partners as well as concerned ILO officials – receive a copy of the full evaluation report via mail or email once it is finalized. Assuring this is the responsibility of the evaluation manager of the project evaluation with oversight of the responsible evaluation officer.

107. EVAL collects and stores all reports of independent project evaluations. A list of all independent project evaluations is made available on the public website of EVAL. Also available on the public website of EVAL are summaries of the evaluation reports, composed of two to three pages that provide the main findings, conclusions, recommendations and lessons learned (in English language only). The list of reports and the summaries is updated every six months.

108. In addition to the website, a CD-ROM with the summaries of independent project evaluations can be obtained upon request to EVAL.

109. Interested parties can receive the full report of an evaluation in soft copy upon request to EVAL. EVAL reserves the right not to release certain evaluation reports if a report is considered of poor quality or if it contains confidential information whose wider dissemination could harm certain individuals. In case the release of a report is denied, EVAL will always provide an explanation.

110. Knowledge dissemination can be effectively addressed in many ways, the most important methods being conferences, workshops, training sessions, or seminars.

111. In the ILO's context, we believe that evaluation lessons learned can and should be disseminated through minute sheets, and/or Annual Evaluation Reports (AERs) as well as posted on EVAL's intranet, public website and the i-track.

112. EVAL produces each year an annual evaluation report which is presented to the Programme, Financial and Administrative Committee (PFA) of the Governing Body in November. The annual evaluation report summarizes the evaluation activities of the Office:

- the progress made in implementing the evaluation policy;
- all independent evaluations carried out in the reporting period; and
- the management follow-up to high-level evaluations of the previous year.

4 Using evaluation results

113. The main product of an evaluation is not simply the final report, but the engagement of key stakeholders throughout the evaluation process, so that the stakeholders can make use of the evaluation's findings and remain open to implementing the suggested recommendations. As presented in previous sections, thinking about the use of an evaluation should start at the planning stage rather than when the final report is submitted. Evaluations should not be commissioned unless the purpose and end users are clear. Potential users, i.e. the evaluation stakeholders, should be consulted in the planning to make sure that their questions are addressed by the evaluators. Users should be involved during the course of the evaluation perhaps at an early workshop to discuss mid-term findings and certainly later to discuss the actual findings and the draft report. It is more likely that users will find evaluations useful if they meet to discuss the main points face-to-face with the evaluators, instead of just receiving a report for comment.

114. Different potential users will be interested in different parts of the report. Some may be interested in recommendations that affect their work, others may be interested in broader lessons. It may be worth planning a communication strategy for the evaluation report to decide the following:

- Who needs information from this evaluation?
- How would they like to receive the information?
- When should communication be organized and effective?

115. The following paragraphs mostly outline the procedures for storing and disseminating evaluation results in the ILO.

116. The project manager, the ILO responsible official, the evaluation manager and the evaluation focal person are encouraged to disseminate the abstract of the evaluation report to other interested individuals inside and outside the Office.

117. The relevant technical specialists in headquarters and the field should also make an effort to disseminate relevant lessons learned to interested officials in the Office. This can be done, for example, via email or a newsletter or by posting noteworthy insights on the website.

118. Evaluation reports should be stored in a systematic manner and the knowledge generated in evaluations should be systematically fed into the design of new projects or the next phase of a project. ILO officials should remember to consult previous relevant evaluation reports when developing technical tools and designing new projects and approaches.

4.1 ILO i-track

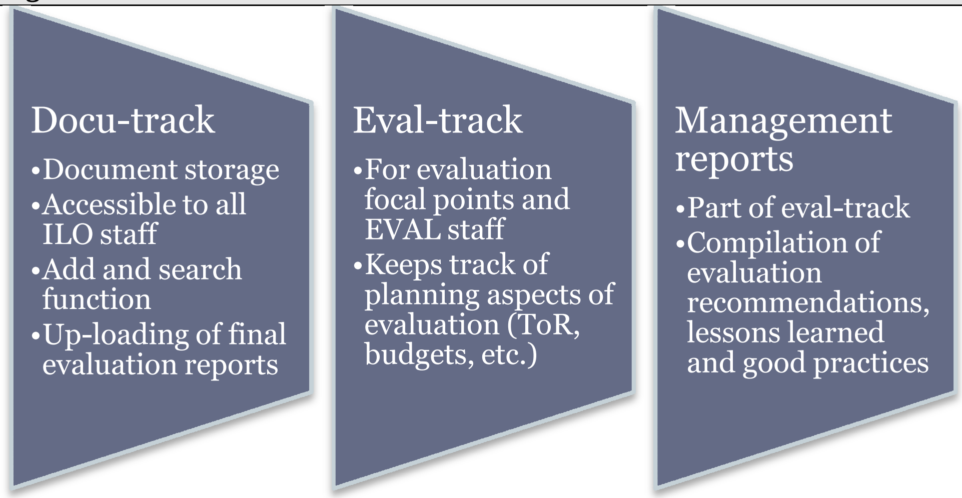
119. A pro-active approach to knowledge sharing is intended to strengthen organizational learning and improvements in ILO technical work and the quality of new project proposals. As hinted at and illustrated in box 5, EVAL has developed an internet-based evaluation database called i-track.

120. The use of i-track is mandatory and helps the ILO create the basis for systematic organizational learning and the application of learning in future intervention design.

121. ILO's i-track consists of three main elements: "Docu-track", "Eval-track" and "Management reports". As shown in Figure 4, "Docu-track" is the placeholder for evaluation

reports classified by year, type, country/region and theme as well as summaries of those reports.

Figure 4 Core elements of i-track



4.2 Management response

122. The purpose of management response and follow-up mechanisms is to strengthen the use of evaluation report findings, to increase stakeholder and management buy-in to the findings of the evaluation, to facilitate in-depth dialogue about evaluation results, and to ensure follow-up of agreed recommendations through formal processes.[24]

123. Evaluations only lead to organizational improvements if recommendations are given systematic management follow-up by line management. Independent high-level strategy, policy and country programme evaluations are presented to the November Governing Body and serve as decision-making papers. An official management response from the Office forms part of the report. A summary of the status of implementation of evaluation recommendations is reported in the Annual Evaluation Report and presented to the November Governing Body of the following year. The Office is accountable to the Governing Body for implementing the recommendations of these evaluations. For that reason, the Director-General has created the Evaluation Advisory Committee (EAC) to monitor and ensure adequate management follow-up to high-level evaluations.[25]

124. In order to systematise the management response and follow-up reporting, the Office has established certain procedures and templates.

125. Within a month after finalisation of the evaluation report, project or line management has to deliver its management response and follow-up action planning, which is being considered by both EVAL and the ILO responsible official. Six months later, the first follow-up report on the implementation of the recommendations is due. For recommendations that need more time for implementation, further reporting on follow-up is needed. Further follow-up reports are usually due every six months or as requested by EVAL or the evaluation focal person.

[24] UN Evaluating Group, Task Force on evaluation quality enhancement/ Engelhardt, A.: Management response and follow-up to evaluation recommendations: overview and lessons learned. March 2009.
[25] Circular 245 from 1 September 2006.

126. The management response and plan of action as well as the first progress report are to be uploaded into the i-track system by the responsible evaluation officer.

127. When the evaluation report is finalized, programme or project managers will want to respond to the recommendations by drawing up a response to show how they intend to act on the recommendations. The management response reflects what is accepted and will be implemented and what is not accepted and why.

128. After a mid-term evaluation, the project manager is responsible for preparing and implementing a follow-up plan. The ILO responsible official has to finalize the management response, follow-up plan of action and first report on implementation. The report should be sent to the regional evaluation officer or the sectoral evaluation officer for uploading into i-track.

129. Table 8 highlights the roles and responsibilities of specific actors for the management response to high-level evaluations.

Table 8 Management response: roles and responsibilities for high-level evaluations

Actor	Roles and responsibilities
EVAL	- Supports Evaluation Advisory Committee in monitoring the management response by requesting and ensuring timely reporting from line management; - Final editing responsibility for Annual Evaluation Report, including section featuring report on follow-up to high-level evaluations; - manages independent high-level evaluations and ensures that evaluation recommendations comply with the quality criteria.
Responsible line manager	- Executive Director (Policy and strategy evaluations) or Regional Director (Country Programme evaluations) to submit management response via EVAL to Evaluation Advisory Committee; - Follow up on evaluation recommendations; - Coordinate implementation with other entities of the ILO as applicable.
Evaluation Advisory Committee	- Monitor and ensure adequate management follow-up - Meet each year in September to review management responses and follow-up reports - Decide whether further follow-up actions are required - Advises Director-General on the information contained in the Office response to evaluation summaries presented to the Governing Body - Advises EVAL on how to feature report on follow-up to high-level evaluations
The Governing Body	- Recipient of the status report on the implementation of recommendations as part of the Annual Evaluation Report as well as of the Office management response transcribed therein.

Further reading about the procedures, timing and templates for management response is provided in "Guidance for management response and follow-up reporting for high level evaluations" and "Explanations to the templates on management response and follow-up reporting for high-level evaluations."

5 Quality standards for evaluations

130. ILO evaluation quality standards are consistent with and directly inspired from the UNEG norms and standards outlined in the "Norms and Standards for Evaluation in the UN system (2005)". According to the UNEG, "each evaluation should employ design, planning and implementation processes that are inherently quality oriented, covering appropriate methodologies for data collection, analysis and interpretation. Evaluation reports must present in a complete and balanced way the evidence, findings, conclusions and recommendations. They must be brief and to the point and easy to understand. They must explain the methodology followed, highlight the methodological limitations of the evaluation, key concerns and evidence-based findings, dissident views and consequent conclusions, recommendations and lessons. They must have an executive summary that encapsulates the essence of the information contained in the report, and facilitate dissemination and distillation of lessons."[26]

131. Additional ILO-specific criteria are outlined in various resources including an Office Directive (IGDS Number 74) and the quality checklist for evaluation reports.[27] There should in particular be:

- separation of evaluation responsibility from line management functions for programmes and projects;
- limited management influence over terms of reference, scope of the evaluation and selection of evaluators;
- transparency and clarity regarding the evaluation process;
- involvement of constituents and others as appropriate, in the planning and reporting process; and
- authority over the drafting of the final report, including findings and recommendations, subject to internationally agreed standards.

5.1 EAC oversight of evaluation use

132. The aforementioned Evaluation Advisory Committee (EAC) is established to provide a mechanism to oversee the use, implementation and follow up to lessons learned and recommendations resulting from ILO evaluation activities. Its objective is to promote institutional follow-up of independent evaluation findings and accepted recommendations and to provide pertinent information and advice to the Director-General on the progress made by the Office in this regard.

133. The scope of its functions includes all independent evaluations with particular emphasis on strategy and policy evaluations, country programme evaluations and major thematic evaluations. The Committee may also consider feedback on follow-up plans and actions taken in respect of a selected number of large technical cooperation projects.

134. All managers are accountable for ensuring proper use of relevant evaluation findings, lessons learned and recommendations. However, without substituting the reporting and accountability obligations of managers, the EAC provides additional assurance to the senior management team and to the Director-General that follow-up to evaluation recommendations is transparent and regularly conducted.

135. The Committee also functions as a forum for internal dialogue on the implementation of the ILO evaluation policy and strategy and, in particular, ensures that evaluations are

[26] UN Evaluation Group: Norms for Evaluation in the UN System, April 2005.
[27] See part 1 of the guidelines Annex 2.

credible and conducted in an impartial and independent way. It may discuss the draft plan for independent evaluations and provide its recommendations to the Director-General or EVAL, as appropriate. The Committee verifies that all independent evaluation reports are made public through appropriate means of dissemination, including the Internet.

136. In relation to the preparation of programme and budget proposals, the Committee provides its views to PROGRAM as to the findings and recommendations of evaluation reports that could be used in developing new programme and budget proposals. The Committee also provides advice on any other evaluation issues that may be referred to by the Director-General.

137. Finally, an external quality appraisal exercise of independent project evaluation reports is to be conducted each year since 2008. The aim is to promote discussion around strengthening the quality of future evaluations. The quality checklist contains some 70 criteria sub-categorized into eleven sub-sections. These are drawn from evaluation norms and standards as laid out in the UN System Evaluation Norms and Standards and the OECD/DAC Evaluation Quality Standards.

Annex 1 Sample title page for ILO evaluations

Evaluation Unit (EVAL)

Evaluation Title Page

- **Project Title:**
- **TC/SYMBOL:** **(use the most recent if phased)**
- **Type of Evaluation:** **(final or midterm)**
- **Country(ies):**
- **Project End:**
- **Evaluation Manager:**
- **Administrative Unit:**
- **Technical Unit:**
- **Evaluation Team:**
- **Date Evaluation Completed:**
- **Key Words:**

Annex 2 Quality checklist for evaluation Terms of Reference

This checklist is a tool for evaluation managers and evaluation focal persons to appraise the quality of terms of reference for independent project evaluations.

Project[28] title: …
TC code: …
Responsible administrative unit: …

Rating dimensions
0 = No (or not addressed)[29]
1 = Clear need for improvement
2 = Sufficient level of quality
3 = Yes (or high quality)
N/A = Not applicable[30]

A		**Introduction and rationale for evaluation**	Score
1		Brief explanation of why the evaluation is being conducted.	
2	6	Brief explanation, what the expected outcome should be.	
3	7	Mention of the type of evaluation (independent or internal) [0/3 only]	
4	8	Mention of the timing of evaluation (mid-term or final) [0/3 only]	
5	9	Mention that the evaluation will comply with evaluation norms and standards and that ethical safeguards will be followed.[31]	

Comments:

B		**Background on project and context**	Score
1	10	Mention of key project data (duration, location, budget, donor, partners).	
2	11	Summary of the project's rationale, internal logic, and strategy approach.	
3	12	If useful, brief outline of economic, political, social, cultural, historical context of the country and how it influenced the project [N/A possible]	
4	13	Brief account of the project's management set-up.	

[28] The term "project" is used in the following as a generic term for any kind of development intervention that is being evaluated, be it a larger programme, several projects, or project components. If two or more projects are being evaluated together, provide titles and codes of all projects.
[29] Some criteria have only a binary dimension of yes/no or there/ not there. These criteria are marked [0/3 only] and can only be rated either 0 or 3 but not 1 or 2.
[30] Only criteria marked [N/A possible] can be rated N/A, all others require a rating.
[31] Reference: UN Evaluation Norms and Standards and OECD/DAC Evaluation Quality Standards. See: http://www.ilo.org/eval/policy.

5	13.1	Brief account of the project's development (major milestones), and current implementation phase (e.g. piloting, towards middle of implementation, in phase out)	
6	14	Mention of any previous evaluations or reviews, if applicable. [N/A possible]	

Comments:

C	**Purpose, scope and clients of the evaluation**		Score
1	15	Explanation of the purpose (objectives) of the evaluation, its primary use and what the evaluation should achieve.	
2	16	Specification of the scope of the evaluation (geographic coverage, whole project or just one component or several projects, one or several phases etc.).	
3	17	Specification of certain issues or aspects the evaluation should specially focuses on (for example gender, exit strategy etc.), if applicable. [N/A possible]	
4	18	Identification of the clients of the evaluation and the main audience of the report.	

Comments:

D	**Key evaluation questions/ analytical framework**[32]		
1	19	Reference to the standard evaluation criteria relevance, effectiveness, efficiency, impact and sustainability or more specifically - Relevance and strategic fit, - Validity of design, - Project progress and effectiveness, - Adequacy and efficiency of resource use, - Effectiveness of management arrangements, - Impact orientation, and - Sustainability.	
2	20	Requirement to analyze the project's performance related to ILO's cross-cutting issues gender, poverty, labour standards, and tripartism/ social dialogue.	
3	21	Listing of the main evaluation questions (specific to project and evaluation purpose).	
4		Suggested analytical framework with sub-questions.	

Comments:

[32] Remark: The evaluator may adapt the evaluation questions and analytical framework, but any fundamental changes should be agreed between the evaluation manager and the evaluator.

E	**Main outputs of evaluation**	Score
1	Main outputs of evaluation are specified (e.g. evaluation design, draft report, final report, stakeholder workshop etc.). Required outputs are draft report, final report and evaluation summary according to ILO template[33].	
2	Statement that the quality of the report will be determined by conformance with the quality checklist for evaluation reports.	

F	**Methodology**[34]	Score
1	Identification of the information needs and possible sources of information.	
2	Description of the method for sampling/ selection of case studies, if applicable. [N/A possible]	
3	Description of the suggested methodology and plan for information gathering and organizing (survey, interviews, case studies, etc.).	
4	Requirement that all data should be sex-disaggregated and different needs of women and men should be considered in evaluation process.	
5	Plan for analysis of data/information	
6	Identification of the conditions and capacities needed to support data gathering, analysis and communication.	
7	Plan for critical reflection processes and for quality communication and reporting of evaluation outcomes (e.g. stakeholder workshop, debriefing of project manager).	
8	Description of the involvement of the key stakeholders in the implementation of the evaluation, including in the finalization of the report.	

Comments:

[33] The ILO publishes all summaries of independent evaluation reports on its website. It has developed a special template for this purpose. (See: http://www.ilo.org/eval/guidance > guidance for project evaluations.) It is the responsibility of the evaluation manager to provide the summary conform to the template. In order to avoid double effort, it is advisable that this template is already used for the summary in the evaluation report.

[34] Remark: Planning the methodology to be used in the course of the evaluation and including it in the ToR ensures transparency and helps planning the budget. The evaluator may adapt the methodology, but any fundamental changes to the methodology should be agreed between the evaluation manager and the evaluator, in particular if they have cost implications.

G	**Management arrangements, work plan and time frame**	Score
1	Mention of who the manager of the evaluation is, i.e. who the evaluation team leader reports to.	
2	Where there is more than one evaluator, definition of reporting lines within the evaluation team. [N/A possible]	
3	Desired composition of the evaluation team (e.g. one international, one local evaluator, participation of ILO or donor representative, gender balance).	
4	Specification of the support needed from the ILO at headquarters, regional, sub-regional and country levels for implementing the evaluation.	
5	A work plan, stipulating each partner's contribution to the evaluation is included.	
6	A time frame, with deadlines for each output or major step in the process is included.[35]	

Comments:

[35] In order to for evaluations to be useful, they should deliver important information into key decision-making processes. Timing is essential. The timing of outputs (final report, stakeholder debriefing etc.) should therefore consider the timing of crucial decision-making events of the main clients of the evaluation.

Annex 3 Checklist for the quality of evaluation reports

This checklist is a tool for the evaluation focal persons and EVAL to appraise the quality of independent project evaluation reports. It can also serve as guidance to the evaluation manager and the evaluator. EVAL is providing the concerned evaluation manager with feedback by returning the completed checklist to him/her. The results of the appraisal will not be made public. Aggregate results for all reports for one year are being reported in the Annual Evaluation Report that is presented to the Governing Body in November.

Project[36] title: …
TC code: …
Responsible administrative unit: …

Rating dimensions
0 = No (or not addressed)[37]
1 = Clear need for improvement
2 = Sufficient level of quality
3 = Yes (or high quality)
N/A = Not applicable[38]

Contents of report

A	Executive Summary[39]	Score
1	Summary identifies the type (internal/ independent), timing (mid-term/ final) and the clients of the evaluation.	
2	Summary provides an overview of the purpose and scope of the evaluation.	
3	Summary provides a brief description of the intervention rationale of the project being evaluated.	
4	Summary provides a brief description of the evaluation method(s).	
5	Summary provides a summary of the main findings and conclusions.	
6	Summary provides recommendations and lessons learned.	

Comments:

B	Project background	Score
1	Description of the problem context and intervention logic of the project, including demand for ILO's support.	

[36] The term "project" is used in the following as a generic term for any kind of development intervention that is being evaluated, be it a larger programme, several projects, or project components. If two or more projects are being evaluated together, provide titles and codes of all projects.
[37] Some criteria have only a binary dimension of yes/no or there/ not there. These criteria are marked [0/3 only] and can only be rated either 0 or 3 but not 1 or 2.
[38] Only criteria marked [N/A possible] can be rated N/A, all others require a rating.
[39] The ILO publishes all summaries of independent evaluation reports on its website. It has developed a special template for this purpose. (See: http://www.ilo.org/eval/guidance > guidance for project evaluations.) It is the responsibility of the evaluation manager to provide the summary conform to the template. In order to avoid double effort, it is advisable that this template is already used for the summary in the evaluation report.

2	If useful, brief outline of economic, political, social, cultural, historical context of the country and how it influenced the project [N/A possible]	
3	Description of project objectives.	
4	Description of funding arrangements for the project (including the donor).	
5	Description of the organisational arrangements for project implementation.	
6	Description of contributions and role of ILO, the project partners and other stakeholders.	
7	Brief description/ review of the project's implementation (major events and milestones)	

Comments:

C	Background on evaluation	Score
1	Description of purpose and primary use of the evaluation.	
2	Description of the scope of the evaluation (for example geographic coverage, only second phase or several projects).	
3	Description of special focus areas (for example gender, collaboration, exit strategy etc.).	
4	Brief description of the operational sequence of the evaluation including dates of evaluation mission.	
5	Mention of who are the clients of the evaluation and the main audience of the report (donors, constituents, implementing parties, …). *	
6	Name(s) of evaluator(s) included. *	
7	Name of evaluation manager with administration unit. *	
8	Cost of evaluation included. *	
9	Acknowledgements included (or as separate section)	

Comments:

* If mentioned elsewhere: rating 1 or 2.

D	Methodology	Score
1	The main evaluation criteria are identified. *	
2	The main evaluation questions are identified. *	
3	Description of the evaluation methods and data collection instruments used.	
4	Brief discussion/ justification of the relevance and validity of the methods and instruments.	
5	If sampling or case studies are being used, an explanation of the sample or case selection is included. [N/A possible]	
6	Description of sources of information/ data used.	
7	Limitations and potential sources of bias are recognised (of methods, case selection, data sources etc.).	
8	Description and rationale for stakeholder participation in evaluation process is given.	
9	The report indicates that evaluation norms, standards and ethical safeguards[40] have been followed. *	

[40] Reference: UN Evaluation Norms and Standards and OECD/DAC Evaluation Quality Standards. See: http://www.ilo.org/eval/policy.

Comments:

* If mentioned elsewhere: rating 1 or 2.

E	Findings	Score
1	All evaluation questions are addressed or an explanation is given for questions that could not be answered.	
2	Findings are relevant to the purpose and scope of the evaluation.	
3	Findings are supported by the evidence presented and are consistent with methods and data.	
4	Factors that have contributed to the success/ failure of the project (or elements) are identified and discussed.	
5	Findings related to the relevance, strategic fit and validity of design of the project are discussed (including adequacy of objectives and indicators).	
6	Findings related to the implementation and the delivery process (activities and outputs) are discussed.	
7	Findings related to the effectiveness of the project are discussed (achievement of outcomes/ objectives or progress made).	
8	Findings related to the adequacy and efficiency of the resource use are discussed.	
9	Findings related to the effectiveness of the management arrangements are discussed (including implementation monitoring, outcome monitoring, backstopping, and collaboration with other projects).	
10	Findings related to the likelihood of the project to have longer-term development impacts are discussed.	
11	Findings related to unintended or unexpected effects are discussed.	
12	Findings related to the (likely) sustainability of the project are included.	
13	Findings on tripartism and the promotion of social dialogue are included.	
14	Findings on the performance of the project in promoting the ratification or application of labour standards are included.	
15	Findings on the effectiveness of the project to reduce poverty are included.	
16	Findings on the gender performance of the project are included.	

Comments:

F	Conclusions	Score
1	Conclusions provide summary judgments about the merit and worth of the project.	
2	Conclusions follow from findings.	
3	Judgments are fair, impartial, and consistent with the findings.	

Comments:

G	Recommendations and lessons learned	Score
1	Recommendations are based on findings and conclusions of the report.	
2	Recommendations are clear, concise, constructive and of relevance to the intended user(s).	

3 Recommendations are realistic and actionable (including who is called upon to act and recommended timeframe).

4 Lessons learned and good practices are firmly based on the evaluation but are generalisations beyond the project being evaluated. [N/A possible]

Comments:

Formal elements

H	Title page	Score
1	Title of project(s) evaluated included. [0/3 only]	
2	TC code(s) of project(s) included. [0/3 only]	
3	Type of evaluation (independent or internal) included. [0/3 only]	
4	Timing of evaluation (mid-term or final) included. [0/3 only]	
5	Name(s) of evaluator(s) included. [0/3 only]	
6	Date of final report included. [0/3 only]	

I	Appendices	Score
1	Terms of reference included. [0/3 only]	
2	List of persons or organisations interviewed included. [0/3 only]	
3	Data collection instruments (like questionnaire) included.	
4	List of publications referenced, cited or mentioned in report included. (List of working documents, websites etc. is optional.) [N/A possible]	
5	Appendices are clearly labelled, presented and referenced.	

J	Formal elements	Score
1	Table of contents is accurate and contains at least all first and second level headers in the report. [0/3 only]	
2	List of appendices is included. [0/3 only]	
3	Lists of tables, figures and charts are included (if more than five in report). [N/A possible] [0/3 only]	
4	List of acronyms/ abbreviations included. [0/3 only]	
5	Tables, figures and charts are clear and labelled correctly.	
6	All data is sex-disaggregated. [0/3 only]	
7	Data is broken down by other social categories (age, ethnic group ...), if relevant. [N/A possible] [0/3 only]	
8	A consistent and accurate system is used for references and footnotes/ endnotes. [N/A possible] [0/3 only]	