iod parc

Independent Evaluation of the ILO's Evaluation Function for the Period 2017-2021

Final Evaluation Report

Acknowledgements

Table of Contents

List of Figures

List of Tables

Annexes

Acronyms and Abbreviations

ACTEMP	Bureau for Employers' Activities
ACTRAV	Bureau for Workers' Activities
AER	Annual Evaluation Report
COVID-19	Coronavirus Disease 2019
CTA	Chief Technical Adviser
DDG	Deputy Director General
DEFP	Department Evaluation Focal Points
DFID	Department for International Development
DG	Director General
DWCP	Decent Work Country Programme
EAC	Evaluation Advisory Committee
EC	European Commission
EQs	Evaluation Questions
ESRC	Economic and Social Research Council
EVAL	Evaluation Office of the ILO
FGD	Focus Group Discussion
FOP	Field Operations and Partnership Portfolio
GB	Governing Body
GE	Gender Equality
GEDI	Gender Equality Discrimination and Inclusion Branch
GEEW	Gender equality and empowerment of women
HLE	High Level Evaluation
HR	Human Resources
IEE	independent Evaluation of ILO's Evaluation Function
IEE TC	Independent Evaluation of the Evaluation Function Technical Committee
IAO	Office of Internal Audit and Oversight
IFC	International Finance Corporation
IFI	International Finance Institutions
ILC	International Labour Conference
ILO	International Labour Organization
IMEC	Industrialized Market Economy Countries
IOE	The International Organisation of Employers
IOM	International Organisation for Migration
IPEC+	International Programme on the Elimination of Child Labour and Forced Labour
ITUC	International Trade Union Confederation
JIU	Joint Inspection Unit
KII	Key Informant Interview

LNOB	Leave no one behind
MOPAN	Multilateral Organization Performance Assessment Network
M&E	Monitoring and Evaluation
M/F	Male/Female
OECD/DAC	Organization for Economic Co-operation and Development/ Development Assistance Committee
P&B	Programme and Budget
PARDEV	Partnerships and Field Support Department
PROGRAM	Strategic Programming and Management Department
QA	Quality Assurance
RBM	Results-Based Management
REACs	Regional Evaluation Advisory Committees
REO	Regional Evaluation Officers
RBSA	Regular Budget Supplementary Account
RBTC	Regular Budget Technical Cooperation
SDGs	Sustainable Development Goals
SIDA	Swedish International Development Cooperation Agency
SWAP	UN System Wide Action Plan
TC	Technical Cooperation
ToC	Theory of Change
ToR	Terms of Reference
UN	United Nations
UN-SWAP	UN System-wide Action Plan on Gender Equality and the Empowerment of Women
UNEG	United Nations Evaluation Group
UNDP	United Nations Development Programme
UNICEF	United Nations Children's Fund
USDOL	United States Department of Labor

Executive Summary

Purpose and scope of the evaluation

This independent evaluation of the evaluation function (IEE) fulfils a decision to provide an independent review and analysis of organisation's evaluation function every five years[1]. The evaluation reviews the role that the evaluation function plays in supporting the ILO's requirements for accountability and organisational learning and informs the next update of the Evaluation Strategy (2022-25).

Specifically, the scope of the IEE is the "evaluation function" which includes:

- The 2017 evaluation policy
- The 2018-21 evaluation strategy
- The centralized evaluation function that includes the ILO independent Evaluation Office (EVAL); The Evaluation Advisory Committee (EAC) and the systems, structures, and evaluations within EVAL's remit (e.g. quality assurance, knowledge management, communications etc.)
- The decentralized evaluation function that includes evaluations of programmes and projects under control of programme, technical and field units, and regional evaluation officers (REOs), departmental evaluation focal points (DEFPs), and the role of certified evaluation managers that support decentralized evaluations.

The evaluation focuses on six interrelated key questions, aligned to OECD-DAC criteria:

1. Impact: To what extent does the evaluation function help the ILO deliver against its current strategy?
2. Effectiveness: What is the effectiveness of the evaluation function during the strategic period?
3. Efficiency & Sustainability: How efficient, sustainable, and fit-for-purpose are the current arrangements and structures of the evaluation function?
4. Relevance: To what extent does the evaluation function respond to international, organizational, and tripartite policies, interests, and information needs?
5. Adaptability: To what extent is the evaluation function able to respond and flex to emerging priorities and stressors?
6. Coherence: To what extent does the evaluation function align with other accountability and learning processes?

The evaluation makes recommendations on how to strengthen the evaluation function in relation to its independence, credibility, utility, operational framework, evaluation architecture, and the enabling environment in which it operates.

The intention of this report is to provide an independent and constructive professional assessment of the evaluation function to support strategic decision-making in the ILO. It recognises the significant strengths and professional credibility of the ILO's current evaluation function and offers recommendations for enhancing the value and critical role of evaluation in driving effectiveness, continuous improvement and learning within the ILO, for the benefit of its constituents and other key stakeholders in the world of work.

---

[1] A technical committee mandated by the EAC (the TC IEE - 2022), was tasked to oversee the independent evaluation of the evaluation function and to ensure that it was conducted in a manner that enhanced the utility of its findings for the ILO while maintaining the independence of the process.

Methodology

The IEE was guided by three principles: **rigour** (a robust methodological approach using mixed methods), **utility-focused** (a participatory and appreciative approach geared to utilization), and **quality** (adherence to ethical standards, international evaluation norms and standards). The methodology applied mixed methods to triangulate across multiple lines of evidence which included:

- Document review of 238 documents and quantitative analysis of datasets pertaining to evaluations and quality assurance, development cooperation, financial expenditure.

- Interviews and focus group discussions with 116 key informants.

- Surveys of i) ILO staff; ii) ILO senior managers; iii) Evaluation Managers, Regional Evaluation Officers (REO), Departmental Evaluation Focal Points (DEFPs); and iv) Governing Body members (former and present) and non-Governing Body constituents.

- Portfolio reviews of a selection of 20 evaluations and three 'Deep dives' to facilitate a cross-cutting view on the perceptions, practices and contextual factors that influence the evaluation function. These were a) evaluation culture; b) evolving trends and practices in other UN bodies; and c) funding partner influence. The findings of the deep dives and portfolio review underpin and inform the analysis and as such their presentation is embedded throughout the findings of the report, with additional detail provided in Annexes F and Q.

The evaluation used a qualitative analytical software, to code and triangulate documents and interview transcripts based upon the evaluation questions. The evaluation team triangulated data gathered through documentation and quantitative datasets, interviews and FGDs, the portfolio review, deep dives, and the surveys. A range of triangulation methods were applied: cross reference of different data sources; triangulation between team members, and verification of findings and information post-data collection. The IEE process has attached importance to the qualitative information gathered on the premise that how evaluation is perceived matters enormously to its use and value within an organisation. As such, perceptions and observations from internal and external stakeholders are important in shaping evaluation culture, the incentives that drive evaluations and are affected by the enabling environment in which they operate.   There were several limitations of the evaluation - set out in the report 1.4.1 and Annex D.

Background

The aim of evaluation in the ILO is to support improvements in programmes and policies, and to promote accountability and learning (EVAL, 2020)[2]. All aspects of evaluation in the ILO are guided by the ILO evaluation policy and the ILO Evaluation Strategy 2018–21 which adhere to the OECD/DAC Principles (OECD/DAC 2019) and UNEG Norms and Standards for Evaluation (UNEG 2016).

The 2018-22 Evaluation Strategy (extended to cover 2022) responded to the findings of the previous 2016 IEE and took a comprehensive approach to bolstering the evaluation function through three main strategic outcomes underpinned by theory of change (Toc) which posits key assumptions about the role of the enabling environment and evaluation culture. The strategic outcomes are:

- **Outcome 1:** Enhanced capacities and systems of evaluation for better practice and use

---

[2] The IEE team interpret learning as incorporating adaptive planning and management of interventions, policies, and strategies.

- **Outcome 2:** Enhanced value of evaluation through the use of more credible and higher-quality evaluations (independence, credibility, usefulness)
- **Outcome 3:** Stronger knowledge base of evaluation findings and recommendations

The figure below sets out the architecture of the evaluation function:

Figure 1 Architecture of the evaluation function
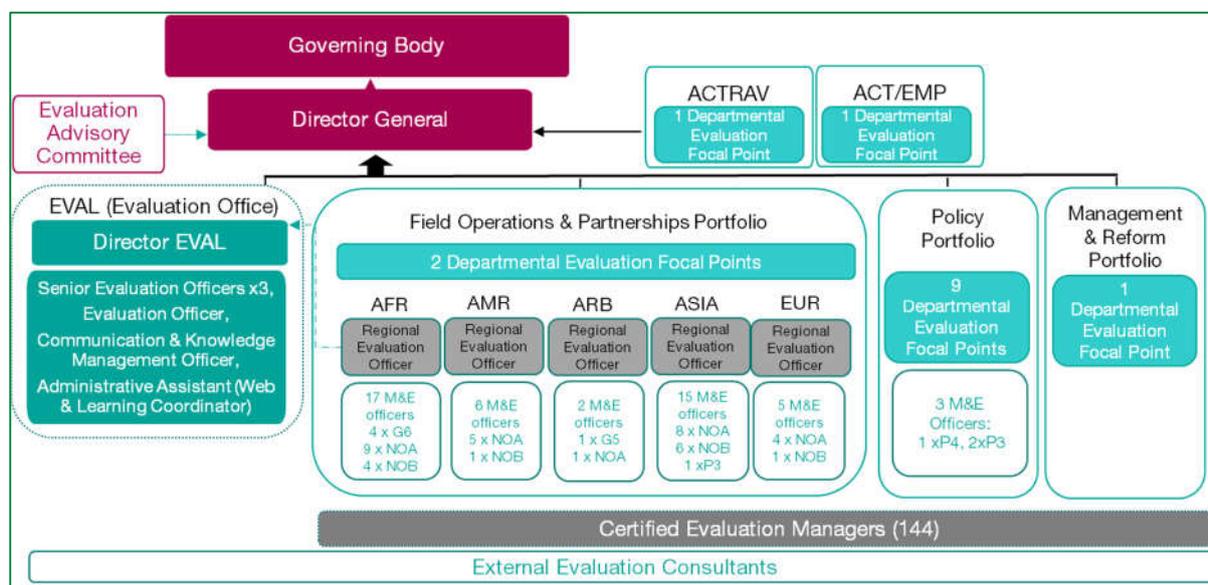
Key Findings

The evaluation function produces centralized and decentralized evaluations[3]. Centralized evaluations, which are at the governance level, are comprised of decent work country programme evaluations, high level evaluations (HLE), and thematic evaluations when there is a specific request. The ILO conducted a total of 429 evaluations during the evaluation period 2017-2021, including: 349 project evaluations (of which 64 were projects over 5 million USD), 30 cluster evaluations (7%), eight country programme evaluations, six thematic evaluations), 20 RBSA funded activity evaluation, six DWCP reviews and six policy and strategic evaluations. The largest share (81 per cent) of ILO's evaluations are decentralized **project evaluations,** with **independent** project evaluations accounting for 53 per cent of the total number of evaluations. Fifteen high level evaluations were undertaken, of which 5 were Decent Works country programmes, an additional 19 synthesis reviews and 4 meta studies were also undertaken. The highest number of evaluations were carried out in Asia and the Pacific and Africa, with 123 and 120 reports respectively. The key findings set out below are a summary overview based on the questions set out in the evaluation framework.

---

[3] All evaluation data was extracted from i-EVAL Discovery April 2022, additional information on synthesis reviews and meta studies is derived from EVAL own data and the ILO website. High level evaluations are recorded in Discovery as strategic, policy, and country programme evaluations.
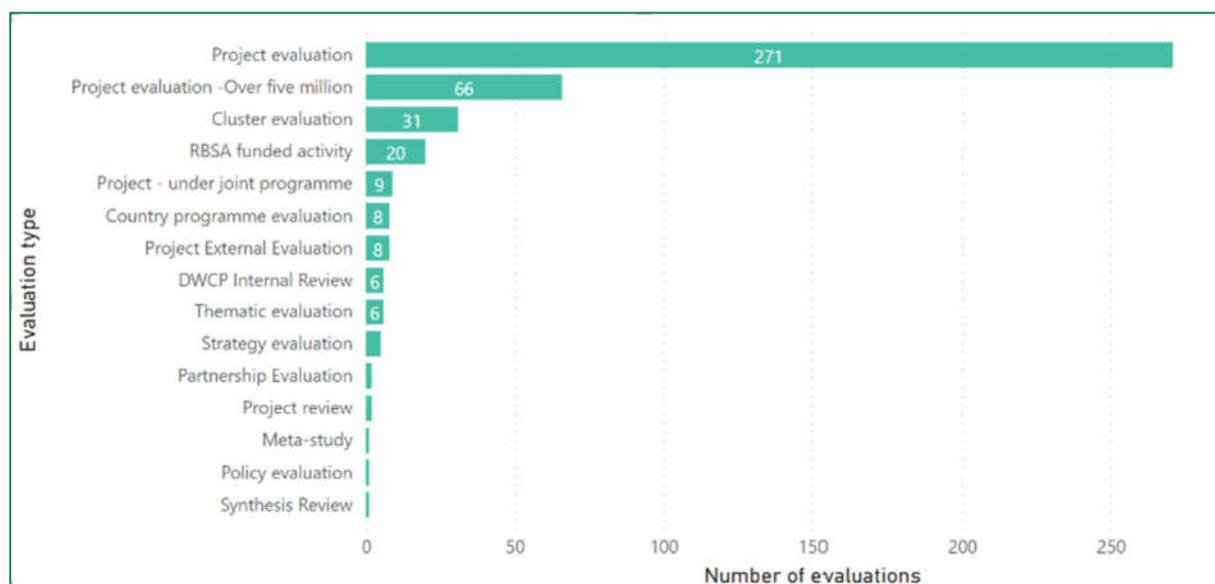
The report and annexes provide far more detail on what is working well, what needs strengthening and opportunities for change. A 'strength of evidence' rating is given against each finding, the detail of which is available as needed.

1. Impact: Evaluation is helping the ILO to shape its strategies and operations but is constrained by the prevailing operating environment - including financial and staffing resources, limited timeframes and the high number of evaluations being conducted. There is scope for better use of evaluative evidence to drive strategic decision-making that could also meet the demands of constituents, staff and partners for more evaluative evidence and learning on ILO's impact.

2. Effectiveness: The evaluation function has achieved or partially achieved 95 per cent of the sub-outcomes set out in its evaluation strategy. Limited resources and inconsistent engagement in evaluation from across the organization constrains the full achievement of targets. EVAL excels at producing regular, timely and data-rich reporting on evaluation performance, evaluations undertaken, findings and recommendations which is highly valued both internally and externally. The current portfolio of evaluations is dominated by project evaluations, this volume inevitably makes demands on the entire evaluation function in terms of managing, quality assuring, and more widely on the ILO in absorbing and using the evidence generated.

3. Efficiency and sustainability: ILO's structural arrangements give weight to the evaluation function and mainstream evaluation across the organization and across regions. However, the efficiency, sustainability, and adequacy of the evaluation structure is threatened by limited investment, organizational incentives and buy-in to evaluation as a source of knowledge and learning. Decentralized evaluations, which comprise most evaluations in the ILO are valued by stakeholders, but the workload on the network, especially the regional evaluation officers, is unsustainable. While there is an increase in and high demand for cluster evaluations, there is a need to further optimise their use. Developing an agreed pipeline of cluster evaluations could improve efficiency.

4. Relevance: The ILO's evaluation function is meeting the expected norms and standards for credibility, quality and independence but has not yet fully attained expectations for utility. EVAL is an active and valued member of the UN evaluation community; there is scope for more joint evaluations and harmonised working. The ILO would benefit from a sharper strategy (including in the choices of what to evaluate) to meet the varied tripartite evidence needs and harness their interest in evaluation evidence. Better ways to involve

constituents and beneficiaries should be explored to benefit from their insights. Evaluations have enhanced efforts to address attention to gender and disability, but full integration relies on greater sensitivity to gender, disability, environmental sustainability, and other cross cutting issues at the project design stage.

5. Adaptability: The evaluation strategy provides a clear strategic vision and outcome areas that have potential to support a transformative approach. The policy guidelines provide a consistent, but less nimble approach to evaluation coverage and prioritization. Guidance developed by EVAL[4] helped the ILO evaluate the effectiveness of its response to the pandemic in the face of significant challenges for delivery, monitoring, and evaluation of planned programmes. The Theory of Change set out in the evaluation strategy does not show clear pathways to a stronger evaluative culture or ways to maximise the impact of evaluative evidence.

6. Coherence: Evaluations broadly meet the accountability and information requirements of key stakeholders but struggle to tell the story of ILO's specific unique contributions to change and to impact. Meta-analyses and synthesis reviews are enabling links to be made across multiple sources of evidence and lessons to be drawn from complementary efforts. The evaluation function needs to move beyond the solid foundation it has built on accountability to embrace the unique learning and improvement potential that evaluation can bring for the ILO.

The Deep Dives yielded several additional insights:

Evaluation Culture

The quality, professionalism, and independence of evaluation is well-recognized and endorsed by ILO staff and stakeholders. There is an overall trust in the integrity and credibility of the evaluation process and products. There is a strong appetite in the ILO to use evaluations to assess impact and a shared commitment to compliance with evaluation responsibilities. The evaluation function benefits from a structure that embeds senior management and leaders in the response to and actions arising from evaluations as well as tripartite constituents in the design and review of evaluations. Mainly due to a constrained enabling environment, time constraints of staff, as well as the sheer volume of evaluations, there is scope to improve the use of evaluations for learning purposes. While independent evaluation for accountability purposes is widely recognised, there is limited evidence of a demand for evaluative knowledge-sharing between different departments and programmes, or highlighting lessons learnt on what has *not* worked. These are critical for a healthy evaluative culture that values transparency, lesson-learning and continuous improvement.

The UN Landscape

The UN Evaluation norms and standards give primacy to behavioural independence and organizational independence of the central evaluation function and in this regard, the ILO's functioning independence, impartiality principles and practices are strong. The ILO has a strong reputation for the high degree of independence of its evaluation function - operationally and financially; for the quality of its evaluations; its adherence to UN policies, procedures, and best practices. However, the requirements set for conducting a large number of evaluations place significant burdens on staff and key stakeholders, which in turn impacts their utility to make use of findings in strategic decision-making and lesson-learning. Some UN agencies take a more flexible and context-specific, criteria-based approach to evaluations without compromising accountability or independence requirements. With

---

[4] See for example: https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_757541.pdf and https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_824659.pdf

adaptations to the threshold, criteria, and scope of independent evaluations (evident in some other UN agencies), the workload of already-stretched ILO evaluation staff could be reduced, leaving more time and capacity for evaluative learning and uptake of recommendations. Despite policy commitments and the development of guidance and more attention paid to gender in quality appraisals of evaluation TORs and reports, it is not yet fully mainstreamed throughout evaluation practice and the ILO's evaluation function falls short of meeting UN-SWAP requirements (UN WOMEN, 2020). In 2020 only four evaluations (out of 46) were quality assessed as meeting the UN-SWAP requirement.

Funding Partner Engagement

ILO's key bilateral and multi-lateral funding partners as well as joint programmes funded through pooled financing arrangements reveal a mix of evaluation, monitoring, and reporting requirements. There are at times divergent views on when and why an independent evaluation should be triggered. While the ILO understandably retains the right to maintain its own standards and policies for evaluation as part of its oversight, and donors at times require separate evaluations to meet their own accountability requirements, there is a risk of duplication and inefficient use of time and resources. For some funding partners, their requirements are not always determined by the size of budget, in contrast with the ILO's policy. Funding partners welcome collaboration and opportunities for strategic coherence in their support for Decent Work. They generally have a high regard for the ILO's professionalism in conducting independent evaluations. They indicated they would trust ILO to advise them of opportunities for streamlining and clustering evaluations - particularly where they could yield valuable insights on the impact of ILO's unique mandate - including social dialogue, tripartism, labour norms and standards.

Progress made since the last IEE 2016

Of the 13 recommendations made in the last IEE 2016, good progress has been made on nine of the recommendations (see report section 4), including strengthening use of evaluability assessments, inclusion of monitoring and evaluation specialists in decent work teams and projects, expansion of the quality assurance system, HLEs and cluster evaluations, introduction of automated recommendations tracker, strengthened knowledge management and dissemination of evaluation results.

However, less progress has been achieved on four recommendations[5] including the transitioning of REOs to EVAL, promotion of gender responsive and participatory evaluation, and a more flexible and diverse portfolio of evaluations. Barriers to full implementation mostly arise in the lack of enabling environment concerning evaluation culture (lack of organisational incentives for participation in evaluation, overreliance on voluntary evaluation management), financial barriers (preventing more flexible funding or control of funding for evaluations), but some are linked to evaluation process and procedures (lack of progress on gender responsive evaluation).

Conclusions

The ILO has a well-developed evaluation function, embedded, and highly regarded with great potential to ensure the organization's decision making, programming design and implementation are underpinned by useful robust evidence. There are some barriers, challenges, and missed opportunities that need addressing in the next strategy period to enhance the value of evaluation, prioritize the evaluation resources, and improve the culture of evaluation and its critical role in learning and continuous improvement in the ILO. There

---

[5] Specifically recommendations 2, 6, 8, 11

are opportunities to strengthen and take the evaluation function to the next level of impact and maturity.

The evidence gathered for this evaluation reveals a system built on a solid strategy, clear mandate, and strong operational processes. It has a degree of fragility given the weighty burden on EVAL staff, Regional Evaluation Officers, Evaluation Managers, external evaluation consultants, and Departmental Evaluation Focal Points to deliver its ambitions.

1. The ILO has a mature, independent evaluation function, with embedded (institutionalised) systems and processes with recognised quality and credibility. EVAL and the REOs strengthened communications and knowledge management efforts are critical and widely recognized.

2. Although EVAL has supported extensive learning exercises, the accountability role of evaluations is more visible than learning and improvement. This has influenced the evaluation culture with many participants of this IEE viewing evaluation as a compliance exercise. There is frustration about the limited time and opportunity to use evaluations and engage in evaluative practice as a vehicle for critical analysis and continuous improvement. Evaluations are not always asking the right questions that could provide valuable insights on what is working well, less well, and why. A strengthened evaluation function focused on strategic lesson-learning will enhance the delivery and effectiveness of ILO's strategy and mandate.

3. There is an expressed demand from the ILO and amongst constituents for evaluation evidence on impact. This could be better harnessed, particularly as there is an appetite for evidence on the story of impact, understanding pathways to impact, and seeing what ILO's unique contribution is to attain that impact. Assessment of impact requires robust baseline, integration with intervention design, monitoring, follow-up, and adequate funding (sometimes beyond the life of the intervention).

4. Despite the perseverance of EVAL and the evaluation network to forge ahead with a transformative ambition (e.g., through the delivery of 30 cluster evaluations during the period of review) this is constrained by barriers in the enabling environment and its own policy and operational guidelines. The volume of evaluations jeopardises the usefulness, learning generated, and crucially the capacity to engage in evaluation results and use them to improve design and decision making. The policy promotes a 'production line' of evaluations which presents risks to the capacity of ILO to ensure high quality useful evaluations. A proportionate approach to evaluation at all levels to ensure optimal use of limited resources is needed, with evaluations that can provide sufficiently robust answers to the questions that are most important to the ILO and can generate the most learning. The evaluation staffing resources are not adequate to meet the demands of the current evaluation strategy and policy or ILO's knowledge management ambition, furthermore the financial system and funding mechanism are too rigid to allow for pooled funding for smooth implementation of portfolio or cluster evaluations, for evaluation funds to be spent once a project has ended (ex-post evaluations, or longitudinal studies), or cost recovery mechanisms to '*payback*' or incentivise home departments of evaluation managers who give up their time to oversee evaluations as called for in the evaluation policy.

5. Independence and impartiality are a key strength and a success for the ILO evaluation function but there are opportunities for a more practical and nuanced approach to upholding independence in the evaluation process. This could involve greater participation of policy departments and field offices to ultimately enhance the relevance, utility and ownership of evaluation processes and results. Harnessing the evaluation skills and knowledge that exist across the entire evaluation function to shape the conversation on effectiveness and performance at the design stage can assist organisational learning. The reliance on EVAL staff for internal real-time quality assurance not only diverts

evaluation resource away from delivery of strategic evaluations, but also risks timeliness of evaluation products.

6. Evaluation is seen as a potentially powerful tool to foster tripartism, social dialogue and build capacity amongst all constituents and intended beneficiaries. Opportunities for engagement of constituents in the identification of questions and to absorb and act on the evidence generated are not being maximised. The length of reports, number of standardised questions (not tailored to their needs and interests), and a high degree of reliance on eliciting written commentary on documents (Terms of Reference, Draft Reports etc.) does not encourage participation of constituents.

Recommendations

The following nine recommendations are presented in brief (the full detailed recommendations are in section six of the report). They arise from the findings and conclusions and have been informed by consultations with ILO staff through the validation and commentary processes.

1. **Maintain EVAL's independence** through its organizational location and make more strategic use of the EAC as an influential body. Use the time of EAC members more strategically to bring sharper discussion of evaluation insights and implications arising.

2. **Use evaluation resources more strategically**, stringently prioritize and plan whilst maintaining commitment to improve results-based management system. Reduce the number of evaluations conducted for smaller projects or where evidence is already strong and ensure ability and agility to direct resources to areas of least evidence, risk, or innovation.

3. **Strengthen the enabling environment**: Resourcing: Ensure adequate financial and human resourcing in terms of level and modality for evaluation to deliver a transformative evaluation ambition. Ensure that the ILO's financial process and mechanisms allow for ease of creation of pooled funds, ex-post evaluation funding, and cluster evaluations.

4. **Strengthen the enabling environment: Invest in building the capabilities and skills** to undertake, manage, and use evaluations. Professionalize, incentivize, and recognize evaluation skills.

5. **Strengthen the enabling environment: Foster a stronger culture of learning.** Promote a culture of collaboration, continuous learning, professional development through evaluation design, delivery and follow up. Enable discussions of failure – what is going wrong and help to understand why. Invest in forging stronger links between monitoring, learning and evaluation that enable evidence-based decision-making, incentivize honest and open discussion on what is working less well, allow for adaptation and continuous learning.

6. **Ensure quality, credibility, and utility of evaluations.** Improve the design of evaluations to focus only on the most relevant issues. Improve the quality of evaluation recommendations with a greater utility focus. Expand the real-time quality assurance of HLEs independent from EVAL.

7. **Invest in knowledge capture and information management systems.** Enable more synergy between evaluations and other learning products such as project reporting, research, and other knowledge products. Introduce an accounting code for evaluation to enable EVAL to track and analyze spend on evaluations. Ensure all project evaluations are accessible on i-EVAL Discovery, allow for better tagging of evaluations to allow search by ILO strategic objectives and results.

8. **Continue and expand ILO's commitment to the UN evaluation system.** Continue to share ILO's evaluation expertise and knowledge with the UN evaluation network and

other UN organizations. To enhance coherence, minimize overlaps and avoid overburdening stakeholders, the ILO should seek opportunities for collaboration and coordination among UNEG members, when appropriate and feasible.

9. **Build on strong foundations to develop a future-fit evaluation strategy and policy**. Update policy guidelines related to prioritization recommendations, quality assurance and independence mechanisms. EVAL should develop its strategic approach to Impact Evaluation – EVAL & EAC should have a stronger voice in determining IEs undertaken as well as project/program leads and funding partners.

Lessons learnt

1. **Internal and external coherence:** The evaluation highlighted that the Evaluation Function is independent but is intrinsically part of a system. The enabling environment, culture of learning and of evaluation, and the quality of programming relies on strong coherence, dialogue, and collaboration at all levels. Coherence across the evaluation function and beyond can help to strengthen ownership and participation of stakeholders, evaluability of programmes, and the opportunity to ensure that evaluations provide useful findings, recommendations, and lessons. Ensuring opportunities to foster dialogue is key to delivering policy goals and to effective and efficient systems.

2. **Integration of gender and cross-cutting issues into all stages of programming and evaluation:** The evaluation identified opportunities to more systematically mainstream gender and cross-cutting issues in all evaluation stages and deliverables. Greater attention, and shared understanding of how these issues can be embedded in programming at all stages can strengthen the integration of ILO policies on these issues.

3. **Embedding evaluative thinking across the organization:** Fostering a culture that rewards transparency about what is working and not to generate insights on what could be done differently. Integrating evaluative thinking at all levels provides an opportunity to strengthen RBM and thinking about how activities are linked to broader mid-term and long-term goals.

4. **Financial and human resource requirements:** Effective implementation of policies and programmes requires adequate human and financial resources. Gaps in resourcing, or overloaded staff members can compromise sustainability. It is critical that policy decisions be matched with adequate human and financial resources

5. **Incentives and accountability or balancing "the carrot" and "the stick":** The evaluation pointed to the challenge of achieving a balance between incentive structures and accountability mechanisms. Institutional systems and accountability frameworks need to continually strive to incorporate both rewards and clear accountability requirements

Main report

1. Introduction

As endorsed by the Governing Body, the ILO sought an Independent Evaluation of its Evaluation Function (IEE) to be conducted in 2022. This will be the third time such an evaluation has been undertaken (first was in 2010, second in 2016 (ILO 2016). The IEE covers the period 2017–2021, with a partial assessment of the work carried out in 2022. The evaluation function is more than just the Evaluation Office of the ILO and as such the IEE investigated the interacting components of the evaluation function. The primary audience for the IEE is the **ILO Governing Body** which includes the ILO's **tripartite constituents** and is responsible for governance-level decisions based on the findings and recommendations of this evaluation. Other primary stakeholders include the **ILO Director-General** and members of the **Senior Management Team, Regional Directors, EVAL,** and the **ILO's development partners**.

This report presents the results of inception, data collection, and reporting phases undertaken by the evaluation team during the period February-July 2022.

The report is structured as follows:

- Section 1 presents the overall purpose of the evaluation, its objectives, guiding questions, approach, methods, and limitations

- Section 2 focuses on the evaluation subject and its organizational context, and the policy and strategy guiding the evaluation function

- Section 3 presents the findings, organized against the evaluation questions, and framed in terms of what is working well, and opportunities for strengthening the evaluation function further.

- Section 4 outlines progress made since the last IEE

- Sections 5 and 6 present conclusions, recommendations and lessons learned.

1.1 Purpose of the evaluation

The Independent Evaluation of the ILO's Evaluation Function (IEE) aims to serve accountability, programme improvement, and organizational learning purposes. It is intended to inform and shape the forthcoming revised Evaluation Strategy, and to support the ILO foster an evaluative culture and practice that adapts and responds to the internal and external strategic and operational challenges.

The IEE includes an examination of the current arrangements and structures of the evaluation function (involving both the centralized and decentralized evaluations), an assessment of the ILO Evaluation Office's (EVAL) performance in the delivery of its strategic goals, and recommendations in relation to the independence, credibility, utility, operational framework and enabling environment for evaluation, with lessons extracted from the implementation of the evaluation strategy during the 2017- 2021 period. The IEE has been guided by the Terms of Reference (ToR) included in Annex A that sets out its' specific aims.

1.2   Evaluation process

The evaluation was conducted in three phases: an eight-week inception phase, a 13-week implementation (full evaluation) phase, and a seven-week finalization phase. Throughout, the evaluation team engaged bi-weekly with the IEE TC and EVAL to ensure ownership of the process and track progress. The IEE TC was closely engaged in evaluation management but emphasized and respected fully the independence of this evaluation.

1.3   Evaluation framework and questions

The purpose and scope as set out in the IEE ToR informed the original formulation of the evaluation framework, guided by the UNEG norms and standards, and drawing on the ILO Evaluation Strategy, accompanying results framework, and Theory of Change. The evaluation questions were refined during the inception phase incorporating feedback received from EVAL, and the IEE TC. The evaluation questions are grouped to inform 'higher level' lines of inquiry whilst remaining faithful to the ToR requirements. Particularly, the evaluation framework focuses on aspects of the conditions that enable the evaluation function (e.g., organizational culture and resources). The key evaluation questions, which were used to organize the findings of the IEE, are provided below, and the full evaluation framework is included in in Annex B. The evaluation questions are as follows:

1. To what extent does the evaluation function help the ILO deliver against its current strategy? [Impact]

2. What is the effectiveness of the evaluation function during the strategic period? [Effectiveness]

3. How efficient, sustainable, and fit-for-purpose are the current arrangements and structures of the evaluation function? [Efficiency & Sustainability]

4. To what extent does the evaluation function respond to international, organizational, and tripartite policies, interests, and information needs? [Relevance]

5. To what extent is the evaluation function able to respond and flex to emerging priorities and stressors? [Adaptability]

6. To what extent does the evaluation function align with other accountability and learning processes? [Coherence]

1.4   Evaluation approach and methods

The evaluation approach was designed to optimize learning whilst also providing accountability and an assessment of the results and impact of the ILO's evaluation function. A mixed methods approach was implemented, i.e., multiple research methods to collect and triangulate qualitative and quantitative data from a range of sources to establish a robust evidence base. The evaluation used six main methods, including the document review and quantitative analysis (e.g., evaluation and quality assurance data, project data and expenditure data), semi-structured key informant interviews (KIIs), focus group discussions (FGDs), portfolio review, four surveys, and three thematic deep dives (exploring evaluation culture in the ILO, evolving trends in the evaluation landscape and funding partner influence on the evaluation function). The findings of the deep dives and portfolio review underpin and inform the analysis and as such their presentation is embedded throughout the findings of the report, with additional detail provided in Annexes F, G and Q.

Each KII and FGD was confidential, and it was agreed that quotes or comments would not be attributed to individuals. The interviews and focus groups provide rich qualitative insights (not always possible to ascertain from quantitative data or documentary evidence) that have provided vital insights in understanding how staff and other stakeholders experience their interaction with the ILO evaluation process and their perceptions of the evaluation culture in the ILO. All evidence streams were weighted equally, and we have drawn on qualitative interview evidence for illustrative purposes. In all data collection and analysis, we have ensured a balanced representation of the people consulted.

The evaluation framework was used to analyze data from different sources and to organize and tabulate it in relation to the evaluation questions, using analytical tools including Excel tabulation and content analysis software MaxQDA. This systematic approach to analysis ensured traceability from findings to conclusions and to recommendations. The evaluation team triangulated data gathered through documentation and quantitative datasets, interviews and FGDs, the portfolio review, deep dives, and the surveys. A range of triangulation methods were applied: cross reference of different data sources; triangulation between team members, and verification of findings and information post-data collection. The team conducted a strength of evidence review using the following criteria (level 1: a single line of evidence or weak triangulation; level 2: limited lines and levels of evidence with moderate triangulation; level 3: Multiple lines of evidence, with strong triangulation). There is more confidence in the findings where the evidence score is two or, especially three.

The evaluation methodology is detailed in Annex D and set out briefly in Fig.1.

Figure 3: Data collection methods

Document review	Interviews and consultations	Online survey	Portfolio Review	3 Deep Dives	Validation
Approx. 238 documents	116 respondents interviewed / in FGDs in total	216 responses: 106 all staff, 33 EMs/REOs/DEFPs and 9 senior managers, 69 GB members & constituents	Sample of 20 evaluations	of thematic areas including Evaluation Culture, funding partner and evolving trends in the evaluation landscape	3 workshops

1.4.1 Limitations of the evaluation

- Low response rate to the survey: the survey was sent out to 1353 staff members and was completed by 148, giving a response rate of 11%. A survey and invitation for interview was sent to 1519 current members and all former members of the Governing Body (GB) and to non-GB constituents which elicited 68 survey responses[6]. In total the surveys solicited a total of 216 responses (see Annex E for full breakdown). The team undertook mitigating actions to minimize the limitation such as issuing reminders, extending the 'window' to participate, and ensuring triangulation with other forms of evidence.

- Depth of participation was constrained by the time available for the evaluation and the limited response of some stakeholders (such as GB members), the timing of the GB in March and the ILC in June contributed to the lack of response and availability for interview. Interviews with the tripartite structure was limited to KIIs with ITUC and IOE, and funding partners as part of the deep dive (in addition to ILO staff from the bureaux of ACTRAV and ACT/EMP). GB members former and current were invited to participate in an interview but did not accept the offer despite reminders. Non-GB constituents were invited to participate via survey only. It was not within the scope of the evaluation to

---

[6] The survey was sent to all current GB members and to previous members to ensure coverage of the evaluation timeframe 2017-2021.

interview beneficiaries or constituents at regional or national levels. The IEE team was not able to secure an interview with representatives of ILO Turin (ITCILO). The evaluation team extended the timeline for interview and the data collection period and issued reminders to encourage participation.

- The evaluation was conducted virtually which limited opportunity for *in situ* observation and face-to-face engagement with ILO staff and stakeholders.

2. Evaluation in the ILO – setting the scene

2.1 Organizational context

2.1.1 Overview

The International Labour Organization (ILO) is the tripartite UN agency that brings together governments, employers, and workers of 187 Member States to set labour standards, develop policies and devise programmes promoting decent work for all. It is the only tripartite UN agency that includes governments, employers, and workers organisations with equal responsibility for setting labour standards, developing policies, and delivering programmes, based on the principle of social dialogue. The ILO carries primary responsibility for UN Sustainable Development Goal (SDG) 8 on Decent Work and contributes to other SDG targets relevant to the world of work.

2.1.2 Strategy and Planning

The period under review spans 2017-2022, covering three ILO strategic periods of 2016-2017, 2018-2021 and 2022-25[7]. In November 2020, the ILO adopted its Strategic Plan for 2022-25, which is currently underway (ILO 2020). It reaffirms commitments to strengthen work in areas where the ILO has a unique mandate - labour standards and tripartism, as well as changes in a wide range of policy issues within the world of work. The Strategic Plan also reflects commitments to improve ILO's organizational

> Of relevance to the IEE, the new strategic plan asserts:
>
> *In the context of continued strengthening of accountability, oversight and risk management, the evaluation function will remain a valuable driver of performance improvement. The focus will be on carrying out high-quality evaluations and fostering a strong evaluation culture, with increased uptake of findings and recommendations.*

performance through enhancing leadership and governance, improving knowledge capacities, increasing cooperation and partnerships, and optimizing the use of ILO's resources.

2.1.3 Resourcing[8]

The ILO is resourced through assessed contributions from member states and voluntary contributions from a range of funding partners - used to support specific global, regional, and national programmes.

---

[7] At its 320th Session (March 2014) the Governing Body decided that the ILO should have a medium-term strategic plan and that it should be aligned with the planning cycle of the United Nations (UN) system. After approving a transitional strategic framework for 2016–17 to bring about that alignment, the Governing Body at its 328th Session (October 2016) adopted the ILO's Strategic Plan for 2018–21. Strategic Plan 2022-2025 P 5

[8] All figures quoted in this section 2.1.3 are from the Programme Implementation report for the 344th Session of the Governing Body ILO 2022 (unless otherwise stated)

The ILO regular budget for 2020–21 was US$790.6 million, of which US$496.6 million was earmarked for the delivery of policy outcomes. In addition, the ILO estimated that extrabudgetary expenditure in the biennium would amount to US$500 million, of which US$30 million would come from the regular budget supplementary account (RBSA). The total actual regular budget expenditure over the biennium was US$749.8 million, 95 per cent of the budget, and extra-budgetary expenditure was US$667.7 million (US$31.6 million from the RBSA), 33 per cent over the estimated amount.

The operational budget for Oversight and Evaluation (Evaluation Office EVAL only) in 20-21 amounted to $8.4m. Of this, roughly $3m was allocated to the Evaluation Office (EVAL), $3.4m went to Internal Audit and Oversight. $620,300 supported the Independent Oversight Advisory Committee; $1m was allocated to External Audit Costs and $300k for the Ethics function. Section 3.3.4 sets out more detail on the budgets dedicated to evaluation.

In summary, the approved figures for 2020-21[9], the ILO's development cooperation support comes from:

- Multi-Bilateral Funding partners, of which the European Union, USA, Netherlands, Germany, and Japan are the largest funding partners.

- Core voluntary funds (RBSA)

- International Financial Institutions

- Direct Trust Funds

- Private / Non-State Actors

- The ILO's interactive Development Cooperation Dashboard [10]sets out the key sources of funding and areas of programming.

2.1.4 Programming and the Results Based Management system

The Strategic Programming and Management Department (PROGRAM) advises the Director-General and assists the ILO management and staff on programme planning, resource allocation and implementation reporting. It operates on principles of results-based management and budgeting.

The ILO's results-based management (RBM) system applies to all stages of ILO's programming cycle including planning, implementation, reporting and evaluation, also shown in Figure 2. In line with the RBM principles and ILO's evaluation policy, all ILO strategies, policies, and programmes, including Decent Work Country Programmes, technical cooperation programmes and projects are subject to evaluation (ILO 2009).

The MOPAN assessment of ILO highlighted a "strong results culture and RBM approach in policy dialogue, planning and implementation in ILO" (MOPAN 2021, 13). The report goes on to say that "it also has a robust and quality-focused evaluation function with the necessary policies and mechanisms in place, while evaluation skills and culture are being strengthened across the organization. However, there remains ample room to improve the integration of RBM across the organization". The new ILO Strategic Plan (2022-25) (ILO 2020) aims to enhance existing approaches, guides, and tools to optimize the use of resources and achievement of results as well as to improve accountability, communication, decision making, and learning.

---

[9]   Top   20   contributors   to   the   ILO's   voluntary   funding   (2020-2021*) https://www.ilo.org/wcmsp5/groups/public/@dgreports/@exrel/documents/image/wcms_701787.pdf
[10] Development Cooperation Dashboard
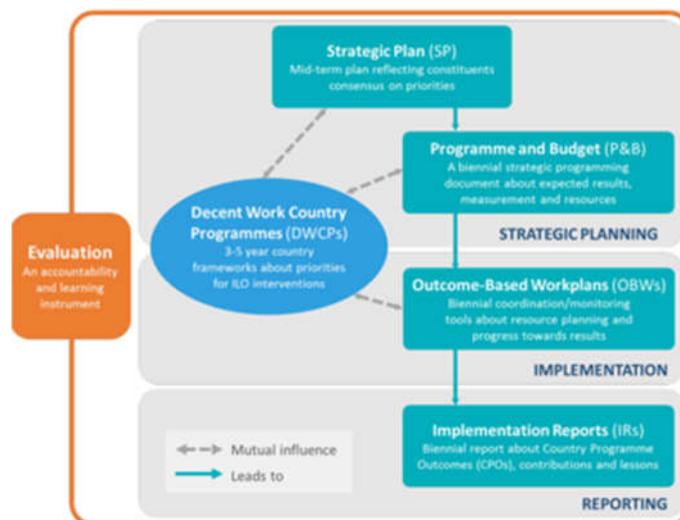
2.2 The ILO Evaluation Strategy and Policy[11]

The aim of evaluation in the ILO is to support improvements in programmes and policies, and to promote accountability and learning (EVAL 2020)[12]. All aspects of evaluation in the ILO are guided by the ILO evaluation policy and the ILO Evaluation Strategy 2018–21 which adhere to the OECD/DAC Principles (OECD/DAC 2019) and UNEG Norms and Standards for Evaluation (UNEG 2016). These two key governance-level documents define the ILO's organizational approach and results-based framework for evaluation. The evaluation strategy is operationalized in the context of the ILO's strategic plan and the biennial programme and budgets (P&Bs). The strategy is a time-bound evaluation strategy, aligned with the ILO's 2018–21 Strategic Plan (ILO 2016) that identifies outcomes and targets through which the Evaluation Policy is gradually rolled out.

2.2.1 The ILO Evaluation Policy

Since 2005, the ILO has had an evaluation policy which has, periodically, been reviewed and updated. The last revision took place in 2017 to address the recommendations of the IEE from 2016 and for an increased adherence to the updated UNEG Norms and Standards (UNEG 2016). The recommendations of the IEE 2016 were also used to inform the current evaluation strategy for the period 2018–2021. Both documents were developed through an extensive consultation process with ILO staff and constituents, using online surveys, focus-group discussions, and interviews (EVAL 2021).

The 2017 Evaluation Policy (ILO 2017) states a desire for a shared vision for the evaluation culture within the ILO: "A sustained, expanding institutional culture of mutual accountability, ownership, transparency, and quality improvement is a strong vision shared by the ILO Governing Body and the Office. An evaluation culture to use evaluation for better

---

[11] The full title of the strategy is: ILO results-based Evaluation Strategy 2018–21, but will be referred to at the evaluation strategy for brevity

[12] The IEE team interpret learning as incorporating adaptive planning and management of interventions, policies, and strategies.

performance, effectiveness, and learning, in the pursuit of the Decent Work Agenda, is at the core of this commitment".

The 2017 policy aims to:

- Reinforce knowledge generation and sharing of the ILO's substantive work, and the processes, approaches, and institutional arrangements for implementing such work
- Strengthen the complementarity between evaluation and other oversight and monitoring functions within the Office
- Clarify standards for engaging constituents in evaluation
- Clarify the division of responsibilities in the ILO for carrying out an evaluation

The policy sets out the types of evaluation that are expected to be conducted, roles, responsibilities (including the structure and role of EVAL), financing and resources, use of evaluation, and capacity development. The guidance document 'ILO policy guidelines for results-based evaluation: Principles, rationale, planning and managing for evaluations 2020' sets out in detail how the policy should be operationalized, including mandatory requirements for project evaluations (see table 1 below).

Table 1: Extract from ILO Evaluation Policy Guidance ILO Policy Requirements for Project Evaluations (EVAL, 2020)

Project budget (US$)	Under 18 months	18-30 months	Over 30 months
Over 5 million	Initial M&E appraisal by EVAL. M&E plan appraisal tool, Independent Mid-term and Final independent. Mandatory: evaluability assessment	Initial M&E appraisal by EVAL. Independent Mid-term and Final independent. Mandatory: evaluability assessment	Initial M&E appraisal by EVAL. Independent Mid-term and Final independent. Mandatory: evaluability assessment
1-5 million	Final independent evaluation	Mid-term (self or internal evaluation) & final independent evaluation	Mid-term (self or internal evaluation) & final independent evaluation
500 000-1 million	Final internal evaluation	Final internal evaluation	Mid-term internal evaluation & final internal evaluation
Under 500 000	Final self-evaluation (part of final project progress report)	Final self-evaluation (part of final project progress report)	Mid-term self-evaluation (part of a project progress report) and final self-evaluation (part of a final project progress report)

2.2.2 The ILO Evaluation Strategy

The ILO Evaluation Strategy 2018 – 21 operationalizes the evaluation policy. It includes a theory of change (ToC) that guides the work of the evaluation function in the ILO with three expected outcomes:

- Outcome 1. Enhanced capacities and systems of evaluation for better practice and use
- Outcome 2. Enhanced value of evaluation using more credible and higher-quality evaluations (independence, credibility, usefulness)
- Outcome 3. Stronger knowledge base of evaluation findings and recommendations

The three outcomes reflect the expectation that the evaluation function will have a more significant influence on the decision-making, credibility, visibility, and performance of the

Organization. These three outcomes are to be achieved by focusing on 13 sub-outcomes and 19 biennial milestones, underpinned by an enabling environment for evaluation at the ILO. The evaluation strategy incorporates the key guiding principles of the ILO Strategic Policy Framework (ILO 2009), which calls upon evaluation to strengthen knowledge development and accountability in the areas of Decent Work, international labour rights and standards and the ILO Declaration on Social Justice for a Fair Globalization, the ILO Centenary Declaration for the Future of Work as well as to enhance the relevance and utility of evaluation to constituents (EVAL 2020).

The evaluation strategy's ToC is presented along two streams: (a) the activities of the Evaluation Office (EVAL) as they feed into an evaluation function; and (b) an organization-wide set of outcomes and impacts within a culture of more comprehensive evaluation in support of the ILO's mandate (see Annex O). As presented in the ToC, the evaluation function seeks to contribute to an effective and efficient ILO, through the delivery of decent work policies and programmes, to enhance the credibility and leadership of the ILO through an embedded accountability, transparency, and evaluation culture; to advance decent work goals by leveraging national and international partnerships to measure the contribution to the SDGs. It is supported by the 2018-2019 Programme and Budget Enabling Outcome B.5. Section 3.5.2 of the IEE presents the IEE's views on the ToC.

2.3 The Evaluation Function

The components of the evaluation function in the ILO include:

- Evaluation policy (2017)

- Evaluation strategy (2018-21)

- The centralized evaluation function including:

    o ILO independent Evaluation Office (EVAL) of seven staff including the Director

    o Evaluation Advisory Committee (EAC)

    o The systems, structures, and evaluations within EVAL's remit (e.g., quality assurance, knowledge management, communications)

- The decentralized evaluation function including:

    o A network of five Regional Evaluation Officers (REOs), and non-evaluation staff acting as 14 Departmental Evaluation Focal Points (DEFPs) and 144 certified Evaluation Managers amongst ILO staff and on a voluntary basis

    o Evaluations of programmes and projects under the control of programme, technical and field units

In addition, there are 48 monitoring and evaluation officers[13][i] spread across country offices and the Policy Portfolio – these roles are primarily designated to specific programmes involving multiple partners and specific funding partners[14], including the ILO's five 'flagship' programmes, these are: Better Work; Jobs for Peace and Resilience Fund; Social Protection Floor, Occupational Health, and Safety, IPEC+ (focused on forced and child labour). A roster

---

[13] For brevity the monitoring and evaluation officers will be referred to as M&E officers, these include monitoring, evaluation and learning officers, monitoring and evaluation officers, monitoring, evaluation, and knowledge management officers. Source: ILO HRD Staff list, May 2022.

[14] The IEE notes that parts of ILO refer to "funding partners" rather than donors. For the purposes of the IEE, these terms are used interchangeably although the emphasis on collaboration in the ILO's terminology is well noted.

of external evaluation consultants supports the evaluation function. Figure 3 charts out the general structure of the evaluation function explained above.

Figure 5: Architecture of the Evaluation Function
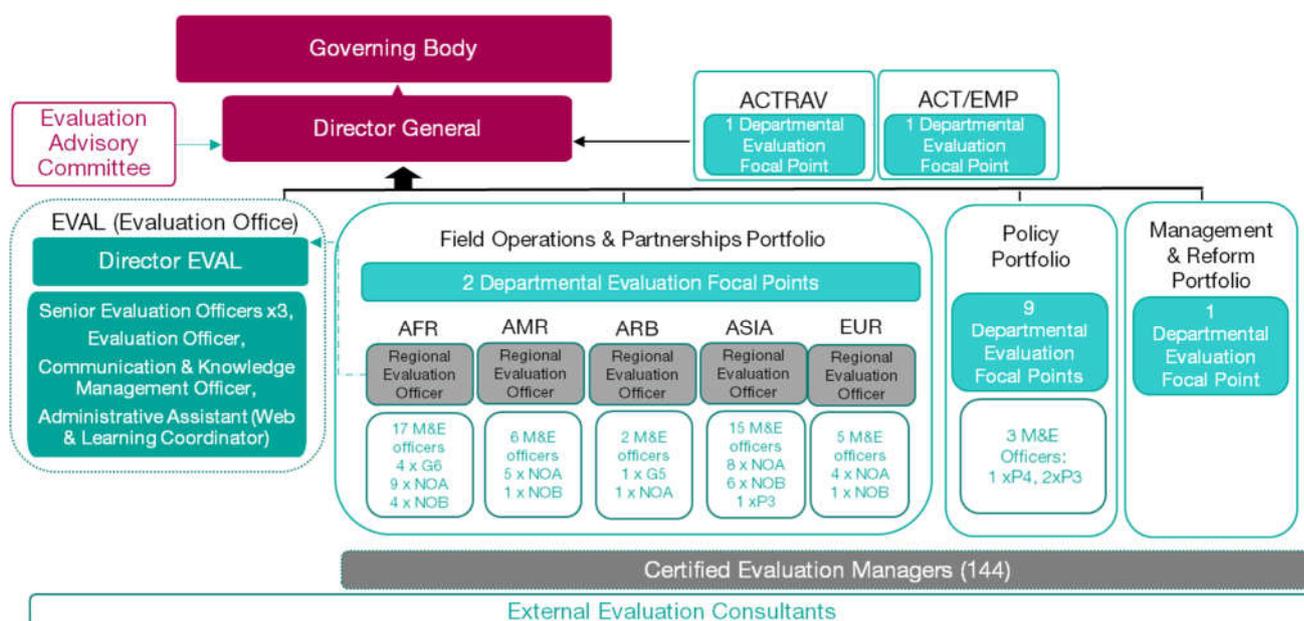
2.3.1 EVAL

EVAL is an independent Office, reporting directly to the ILO Director-General and the Governing Body, and is the heart of the evaluation function in the ILO. EVAL's structure and modalities of operation are designed to protect its independence through dual reporting lines. EVAL is mandated to manage the evaluation function and to ensure implementation of the evaluation policy.

The Evaluation Office had its origins in the central Evaluation Unit (also known as EVAL), which was established in 2005, as part of the ILO's management and administration structure. In 2011, based on the recommendations of the 2010 IEE, EVAL was placed directly under the DG's Office in the ILO's organizational structure. In 2014, as part of the DG's reform agenda, it was renamed the ILO's Evaluation Office. EVAL comprises: the Director, three Senior Evaluation Officers, an Evaluation Officer, a Communications and Knowledge Management Officer and Administrative Assistant (Web and Learning Coordinator).

EVAL is responsible for selecting and undertaking High Level Evaluations (HLEs) based on consultations and reporting on their findings to the Director General and the Governing Body. The EAC oversees the implementation of follow-up to recommendations of HLEs and the related management responses and EVAL provides a summary of their deliberations to the GB through the Annual Evaluation Report (AER). In terms of activities, the evaluation function is organized around nine inter-related broad groups, (Annex I provides a mapping of activities undertaken by EVAL against the strategy outcomes):

- Production and management of evaluations
- Oversight to ensure quality and relevance
- Institutional accountability
- Knowledge management and communications
- Capacity development
- Engagement in international evaluation for a
- Wider Network: Decentralized Evaluation Staff and Focal Points.

- Organization learning
- Results-based management

2.3.2 Decentralized Evaluation

The decentralized evaluation structure is supported by a network of certified evaluation managers, regional evaluation Officers (REOs), departmental evaluation focal points (DEFPs), with EVAL providing oversight and quality control over the whole cycle with final approval authority and accountability for the quality of the evaluation products.

At the regional level, the network comprises full-time designated REOs to support the planning and implementation of evaluation activities at regional and country levels: five REOs report to regional management, with a technical reporting line to the Director of EVAL. For the implementation of decentralized evaluations, the REOs identify evaluation managers within the ILO[15]. At the departmental level, the network comprises designated departmental evaluation focal points (DEFPs) to provide support, in addition to their other normal tasks, to the planning and implementation of evaluation activities.

2.3.3 The Evaluation Advisory Committee

The EAC reports directly to the DG's Office and is made up of ten members: the directors of PROGRAM and PARDEV, two directors of technical departments, regional directors, and one official from the DG's Office. Its chair is the Deputy Director General (DDG) for Management and Reform. The Director of EVAL serves as the secretariat of the EAC. The EAC was established to provide a mechanism for overseeing the use and implementation of follow-up to lessons learned and recommendations resulting from the ILO's evaluation activities. Its objective is to promote institutional follow-up on independent evaluation findings and recommendations. The Committee is also designed to function as a forum for internal dialogue on the implementation of the ILO evaluation policy and strategy and, in particular, to ensure that evaluations are credible and conducted in an impartial and independent way (ILO 2006).

2.4 The evaluation portfolio

Since 2005, the ILO evaluation function has incorporated a combination of centralized (governance-level) and decentralized evaluation responsibilities. Independent strategy and Decent Work Country Programme (DWCP) evaluations are governance-level centralized evaluations, which are conducted or managed directly by EVAL. EVAL also undertakes thematic evaluations whenever there is a substantive request.

High-level evaluations, undertaken for the Governing Body, aim to generate insights related to organizational performance in the context of the RBM system. These contribute to high-level decision-making on policies, strategies, and accountability. The ILO's senior management and the Governing Body participate in the process of identifying priorities for evaluation, as well as determining the timing and intended use(s) for each evaluation. To this end, an annual process of informal consultations on the selection of topics for high-level strategic evaluations and their terms of reference (ToR) is organized. A rolling three-year evaluation programme of work with proposed HLE topics is then presented to the Governing Body for its endorsement.

Decentralized evaluations include those of projects and programmes that are managed through departments and regions. Evaluation management is provided by certified evaluation

---

[15] The availability of evaluation managers varies between regions. In some regions the REO manages the majority or all evaluations.

managers, with EVAL providing oversight and quality control. Decentralized evaluations may include thematic evaluations, project evaluations, Regular Budget Supplementary Allocation (RBSA) evaluations, impact evaluations[16], and joint evaluations, as well as all forms of internal evaluations, including self-evaluations. The objects of these evaluations include development cooperation projects, country programmes, and reviews of technical interventions from all sources of funding, including from RBSA and Regular Budget Technical Cooperation (RBTC). The RBSA modality provides 4.6 per cent for monitoring and evaluation, mostly allocated to the regions and supervised by EVAL and PROGRAM. Fully unearmarked voluntary funding as well as light-earmarked voluntary funding enable the ILO to link the funding closely to its strategic priorities.

Between 2017 and 2021, 429 evaluations were conducted across the ILO, section 3.2 sets out an overview of this portfolio[17] .

3. Findings

3.1 To what extent does the evaluation function help the ILO deliver against its current strategy?

> **Key finding 1: Evaluation is helping the ILO to shape its strategies and operations, but the use of evaluative evidence is not systematically embedded in programming. There is a need to harness the demand that exists within the ILO and its constituents for evaluative evidence and learning, especially on ILO's impact.**
>
> Strength of evidence: Level 3[18]

3.1.1 The extent to which evaluation supports and informs the ILO's strategic direction, policy, and programming

The ILO's Strategic Plan 2022-25 states that in the context of continued strengthening of governance, accountability, oversight and risk management, the evaluation function will remain a valuable driver of performance improvement. It calls for a focus on carrying out high-quality evaluations and fostering a strong evaluation culture, with increased uptake of findings and recommendations. It highlights the importance of improving knowledge capacities and cites the recommendations from the high-level evaluation of ILO Research and Knowledge Management Strategies to improving how the ILO generates, shares and brokers knowledge. It calls for innovation in working methods to promote teamwork, knowledge-sharing and internal communications as a means of optimizing the use of ILO resources. The ILO's vision in this Strategic Plan is to strengthen its governance mechanisms and capacity to deliver quality services that meet the needs of its constituents, based on their expressed guidance and the availability of objective evidence.

Evaluations can play a critical role in evidence-based decision-making - particularly where decision-makers are consulted on the questions evaluations are seeking to answer, and where they value and trust the insights and information that emerge from evaluations. This IEE investigated the extent to which the right kind of evidence is being generated and used.

---

[16] EVAL provides ex-post quality control over impact evaluations – not oversight which is during the process of planning and implementing evaluations.

[17] It was not possible to undertake an analysis on the types of evaluation approaches, methods or designs used as this data is not captured on i-EVAL Discovery.

[18] Based upon strength of evidence review: 1: a single line of evidence or weak triangulation; 2: limited lines and levels of evidence with moderate triangulation; 3: Multiple lines of evidence, with strong triangulation

What is working well

The demand for evidence: There is appetite for evaluations that can measure the impact of ILO's work from within the ILO, from external stakeholders, and across the tripartite structure. Stakeholder representatives recognized the importance of evaluations in measuring and improving performance and value the useful products of the evaluation function, in particular when preparing for ILO Governing Body discussions. The work of EVAL to distil lessons was widely recognized positively by evaluation respondents.

Evidence generation and uptake: Evaluative evidence is being generated across the ILO at all levels from project evaluations to country evaluations, through to thematic and strategic evaluations. There is comprehensive coverage of projects (via the policy guidelines), and non-project strategic level work of ILO covered by the HLE annual planning and consultation process with GB members. The evaluation function also has mechanisms and processes for ensuring that evaluative findings bring about change - this is particularly through the discussion of HLEs by EAC and GB. Respondents demonstrated they understood the necessity and potential influence of the EAC. Targets set in the Evaluation Strategy for four meetings per year with discussions on 75 per cent of high-level evaluations have been exceeded.[19]

During the review period, EVAL has steered an initiative to maximize uptake of results via the introduction of mandatory evaluability assessments (EAs), (since the GB session in November 2017, EAs are mandatory for projects above USD 5 million). The ILO created an evaluability diagnostic tool to assess the evaluability of Decent Work Country Programmes in the context of the SDGs and aligned to the UNCF requirements, the instrument was piloted, and collaboration was held with PROGRAM to embed the tool into the DWCP design, and mid-term reviews. Despite a focus on evaluability, EVAL has received a low number of EAs completed in the reporting period. In some cases, diagnostic tools were used as guidance rather than being applied in full.

Evidence of evaluations making a difference: Evaluations have contributed to substantive policy change and funding decisions. There are multiple examples of evaluations that have put forward evidence of what has worked, and the factors that have supported and constrained performance[20]. According the EVAL's AER, seven out of nine high-level evaluations (77 per cent) were reflected in strategic guidance documents for the Office during the Evaluation Strategy's implementation period. Decentralized evaluations are often used by project and programme managers, and by funding partners to make decisions about programme extensions.

Acting on evaluation recommendations: Steps have been taken to enhance management response to evaluation recommendations. To track the extent to which recommendations are taken up, EVAL has strengthened its tracking of management responses to recommendations through the automated management response system (AMRS) launched in 2018. This enables a detailed and insightful analysis of recommendation follow-up and the scope, timeline, and resource implications of recommendations. According to the 2020-21

---

[19] Average 81 per cent over the period under review.

[20] High-level independent evaluation of ILO's strategy and action for promoting fair and effectives labour migration policies, 2016-20 is frequently cited as a good example of an evaluation that was able to put forward evidence not just on what had worked but why and how. The Sustainable enterprises HLE (2020) was used to inform the Programme and Budget (P&B) proposals for 2022-23, input into a new strategy on sustainable enterprises, and inform the work plan of the Green Jobs Programme. Another HLE that had an impact according to the ILO was the HLE Research and Knowledge Management (2020) which has helped to shape P&B proposals noting that the HLE provides important guidance on how to strengthen knowledge management. HLE capacity Development (2019) as evidenced by the ILO institution wide strategy and HLE GEM (2021) (in the portfolio) as evidenced by the new APGE are other examples, that shows a degree of utility and use of evaluations for enhancing organisational effectiveness.

AER, progress in this area has been partially achieved. While 93 per cent of required management responses to independent evaluations were achieved, a total of 32 per cent of recommendations have been completed, and 46 per cent have been partially completed by management during the reporting period. In the case of HLEs, reporting on implementation status of recommendations is mandatory. During the Strategy period, EVAL has proactively ensured strategic discussions with the EAC on the response to recommendations from the HLEs.

Opportunities for strengthening

Evaluating impact: There was consensus amongst respondents that ILO evaluations focus on showing relevance and effectiveness of the ILO's work and that it is more difficult to measure, or attribute impact. The normative nature of the ILO's work, short project timeframes, and the limited availability of longer-term monitoring after a project ends, constrains the extent to which impact can be evaluated. EVAL has developed specific guidance for adapting evaluation methods to the normative function (EVAL June 2020).This has supported an increase in evaluations that assess normative standards (EVAL/Luc Franche 2020). Nevertheless, the nature and funding modalities of ILO projects, and limitations in tracking longer-term outcomes make measuring impact on beneficiaries challenging: when evaluating impact it is important to design the evaluation in parallel with designing the intervention, and also requires funding partner commitment or access to non-project funding mechanisms.

The EAC with its senior level participation and strong governance arrangements, has the right people 'at the table' with influence and knowledge to bring about necessary changes[21]. However, the EAC members interviewed expressed a limited time and capacity to fully digest reports about follow-up to evaluation recommendations. The 2016 IEE recommended the EAC capacity be supported through introducing regional EACs to replicate and complement the central EAC mechanism. Unfortunately, that has not been achieved in full - only the Latin America and the Caribbean (Americas) region has established a regional monitoring and evaluation committee.

Maximising utility of HLEs: HLEs are complex, multi-sector endeavours which require significant effort and engagement across the ILO and externally. The quality and utility of the HLE's is premised upon full stakeholder participation in the process and integrating feedback into the final report.[22] The high visibility of the HLEs necessitates a higher level of effort, quality assessment, and time for thorough participation and meaningful feedback to minimize risk of poor quality and utility.

Evaluation recommendations: Although the evaluation function has seen an upward trend in the proportion of management responses to evaluation recommendations from decentralized evaluations completed; respondents pointed to challenges in implementing recommendations. Currently only 32 percent of evaluation recommendations are fully completed. Some of the reasons given for not implementing and using recommendations:

- Recommendations that are too project specific cannot easily generate insights for other regions, policy, or programme areas, whereas recommendations that are too general have limited utility to the evaluand.

---

[21] During the commentary phase of the evaluation report, a suggestion was made to include representatives from ACT EMP and ACTRAV on the EAC. The Employers' Group, in recent GB sessions, have highlighted that the Directors of ACT/EMP and ACTRAV are not part of the EAC and hence unable to offer intelligence and evidence if follow-up/implementation of recommendations are making a difference to employers and workers' organizations. Please refer to recommendations.

[22] The Gender HLE was cited in interviews as an evaluation that did not fulfil staff and constituents' expectations for utility and quality.

- Recommendations that do not reflect ILO's operating and institutional context detract from credibility and are not built to support to support actionable steps in response.

- Despite reported challenges, analysis of the AERs shows that many of the recommendations have few resource implications and are geared toward a short- or medium-term timeframe which affects the feasibility of implementation.

- Whilst the ex-post Quality Assurance results of decentralized evaluations for 2020 rated on average 5 (satisfactory), both the external QA assessors and respondents noted weaknesses in recommendations. Respondents reported that recommendations suffered due to the use of consultants with an inadequate knowledge of the contextual and structural context in which the ILO operates or lacking relevant technical expertise. External assessors noted substantive problems with recommendations (including lack of consideration of context, too numerous, lacking detail, poorly presented, not linked to finding etc.).

Barriers to uptake: There are a several interrelated factors that constrain the ILO to use evaluative evidence to shape programming and decision making:

- Volume of evaluations generated (429 in the review period) means that staff, managers, and constituents have limited absorptive capacity.

- Nature of the majority of the evaluations are low budget (see section 3.3.4) with too short timeframes to adequately assess what is often long-term nature of impact achievement. Design and nature of projects – often means that the timescales are too short to generate meaningful impact.

- Inflexible funding to undertake ex-post evaluations, lack of longitudinal studies and impact evaluations[23].

- Evaluation products may not be presenting answers to the right questions or may not be shaping findings and recommendations to maximize utility and application. Evaluations are not asking the 'bigger' questions and not demonstrating the impact (or lack of impact) this was a view expressed by the funding partners and staff alike.

- The predominant type[24] of ILO evaluations do not allow for real time learning and the identification of poorly performing interventions with adequate feedback loops (MOPAN, 2020). This aligns with the findings of the recent MOPAN assessment. EVAL is working with SIDA to embed more responsive evaluation, but the use of adaptive management has not seen widespread application.

3.1.2 Contribution to learning for decision making and continuous improvement

Evaluations offer opportunities for learning and can be an important tool to answer critical questions - understanding what works and doesn't work in specific circumstances to achieve certain objectives.

What is working well

Creative communications and dissemination: EVAL is delivering high-quality knowledge products and meeting its strategy targets. These include think pieces, synthesis reviews and meta-studies, convening learning series to discuss findings and recommendations –

---

[23] Thematic synthesis reviews and meta-analyses on ILO's decent work results and effectiveness of ILO operations are conducted by EVAL on a yearly basis.

[24] Majority of the ILO's evaluations are summative (rather than real time or developmental) end of term, process, or performance oriented, ILO's mid-term evaluations can be more useful in this respect.

particularly from high-level evaluations. The knowledge products summarize evaluation findings, achievements, lessons learned, and good practices from evaluations. There are also good examples of communications products highlighting the perspectives of project participants. Thirty knowledge products were produced over the Evaluation Strategy's implementation period this compared to the previous strategy period with 22 think pieces, meta studies and synthesis reviews delivered. The REOs (particularly in the Americas region) are also actively generating and disseminating knowledge products which are creative and highly visual thus enhancing likelihood of uptake.

The ILO's effectiveness in applying lessons learned increased during the COVID-19 crisis. The organization has applied key lessons from previous crises, notably the 2008-9 financial crisis, to its response to the COVID-19 pandemic. The organization consolidated evaluative lessons from its past response to the 2008 economic and financial crisis through its learning series and knowledge sharing platform. The interviews suggested that the pandemic has forced the ILO to become more responsive in making improvements based on lessons learned. (MOPAN 2021)

Evaluation reports are made available on the i-eval Discovery platform and are available for staff to access[25]. The ILO's Development Cooperation Internal Governance Manual (ILO 2021) encourages staff to use the i-eval Discovery platform to search for findings, lessons, good practices, recommendations and management responses from prior and other relevant evaluations when designing new programmes. EVAL shares information about evaluations in newsletters circulated to ILO staff and holds regular discussions to share information about evaluations with the evaluation network. High-level evaluation findings and recommendations are discussed at the Governing Body. This evaluation's survey results show that 87 per cent of respondents thought that ILO evaluations were moderately or largely successful in enabling organisational learning.

Constituent consultations: ILO tripartite constituents are consulted on the terms of reference for most evaluations, and they are asked for their comments on findings and draft recommendations before these are finalized. These are done primarily through ACTRAV and ACTEMP that provide secretariat functions to the ITUC and IOE, who reach out to their national constituents and technical specialists on thematic issues, where feasible. For decentralized evaluations, constituents involved in specific programmes and projects are consulted at the country level.

Opportunities for strengthening

Learning from evaluations: The ILO is able to identify lessons learned from evaluations, but uptake of these lessons can be a challenge for the organization, this was recognized by EVAL in their self-assessment and by MOPAN (MOPAN 2021) which found that there was little evidence of a culture of strategic evaluative thinking or opportunities to reflect on lessons from failures as well as successes.

Integration with RBM: Evidence suggests that evaluation findings, lessons learnt, and recommendations are not systematically embedded into the design or adaptation of programmes. Although lessons learnt from mid-term evaluations provide an essential opportunity to inform programming, this was rarely mentioned by respondents (beyond one person stating that mid-term evaluations were used to improve results frameworks). One funding partner did state that midterm evaluations were useful to inform funding decisions. Whilst final decentralized evaluations often provide the basis for subsequent phases of projects, undertaking real-time evaluation to make course corrections was not practiced

---

[25] i-eval Discovery is also publicly accessible via the ILO website.

widely. The SIDA funded work on adaptive programming and management being was cited as one example of this type of approach, which is increasingly being used by other agencies.

Feasibility and evaluability: Despite efforts by EVAL and PARDEV to encourage the use of evaluability assessments, this remains an area for improvement. There is a lack of project baselines and monitoring systems which constrain the value and evidence generated from evaluations. Identification of the organisational or programme learning needs at the early stage of the evaluation process is vital for successful generation of relevant lessons.

Communications: As stated above, the IEE found that stakeholders largely found lessons learnt and findings were presented in an accessible and useful way but that stakeholders did not necessarily engage fully with the products generated by the evaluation function.

> …there remains ample room to improve the integration of RBM across the organisations… the baselines in programme and project formulation is not yet mandatory for all interventions, which limits the ability to set results targets on a sound evidence base…monitoring and reporting practices need to be strengthened to address underperforming projects and programmes. (MOPAN, 2021)

According to EVAL's self-assessment (2022), the evaluation function has not fully harnessed the potential to share results and lessons with partners. The EVAL self-assessment reports that only on rare occasions are evaluation results and lessons learnt shared with other UN agencies and multilateral partners (most recently as part of the OECD-DAC COVID-19 coalition group). EVAL sometimes makes presentations about its work via UNEG, notably as part of the Evaluation Practice Exchange.

Enhancing engagement with constituents: In relation to workers' and employers' organisations, their engagement could be more cognizant of their limited time and capacity to contribute to evaluations. Several interviews highlighted that ILO constituents are often asked to comment on programmes they know nothing about, and at times when they do provide comments or raise concerns, there is little time to address these and integrate them into the final report. Where there are language and literacy barriers, paper-based (or virtual narrative forms of) consultation exercises are less than ideal.

3.2 What is the effectiveness of the evaluation function during the strategic period?

> **Key finding 2: The evaluation function has achieved or partially achieved 95 per cent of the sub-outcomes set out in its evaluation strategy. Limited resources and inconsistent engagement in evaluation from across the organization constrains the full achievement of targets. EVAL excels at producing regular, timely and data-rich reporting on evaluation performance, evaluations undertaken, findings and recommendations which is highly valued both internally and externally**
> **Strength of evidence: Level 3**

Overall, the sub-outcomes, milestones, and targets for implementing the evaluation ambition set out in the Evaluation Strategy have seen significant progress during the strategic period with 95 per cent of sub-outcomes achieved overall. As of August 2021, 68 per cent of sub-outcomes (13 out of 19) have been achieved, 26 per cent (5 out of 19) have been partially achieved, and 5 per cent (one sub-outcome) has not achieved progress. A synthesis of progress during the strategic period is summarized in Annex K. Progress against strategy outcomes is summarized below.

Outcome 1: Enhanced capacities and systems of evaluation for better use: Over the strategic period, there has been consistent delivery of evaluation with a substantial volume of evaluations completed. Evaluative capacity across the ILO and regions has been enhanced through a consistent demand for evaluation capacity building, which is met by a well-regarded training provision. There is considerable momentum behind the evaluation policy, and the mechanisms for implementation across the regions. With this, EVAL ensures adherence to the targets set in the evaluation policy. It has been more challenging to meet targets for internal evaluations and evaluability assessments. Although these are monitored, these are premised on an appetite for evaluative information rather than accountability requirements.

Outcome 2. Enhanced value of evaluation through the use of more credible and higher-quality evaluations: The quality and credibility of evaluations is supported by multiple layers of quality assurance, supported by up-to-date, comprehensive quality guidelines and knowledge products. EVAL developed guidance on conducting cluster evaluations which have seen an uptick in the strategic period and other guidance on approaches, methods and models. Although the decentralized evaluation function is strong, the capacity for decentralized evaluations is stretched. The indicator for use of impact evaluations that are implemented by technical departments has not achieved progress.

Outcome 3. Stronger knowledge base of evaluation findings and recommendations: EVAL has invested significantly in systems for communicating and following-up on evaluation findings and recommendations. EVAL's four meta studies and 20 synthesis reviews, think pieces, annual evaluation reports (AER), and evaluation specific products offer increased visibility, and more targeted communication of evaluation findings. Although EVAL has invested in a system to ensure follow-up to evaluation recommendations, this remains an area that has not fully achieved progress.

3.2.1 Evaluations conducted 2017-2021

The evaluation function produces centralized and decentralized evaluations. Centralized evaluations, which are at the governance level, are comprised of decent work country programme evaluations, high level evaluations (HLE), and thematic evaluations when there is a specific request. The work programme for HLE topics is the product of annual informal consultations which determines a rolling four-year work programme. This process supports a consistent work programme of HLEs (three per year).

According to data held on i-eval Discovery[26], the ILO conducted a total of 429 evaluations during the evaluation period 2017-2021. The largest share (81 per cent) of ILO's evaluations are **decentralized project evaluations**, including: 349 project evaluations (of which 64 were projects over 5 million USD). Other evaluations conducted include: 30 cluster evaluations (7%), eight country programme evaluations, six thematic evaluations, 20 RBSA funded activity evaluations, six DWCP reviews and six policy and strategic evaluations. **Independent** project evaluations accounted for 53 per cent of the total number of evaluations conducted. Fifteen high level evaluations were undertaken, of which five were Decent Works country programmes, a further 19 synthesis reviews and 4 meta studies were completed.

Annex G provides an overview of the evaluation portfolio along with table 2.

Table 2: Number of evaluations per type (data i-eval Discovery, analysis IEE team)

Evaluation Typology i-eval Discovery	2017	2018	2019	2020	2021	Total	%

---

[26] Evaluation data was extracted from i-EVAL Discovery April 2022, and supplemented by EVAL records and ILO website for high level evaluations (as this is not a distinct category in Discovery), meta studies and synthesis reviews.

DWCP Internal Review	3		1	2		6	1.4
Project review	0	0	1	1	0	2	0.5
RBSA funded activity	4	6	3	6	1	20	4.7
Project evaluations (inc. external evaluations, partnership evaluation, joint)	61	83	80	61	64	349	81.4
Cluster evaluation	2	5	5	9	9	30	7.0
Country Programme evaluation	3	2	1	1	1	8	1.9
Policy evaluation				1		1	0.2
Strategy evaluation	2	3				5	1.2
Thematic evaluation	2	2		1	1	6	1.4
Synthesis Review		1				1	0.2
Meta-study				1		1	0.2
	77	102	91	83	76	429	100
In addition, EVAL conducted 19 synthesis reviews, 4 meta-studies (2017-2020)							

Figure 4 shows the number of completed evaluations by evaluation nature as captured in i-eval Discovery for the period of 2017-2021.

Figure 6: Number of completed evaluations by type, 2017-2021 (data from EVAL, analysis IEE team)
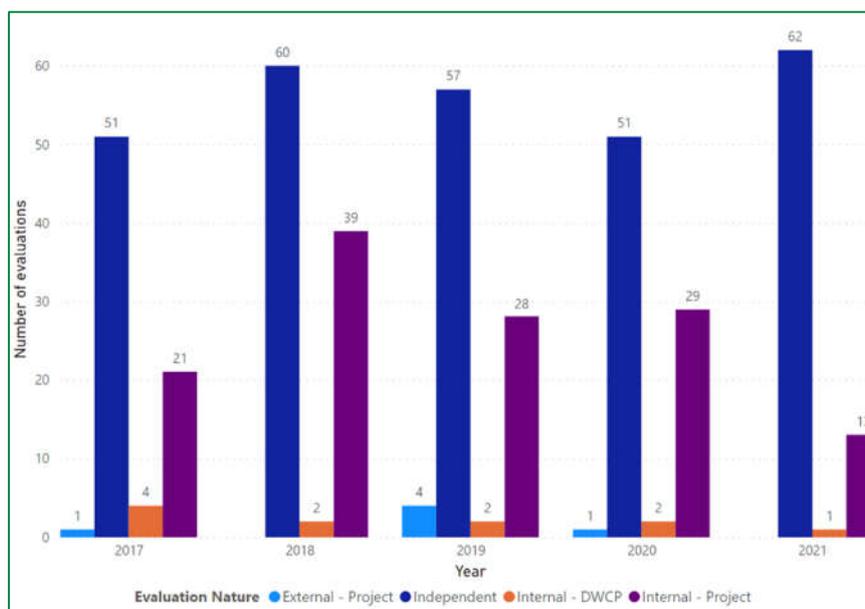
In terms of evaluation management, 43 per cent (185) of the evaluation portfolio was managed by REOs, 3 per cent (15) by EVAL and the remainder (54%) by evaluation managers.

Geographic coverage: In terms of geographic coverage, Asia and the Pacific and Africa have the largest project and programme evaluation portfolios - conducting 123 and 120 evaluations 2017-21 respectively. The Americas comes third (see fig 5. reflecting the number of DC projects during the same period[27]).

---

[27] Note there is not a one-to-one relationship given the timing of projects and evaluations
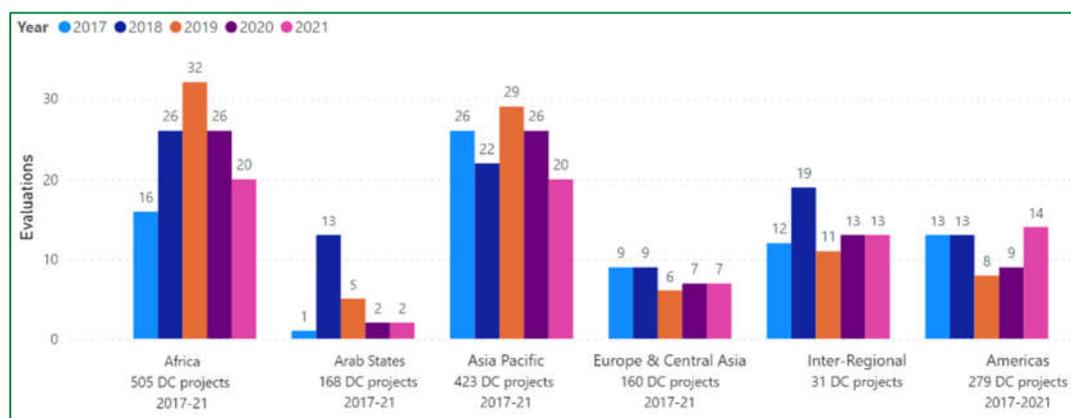
Impact evaluation: The Evaluation Strategy and Policy call for credible impact evaluations to be carried out in ILO as part of knowledge building on effective policy interventions. They are a complement to mandatory performance evaluations at all levels for demonstrating the impact of ILO work. Impact evaluations are the responsibility of technical departments rather than EVAL. A specific sub-outcome and indicator has been established in the evaluation strategy to track **quality** of impact evaluations (not number undertaken). Given the lack of a complete and accurate dataset[28] it is not possible to confirm the number of impact evaluations conducted, though the AER summary provided by EVAL states that a total of 17 impact evaluations and studies were completed in the past five years. Guidance and support from EVAL to impact evaluations is provided on a demand-basis to ensure the quality of impact evaluations. This is done primarily through guidance notes and the Impact Evaluation Review Facility initiated in 2016[29], records show this was used four times during the period under review. EVAL has a dedicated email for enquiries and has held one community of practice session during the review period.

Cluster evaluation: Clustering evaluations by theme, geography, and policy area is articulated in the evaluation strategy as a way to improve the evaluation efficiency and strategic value. As noted in the 2016 IEE, cluster evaluations have consistently been on EVAL's agenda as a proposed solution to addressing the need for evaluations that respond to learning, accountability, and decision-making needs of the ILO (ILO, 2011). To support this, the 2018-2021 Evaluation Strategy calls for "more strategic evaluations of projects and programme activities under identical or similar themes, programme frameworks and locations by means of clustering and integrated funding" as part of a transformative approach to evaluations (ILO, 2018). EVAL has worked to facilitate cluster evaluations by developing guidelines for the design and conduct of cluster evaluations. Clustering of evaluation is seen as part of implementing the more transformative approach to evaluation.

What is working well

Consistency in delivering against strategy targets: EVAL has established systems and mechanisms to create momentum behind the Evaluation Strategy and Policy which supports

---

[28] Impact Evaluations are not categorized as an evaluation type on i-EVAL discovery, instead a spreadsheet is maintained by EVAL to be updated by staff across the ILO engaged in impact evaluation.

[29] EVAL has developed an Impact Evaluation Review Facility (IERF) to offer support to those in the ILO who are considering, beginning or are in the process of implementing impact evaluations. The IERF is able to provide feedback, ideas, and suggestions for proposed and ongoing work on impact evaluation. It is also able to suggest relevant consultants, further technical support, and relevant technical resources. See EVAL guidance note 2.5 2020

adherence to strategic targets set out in the Evaluation Strategy. The evaluation function has consistently achieved its targets for independent evaluations which are overseen by EVAL. Between 2018-2021 EVAL provided oversight and quality control for a consistent pipeline of evaluations. Between 2018-2021, the evaluation function has produced a considerable volume of evaluations which offer good coverage of ILO's activities.

Cluster evaluations have seen an uptick in 2020, both in absolute numbers and as a proportion of all evaluation types which means that the evaluation function is moving towards realizing the ambition of its strategy. Evaluation managers surveyed in 2020 endorsed the view that clustered evaluations "improve cost and time efficiencies, reduce evaluation fatigue and enhance understanding among the tripartite constituents of the advantages of looking at the bigger picture of aggregated results at the country or thematic levels or by funding partner" (ILO, 2021). For this independent

Figure 8 Survey responses on number of 'strategic evaluations conducted in the ILO

High Level Evaluations or cluster evaluations

Too many 7% | Just right 65% | Not enough 28%

evaluation, nearly 75 per cent of survey respondents[30] thought that the ILO is conducting a sufficient number of cluster or HLEs (fig.6). The IEE team view the move to more cluster evaluations as a welcome development and a step in the right direction to a more strategic use of evaluation effort.

Table 3: Number of cluster evaluations data from EVAL, analysis IEE team

	2017	2018	2019	2020	2021	Total
# Cluster evaluations	2	5	5	9	9	30
Total # evaluations	77	102	91	83	76	429
Overall %	3%	5%	5%	11%	12%	7%

Opportunities for strengthening

Internal evaluations, which require oversight from regional and departmental directors rather than EVAL, have not met submission targets for the strategic period. Coverage for internal evaluations has increased but is still behind targets. Conduct of internal evaluations relies upon intrinsic motivation rather than external pressures or policing.

Dominance of project evaluation: The current portfolio of evaluations is dominated by individual project evaluations accounting for 81% the portfolio. There is a need to balance the need for project level information with ILO activities outside development cooperation programming. The volume of evaluations produced inevitably makes demands on the entire evaluation function in terms of managing, quality assuring, and more widely on the ILO in absorbing and using the evidence generated.

The use of cluster evaluations has not yet been optimized: it remains challenging, as noted in the AER and reinforced through stakeholder interviews, to achieve an optimal proportion of cluster evaluations, and offset the reliance on individual project evaluations with more strategic evaluations. Stakeholders report that despite their potential to offer economy of scale, they remain challenging to design, plan for, and implement given constraints at the design stage, and often disparate levels of information available from different projects or country offices that may be included in a cluster evaluation. Cluster evaluations have

---

[30] Covering all staff, senior management, evaluation managers, departmental evaluation focal points and regional evaluation officers

additional transaction costs, which require closer coordination across ILO departments to plan for cluster evaluations from the outset. For cluster evaluations to be conducted, there must be an appetite from funding partners at the project planning stage, and there must be up-front planning with policy/programme departments and with PARDEV to ensure alignment of project monitoring and evaluation frameworks and identify opportunities for more strategic evaluations.[31]

On impact evaluations, there is a need to balance the 'bottom-up' demand driven decision making process of when to conduct Impact Evaluations with more strategic direction and planning from EVAL and EAC – to ensure the most strategic use of evaluation resources and staff time. The nature of development cooperation funding can restrict the ability to conduct impact evaluations that require data collection once the project has ended. As such, overcoming these challenges require a concerted and targeted strategic approach (see recommendation nine). Impact evaluations are not adequately captured on the i-track system and reliance on self-completed spreadsheet is not adequate to manage information needs. Analysis of the records provided by EVAL show that the Impact Evaluation Review Facility appears to be underutilized, either due to declining number of evaluations, alternative mechanisms being used or lack of awareness of the service or recognition of need.

3.2.2 UNEG quality and credibility standards

The IEE looked specifically at implementation of the UNEG norms and standards 2016 (see list provided in Annex M). The norms and standards exist to ensure that UNEG members adhere to shared basic principles and apply best practices in managing, conducting and using evaluations. This section of the report specifically presents the findings pertaining to quality and credibility standards[32]. The norms and standards pertaining to independence are presented in section 3. 3.3, and a summative assessment of alignment to all UN norms and standards is presented in section 3.4.1.

What is working well

The evaluation function benefits from multiple layers of quality assurance, supported by up-to-date, comprehensive quality guidelines. The ILO evaluation function has quality guidelines for the conduct and quality assurance for centralized and decentralized evaluations. ILO's guidance has evolved over time and currently is comprised of a suite of complementary materials for managing, ensuring resourcing and quality of different types of evaluations and technical guidance across a range of key topics relevant to ILO. Stakeholders from peer UN agencies indicated that ILO guidance has informed their own evaluation guidance materials.

EVAL is responsible for ensuring the proper implementation of the evaluation policy, aligning with funding partners' requirements, and for evaluation quality oversight. Keeping up with emerging OECD and UNEG standards, EVAL has revised guidance notes on the use of OECD DAC criteria and cross-cutting issues. The guidance notes are a critical pillar of EVAL's support to quality assurance of evaluative processes and products and consist of templates, checklists, protocols, and guidance notes. In the latest update to the guidelines (2020), which

---

[31] EVAL is commissioning an Ex-post quality review of impact evaluations and a Review of the approach to clustered evaluations and their strategic value. The former is expected to provide insight on the factors driving trends in evaluation practice and uptake for these types of evaluations. The latter is focusing on documentation of the specific contribution and experience with clustered evaluation, including the type of strategic findings that clustered evaluations provide https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_837829.pdf.

[32] Ten general norms should be upheld in the conduct of any evaluation; the four institutional norms should be reflected in the management and governance of evaluation functions. The associated standards support the implementation of these normative principles.

were developed with expert input and with the help of a communication consultant, the guidance notes were organized around five themes rather than being presented as a long list. The electronic version has been redesigned for greater ease of use and substantive updates to the content on evaluation design, cluster evaluation, social dialogue, and gender.

Management and quality control of project level and decentralized evaluations benefits from multiple layers of quality assurance, as shown in fig. 7.

Figure 9 Summary of quality assurance process and roles

Evaluation Managers come from across the ILO and are either appointed or express interest in carrying out this role. The Evaluation Managers ensure quality of the evaluations guided by comprehensive evaluation quality guidance and resources, including the Certification for Evaluation Managers. According to EVAL's internal records, Evaluation Managers tend to manage on average 0.71 evaluations per year (2017-2120 period) after obtaining their certification (ILO,

Figure 10 Quality Trends by QA criteria 2019 and 2020 (data EVAL , analysis IEE
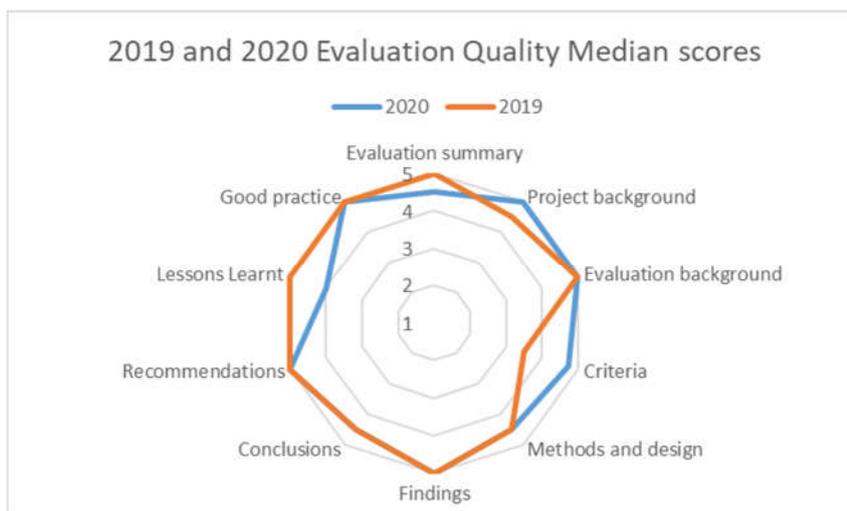
2022). The role is undertaken in addition to their full-time roles. In the instance that an evaluation manager cannot be identified, evaluations are managed by REOs or DEFPs. The REO provides quality backstopping for DWCP reviews and project evaluations[33]. REOs will quality assure the evaluative products for between 6 and 31 evaluations in a given year. All evaluative products for both centralized and decentralized evaluations are quality assured by EVAL. This translates to an average of 57 independent evaluations per year, in addition to support to internal evaluations, and the production of synthesis reviews and meta-studies. Ex-post quality assessment moved from being conducted annually by an external assessor to ex post quality assessment conducted - on a rolling basis. According to external ex-post quality assessment commissioned in 2021 and validation by MOPAN and this IEE, ILO's evaluations have largely met quality standards which are aligned to UNEG standards and adopt OECD DAC criteria (Table 4).

Table 4: Quality trends from external quality assessment processes compiled for IEE

	Quality assurance	Scope	Year	Sample size	Quality rating	Strengths	Challenges

---

[33] The REO manages DWCP reviews as part of the Regional Programme Unit.

High level Evaluations	External ex post quality assessment	2015-2021	2021	10 HLE	90 per cent received a rating equal to or above "somewhat satisfactory"	Satisfactorily incorporated specific approaches to addressing social dialogue, tripartism and the SDG	Some improvements are still required to ensure a better systematic reflection of international labour standards.
Centralized and decentralized evaluation	MOPAN	2016-2020	2020	Reference to Annual Evaluation report 2019-20	N/A	Systematic use of evaluation guidelines	Limited availability and therefore use of monitoring information
Project evaluation reports	External ex post quality assessment [34]	Jan-Dec 2020	Nov 2021	46 project evaluation reports	100 per cent of reports completed received a rating equal to or above "somewhat satisfactory; they contained an average of 91 per cent of the components that are essential for an evaluation report.		

Analysis of the ex-post quality assessments conducted in 2019 and 2020[35] reveals that (independent) evaluations scored satisfactory or somewhat satisfactory across the majority of indicators (fig 8).

Opportunities for strengthening

There needs to be a continued concerted effort to improve quality assessment scores for evaluation on gender and cross cutting issues such as disability.

In terms of achieving system-wide gender targets for evaluation measured through the UN-SWAP (UN System-wide Action Plan on Gender Equality and the Empowerment of Women), the IEE analysis reveals that of the 39 evaluations conducted in 2019 and 46 evaluations conducted 2020 the UN-SWAP score was 4.31 points and 4 respectively out of a maximum of 9 this translates as "*approaches requirement*". In 2020 only four evaluations were quality assessed as meeting the UN-SWAP requirement.

Table 5: Quality assurance scores for 2020 evaluations UN-SWAP (data EVAL, analysis IEE)

UN SWAP Score assessed by Quality Assurer	2020 evaluations QA score
Meets requirement	4 (8.7%) out of 46 evaluations
Approaches requirement	23 (50%) out of 46 evaluations
Misses requirement	19 (41%) out of 46 evaluations

In 2020 the presentation of lessons learnt achieved the lowest median score – lessons learnt are judged to be highly satisfactory when they present significant non-trivial matters, concisely capture context, are applicable in different contexts, target specific users, and suggest what should be repeated or avoided in future.

Consideration of cross cutting policy drivers[36], inclusion for persons with disabilities in the evaluation questions, and the inclusion of evaluation questions to examine the ILO's

---

[34] Universalia (2021), Results and reflections from a quality appraisal of ILO evaluations, 2020 i-eval Think Piece No. 21: https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_830075.pdf

[35] (EVAL provided quality assessment data by criteria for 2019 and 2020 only. Criteria provided in Annex N)

[36] Prior to 2020, CCPDs encompassed the following: international labour standards, tripartism and social dialogue, gender equality and non-discrimination (including disability), and a just transition to environmental sustainability. During 2020-21, CCPDs were discontinued and the marker on gender equality and non-discrimination was introduced.

delivery of mitigating effects against COVID 19, were the poorest scoring elements of the ex-post quality assessment of evaluations undertaken externally for 2019 and 2020[37]. It will be necessary for the ILO to reflect on the results of the 2021 ex-post quality assessment once completed to form a complete picture of trends and signs of improvement.

Part of the quality criteria includes an assessment of whether the evaluation has adequately described unintended or unexpected outcomes of the intervention, this would usually be where the evaluators set out any outcomes achieved (either positive or negative) that were not foreseen in the design of the intervention, despite inclusion in EVAL's guidance on writing ToRs evaluations scored poorly in 2019 and 2020 with the median score of 1 (highly unsatisfactory) for 2020[38].

The established QA process requires that EVAL oversees a large volume of evaluations. With final oversight of all evaluation reports that come through the pipeline, the volume of evaluations that EVAL must quality assure is considerable. With all quality assurance taking place 'in house' by a relatively small evaluation office, evaluation personnel within EVAL are stretched. It should be noted that thus far this has not had an impact on the overall quality of evaluations, as demonstrated in fig. 9. Instead, the impact is felt by staff who raise concern about the sustainability of the volume of work that EVAL manages on a day-to-day basis given their finite time and human resources. The IEE team note that the use of median scores to establish the overall rating of the evaluation product risks masking fundamental flaws in evaluations – other UN agencies and bilateral organizations use weighted indicators and critical 'red lines' which can provide a robust quality check mechanism.

The evaluation function must balance extensive and varied interests, with limited time and budget in implementing the evaluation policy guidelines, the development of terms of reference for each evaluation involves consultation with ILO's diverse internal and external stakeholders and constituent groups. Evaluations try to serve interests of different stakeholder groups often by asking a wide range of questions[39], or by focusing on specific technical issues (e.g., market systems development, social protection, employment investment). This creates an extensive scope for evaluations which creates what some stakeholders referred to as "an impossible task" or at least a "tall order." In balancing stakeholders needs, evaluation design by evaluation teams must consider the need to balance evaluation skills with the need for specific technical areas within finite time and resources for an evaluation.

3.2.3 Developing evaluation capacity and competencies

ILO's evaluation function strengthens the evaluation capacity across the ILO through the guidelines, checklists, and training provision for ILO stakeholders and constituents. EVAL partners with ILO International Training Centre (ITC) to provide evaluation trainings. The cornerstone of EVAL's capacity building opportunities are the following training programmes, with table 6 showing the achievements of targets:

The Evaluation Manager Certification Programme (EMCP) has been conducted on an annual basis since 2011. During the height of the COVID-19 pandemic, the training was delayed and training modules reconceptualized for use in an online format. The programme prepares ILO staff to manage evaluations.

---

[37] Covid 19 inclusion median score 2020 Highly unsatisfactory, with 32% evaluations scoring 1. Consideration of persons with disabilities median score highly unsatisfactory with 73% evaluations scoring 1. Cross cutting policy drivers: median score 1 highly unsatisfactory with 82% evaluations scoring 1.

[38] 67% scored 1 (highly unsatisfactory) or 2 unsatisfactory for 2020 evaluations

[39] The ILO's Evaluation Policy Guidelines for results-based evaluations (p33) encourages evaluations to examine all six DAC criteria, when insufficient data prevents the evaluation of certain criteria the evaluator should determine which criteria to use.

Evaluation Training for Tripartite Constituents. In November 2018, EVAL launched a training programme for ILO constituents on evaluating the Decent Work Agenda in the era of the Sustainable Development Goals (SDGs). During the Evaluation Strategy's implementation period, EVAL has mainstreamed evaluation into training initiatives for all three constituent groups, to maximize their contributions to evaluation processes at the country, regional and global levels.

The Internal Evaluation Certification Programme (IECP) was developed in 2017 for ILO staff who may need to conduct internal evaluations of ILO projects and programmes as a tool to support engagement and greater use of internal evaluations. Given that the IECP training has been phased out and will be available as an online offer, this section focuses on the EMCP and the evaluation training for Tripartite constituents.

Table 6: Capacity Building initiatives

	Launched	Target groups	# Trained 2018-2021	Effects of COVID-19	Status
EMCP	2011	ILO staff	133	Courses delayed in 2020. Course moved online. In person training scheduled for July 2022.	Targets achieved.
IECP	2017	ILO staff	25	Modules moved to online format	Cancelled due to limited demand for internal evaluations. Course superseded with an advanced ECMP.
Evaluation training for constituents	2012	Constituents	297	Online materials developed	Targets achieved. Participation predominantly from workers (62 per cent). This compares with the previous review period with 1052 constituents training 2010-2017.

What is working well

Demand for capacity building: There has been a consistent increase in the number of ILO staff trained to be evaluation managers through ITC ILO and demand for evaluation capacity building training programmes continues. ILO's Evaluation Function is achieving or exceeding training targets through the EMCP. It benefits from an active pipeline of interested participants. The training offer is well-regarded externally. Stakeholders from peer UN organizations pointed to the ILO EMCP as a strength of the evaluation function. While the COVID-19 pandemic interrupted the continuity of the course offered by the Turin-based International Training Centre of the ILO (ITCILO), there was a high subscription rate for the June 2022 Evaluation Manager training and recognition across the ILO of its value. ILO staff noted that it offers value not only for being able to manage an evaluation, but also to build staff competency to "think differently." Specifically, staff pointed to the value of participating in the training to understand causal relationships in project design, and to understand Theory of Change for better project design.

Participation in the ECMP comes predominantly from regional staff, there are few technical specialists (16 out of a total of 122) from programme units. Of the evaluation managers polled[40] for the IEE 2022, 76 per cent agreed that they have an interest in receiving more training and capacity building. Based on EVAL data, 91 out of the 122 (74 per cent)[41] certified

---

[40] A rapid poll was undertaken at the end of the FGD with Evaluation Managers
[41] This number excludes REOs and is based upon the number of staff that have achieved certification. Not all REOs, many of whom are already experienced evaluators, have been through the ECMP process.

evaluation managers went on to manage an evaluation in the first year after participating in the EMCP (fig 9.).

Figure 11 Evaluations managed by certified evaluation managers, data EVAL analysis IEE team

Evaluations managed by certified evaluation managers after certification

Table 7: Number of evaluations managed by evaluation managers 2017-21, data EVAL analysis IEE team

Number of evaluations managed 2017-21	Number of evaluation managers
1 evaluation	91
2 or 3 evaluations	34
4 or 5 evaluations	5
6 -8	3
9-12 evaluations	2

On average, staff that have been through the EMCP managed 1 or 2 evaluations (average 1.58), see table 7. From the evaluation database records of the evaluations managed outside of EVAL, 91 evaluation managers only managed 1 evaluation (there is one evaluation manager who accounts for 5 per cent of all evaluations conducted).

Training programme for Tripartite Constituents: The training programme for ILO constituents is part of the evaluation strategy's emphasis on contributing to the set-up and use of participatory SDG evaluation systems at national level and enhance their ability to measure SDG progress. The learning objectives of the programme focus on the role of the SDGs and their link with DWCPs and the role of constituents in review and reporting on SDG progress. The training programme also serves to socialize the value of evaluation for their work. Led by EVAL, training initiatives with constituent groups now have an evaluation component. Constituent representatives report that there has been some uptake of capacity building opportunities from EVAL. The extent that the capacity building is supporting demand for evaluation as part of DWCP continues to be monitored. There are examples from the field of constituents engaging in national development processes and SDG processes. Progress in this area hinges on continuing to mainstream evaluation in other training and capacity building programmes. As noted in the 2020-21 AER, constituents were involved in evaluation activities in 66 per cent of cases in decentralized project-level evaluations.

Opportunities for strengthening

Although there is a consistent demand to participate in the EMCP, stakeholders report that the ILO is still lacking adequate incentive structures for staff to want to manage evaluations after participating in the trainings. In addition, even though staff are trained, there is still a paucity of available evaluation managers. There are few incentives for follow-up to the training programme, for example, no requirement that EMCP participants take on the role of evaluation manager. While EVAL plans an enhanced EMCP programme ('EMCP+'), at present the certification is a one-off programme, including a practicum which does not have refresher offerings.

GB members and constituent (93 per cent of IEE survey respondents) stated that they had not received any evaluation training in the past five years. There is scope to increase evaluation knowledge among Tripartite Constituents (80 per cent of respondents expressed interest in receiving training) by increasing awareness of training opportunities among Constituents. While post-training exit surveys are broadly positive, the trainings may not

respond to the information needs of constituent groups or consider their limited bandwidth to participate in these trainings. As identified in the 2021 AER, tailored evaluation training initiatives with under-represented constituent groups should be enhanced, in collaboration with other entities to ensure better coverage of evaluation capacity among all constituent groups (ILO, 2021).

Table 8: Strengths and challenges of EMCP and Constituent trainings

	Strengths	Challenges
EMCP	Appreciation of value of EMCP	Limited participation of technical specialists in EMCP
	Demand for EMCP	Participation in EMCP does not guarantee taking on role of EM
	Opportunities for evaluation thinking to support programming more broadly	Evaluation is a professional area. Concern over the capacity required to comment on quality.
Constituent training	Constituent training geared toward policy influence	Limited evidence of training translating to participation in evaluations
	Constituent training is mainstreamed in other training offers	Limited evidence of the extent the format and content of trainings is most useful to constituents

3.3 How efficient, sustainable, and fit-for-purpose are the current arrangements and structures of the evaluation function?

Key Finding 3: ILO's structural arrangements give weight to the evaluation function and mainstream evaluation across the organization and across regions. However, the efficiency, sustainability, and adequacy of the evaluation structure is threatened by limited investment, organizational incentives and buy-in to evaluation as a source of knowledge and learning. Strength of evidence: Level 3

3.3.1 Structural arrangements and organizational relationships for accountability, organizational learning, and planning

ILO's evaluation function is driven by an independent evaluation office (EVAL) with seven staff members, supported by an evaluation network comprised of evaluation focal points (DEFPs) within each department, a REO within each region, and certified evaluation managers from across the organization. The function also relies on external evaluation consultants, recruited through a public procurement process for each evaluation, who are contracted to conduct independent evaluations. Oversight of evaluation follow-up is provided by the EAC which is composed of senior managers from across the ILO. The structures of the evaluation function are built to support accountability and cross-organizational learning.

What is working well

In contrast to other evaluation offices of broadly comparable maturity and operational scope, EVAL has a decidedly more 'hands on' approach in the quality assurance and oversight of evaluation products. It conducts significantly more evaluations per year and has a more extensive decentralized structure [42]. It also differs in its use of certified volunteer evaluation

---

[42] UNFPA, UNESCO

managers to embed evaluation across the organization, rather than relying solely on designated evaluation staff. This could result in efficiencies and opportunities to embed evaluation across the organization. However, this is a significant undertaking for evaluation managers and could further compromise its sustainability, given current workload challenges.

Accountability mechanisms for the evaluation function are strong. The evaluation function has a direct reporting line from the EVAL office to the Director General and Governing Body and is supported in its mission by the EAC. In addition, the use of external evaluation consultants supports credibility and sustainability of the evaluation function.

ILO's Evaluation Network, comprised of DEFPs, REOs, and internal certified evaluation managers is a strength in the design of the ILO's evaluation structure. Organizational learning is supported through the hybrid nature of the evaluation function and the potential for exchange between EVAL and other departments. This comes through the inclusion of staff from across the organization, particularly through the representation of senior leadership in the EAC. The DEFPs serve as ambassadors for evaluation in their departments. They are appointed by the department director, offering valuable evaluative insights on technical policy and programme areas. REOs in-country have value and understanding of what is needed from the ground and translate what is happening within EVAL back to the region. The evaluation manager system is a strategic and cost-efficient way to, as one stakeholder interviewee suggested, "democratize" evaluations across the organization so that evaluations are not just the responsibility of a small segment of the organization.

Opportunities for strengthening

The evaluation management system, as noted previously, suffers from practical and structural limitations which it must overcome to be more effective. This particularly means encouraging staff within departments and regions to assume additional responsibilities as evaluation managers or getting line managers to approve the release of their staff for evaluation management tasks. Currently the role is assumed by certified ILO staff who either have previous experience in evaluation, an intrinsic interest in developing additional skills and widening their perspective of ILO programming, or in some cases, where the role is encouraged by a line manager as part of career development. Table 9 sets out the opportunities for accountability and learning supported by the hybrid structure as well as risks for efficiency and sustainability.

Table 9: Assessment of hybrid structure opportunities for accountability and learning and risks

Component of hybrid structure	Opportunities		Risks
	Accountability	Organisational learning	Efficiency and sustainability
EVAL	Direct reporting line to DG and GB adds weight to evaluation function	Knowledge generation, oversight, and capacity building role for the whole ILO	Small office with extensive responsibilities.
EAC	Example of good governance with strong accountability mechanisms.	Participation from across the ILO supports learning from across the organization.	The volume of EAC's agenda means that EAC limits the time for strategic discussions.
DEFP	Supports compliance and participation from across organization in centralized evaluations.	Representative for evaluation across departments. Positioned to share policy/programme insight with EVAL and versa.	Evaluation is an added responsibility for the DEFP. Performance indicators are added to DEFP performance cycle, but the role is not allocated additional time or resources.
REO	Technical reporting line to EVAL Director exists.	Field presence supports learning from regional contexts.	Responsible for largest volume of evaluations. No first-line performance reporting to EVAL.
Evaluation managers	There no formal reporting lines but responsibility for evaluation compliance.	Mainstreaming of evaluation across the organization has potential to enhance learning exchange.	Evaluation management is time consuming and is voluntary. There is a lack of EMs which risks sustainability of function.

Evaluation consultants	Independent evaluation consultants support credibility of evaluations.	External support does not necessarily feed into organisational learning	Public procurement for all independent evaluations. Long term agreements could support efficiencies in recruitment.

3.3.2 Evaluative culture to enable the successful delivery of the evaluation strategy objectives

Creating a strong evaluation culture requires an assiduous balance between oversight and support, balancing the roles of "*watchdog*" and "*guide dog*" while avoiding the role of "*lapdog*" (Schwandt July 2019, 25, no. 3). Evaluation culture is referenced explicitly in the ILO's evaluation strategy as an enabling condition for an effective evaluation function. Evaluation culture, within the context of this IEE refers to the attitudes and behaviors that contribute to how people think about and interact with the ILO evaluation function (Boris Groysberg Jan/Feb 2018). As noted in the World Bank Evaluation Principles, "building an evaluation culture is not only about strengthening the supply side (that is, the quality and focus of evaluations) but also about strengthening the demand side—for example, by building a common understanding of the role of evaluative evidence in learning and accountability" (World Bank Group/IEG 2019). More broadly, the UN Secretary-General has identified the transformation of Organizational culture as a central tenet of UN reform in enabling the Organization to better deliver results and has called for a culture focused on results, risk management, tolerance for mistakes, innovation, empowerment, and accountability. As such, the evaluation function is considered a valuable driver of performance improvement. The 2020 OIOS-IED report on organizational culture (which did not include UN specialized agencies like the ILO within its scope) found that "a poor culture of evaluation was a top inhibiting factor for the increased use and utility of the evaluation function as an accountability tool."

In this vein, the ILO's Evaluation Strategy calls for a "culture of more comprehensive evaluation in support of the ILO's mandate" and states that a "strong evaluation culture is required to achieve better organizational performance, effectiveness and learning." The Strategy points to several complementary factors required to support an evaluation culture within ILO which include capacity building opportunities for ILO staff and constituents, tools and instruments to improve monitoring and evaluation systems, a strong, functioning evaluation network, and commitment from across the organization in the form of adequate capacities and resources for evaluation activities, and consistent response to evaluation recommendations at the Governing Body level, and institutional incentives to use and learn from evaluations. Enabling factors for the evaluation culture, implied in the Evaluation Strategy, are set out in Table 10.

Table 10: Enabling a strong evaluation culture

Components required for a strong evaluation culture	Successes	Challenges
Capacity building opportunities for ILO staff and constituents	Strong training offer from EVAL	Training, though practicum based, is one-off and does not require that stakeholders engage in evaluation training thereafter.
Evaluation network	The evaluation network has representation across the ILO	Evaluation responsibilities (Evaluation manager) rely on staff to take on responsibilities on top of their full-time staff positions.

Tools and instruments to support implementation of evaluation	Extensive guidance materials for design and implementation of evaluations	Managing evaluations can be a complex endeavour requiring sufficient time to absorb and implement best practice guidance
Consistent response to evaluation recommendations	Tracker for management response to evaluation recommendations	Recommendations are often considered too general to be implemented by a project or too specific to be implemented widely
Incentives to use and learn from evaluations	Stakeholders report that evaluation is an important tool for compliance	Integration of lessons learned in project design is a work in progress
Access to evaluation products	I-Eval Discovery provides a public, searchable database of evaluations	Survey responses suggest i-eval Discovery is not entirely accessible. [43]
Timeliness of evaluation products	Evaluations are largely delivered on-time	Short project timeframes (e.g., 18-24 months) detracts from the value of evaluation findings.

Aspects of a strong evaluation culture, including self-evaluation, evidence-based learning, and knowledge transfer are supported through the evaluation strategy (see Mayne for a presentation of evaluation culture, seen through a wider lens (Mayne 2010).

What is working well

Beyond the "hardware" that supports a good evaluation culture, there are less-tangible aspects that determine the embeddedness and value of evaluation for the ILO, the IEE found:

An appetite to understand what is being achieved: There is a demand for evidence on how ILO is contributing to its mandate both at governance and at the project level[44].

A shared commitment to compliance: Staff members expressed that there is a shared commitment to complying with evaluation responsibilities.

Endorsement of evaluation quality: The quality, professionalism, and independence of evaluation is well-recognized and endorsed by ILO staff and stakeholders. There is an overall trust in the integrity and credibility of the evaluation process and products. Although there are anecdotal exceptions, overall, ILO staff and stakeholders endorse the quality of evaluators and evaluations.

Cross-organizational leadership by design: The evaluation function benefits from a structure that embeds senior management and leaders in the evaluation function. The role of the EAC in engaging with HLEs is set up to minimize the risk that evaluation would but siloed from the organization.

Synthesis of evidence: Meta-analysis of the ILO's work has been conducted for 10 years. The synthesis reports aim to provide insight into lessons learned; what works and why. This contributes significantly to the availability of evaluative lessons from across ILO programming.

Opportunities for strengthening

Placing value on evaluation: Although the evaluation function offers efficiency gains through working through the evaluation manager model and supports mainstreaming across the

---

[43] Survey question 7: Only 50 per cent of Evaluation Managers, DEFPs and REOs state that the online EVAL database is known and easy to access and use. 28 per cent report that it is not easy to access and use and 22 per cent state that they do not know.

[44] See section 3.1

organization, this is not supported by incentives or resourcing which would enable greater participation and valuing of evaluation.

Uptake of internal evaluations: As noted in consecutive AERs, targets for internal evaluations have not been reached even though internal evaluation being a key part of the evaluation infrastructure. For self-evaluations by programmes and projects, EVAL, however, offers leadership in self-evaluation through its comprehensive self-assessment.

Uptake of evaluation recommendations: In spite of clear internal governance documents, programme and policy departments, there is still a need to strengthen evaluation follow-up from across the ILOs. Partly due to a perception that evaluations do not capture institutional and operating constraints, and partly due to a culture where evaluation is not always part of decision-making.

Build trust and strengthen the enabling role of EVAL: EVAL straddles the role of oversight and as purveyor of evaluation knowledge. There is an opportunity for EVAL to emphasize its role in learning across the organization to build active customers and users of evaluation, and to promote and embed evaluative thinking[45] and practice. This does not need to be in tension with EVAL's oversight role but can be an important reinforcement of the need for and value of strengthening evaluation across the organization.

Limited demand for evaluative knowledge: Although stakeholders widely consider evaluation important for accountability purposes, stakeholders report that there is a tendency for programmes to generate their own research, rather than to look to evaluations as a useful source of knowledge. A critical test of a positive evaluation culture is the degree to which the design of new initiatives, decision-making, learning, and compliance systems draw on evaluation evidence from across an organization. The IEE found that such an evaluation culture is not fully embedded in the ILO.

Embrace and learn from failure: Learning from failure both from evaluations and across the evaluation function is critical for a strong evaluation function and relies on coordinated efforts from country offices and programming units (see section 3.1].

3.3.3 Independence of the evaluation function

Independence of evaluation is necessary for credibility, influences the ways in which an evaluation is used and allows evaluators to be impartial and free from undue pressure throughout the evaluation process (UNEG Norms and Standards 2016). Independence protects the integrity of the assessment process, enhances its credibility, minimizes bias, and provides fresh perspectives on the policies and programs being evaluated (Mayne 2010). UNEG Norm 4 elaborates a requirement for behavioural independence and organizational independence of the central evaluation function.

What is working well

ILO upholds independence norms and standards: The ILO EVAL function's independence has been recognized in both the MOPAN study and the 2014 JIU study of the evaluation functions in the UN system. According to JIU (2014:15), ILO is now one of only six UN entities to have achieved the highest level of structural independence of 24 assessed entities in the UN system. According to latest ILO MOPAN Assessment (2020) the evaluation function of the ILO is operationally and financially independent. Following the IEE 2010, the ILO transformed its Evaluation Office into an independent unit.

---

[45] Evaluative thinking is defined as critical thinking applied in the context of evaluation, motivated by an attitude of inquisitiveness and a belief in the value of evidence, that involves identifying assumptions, posing thoughtful questions, pursuing deeper understanding through reflection and perspective taking, and informing decisions in preparation for action. (Jane Buckley 2015)

As ensured by the GB decision in 2011 and enshrined in the Evaluation Policy of 2017, the Evaluation Office reports directly to the Director-General and the Governing Body and is funded through the regular operating budget. It has full discretion in deciding on the evaluation programmes and is also directly responsible for high-level evaluations. EVAL continues to have a three-layered oversight and quality control of evaluations (including EVAL, REO/DEFP, and the Evaluation Managers - EMs) that uphold and promote independence and impartiality throughout the evaluation processes. These principles were fostered during the Strategy period through the issuance of revised policy guidelines[46] and guidance material to ensure standardized evaluation practices adhere to the updated UNEG's Norms and Standards (UNEG 2016) and ethical guidelines (UNEG, Ethical guidelines for evaluation 2020).

Table 11 sets out the IEE team assessment of ILO's application and fulfilment of the UNEG Norms and standards relating to independence and impartiality[47].

Table 11: ILO application and fulfilment of the UN norms and standards on independence and impartiality

UNEG NORM 4 INDEPENDENCE
Structural independence
✓ The director of EVAL heads an independent evaluation function.
✓ EVAL reports directly to the DG and the GB.
✓ EVAL is mandated by the GB to manage the evaluation function and ensure proper implementation of the evaluation policy.
✓ An evaluation advisory committee (EAC) provides a mechanism to oversee the use, implementation, follow-up to lessons learned and recommendations resulting from ILO evaluation activities and provide advice to the director-general on progress made by EVAL.
✓ Decision making on evaluations outside of EVAL – i.e., decentralized evaluations is guided by evaluation policy not the employees directly responsible for implementing interventions
✓ Decentralized evaluations subject to internal real time quality assurance by REOs and EVAL, and external post-hoc quality assessment. **Centralized evaluations are subject to internal peer review and post-hoc external quality assessment.**
The evaluation function sets the evaluation agenda
✓ The Director of EVAL has full discretion over EVAL commissioned evaluation selection and approval of evaluation reports to the Board.
✓ Mechanisms are in place to ensure that evaluations commissioned outside of EVAL are free from undue influence and that reporting is unbiased and transparent.
✓ Evaluations are submitted directly for consideration at the appropriate level of decision making for the subject of evaluation. The operational framework, set out in the ILO Policy Guidelines for Results-Based Evaluation, enables evaluations to be submitted directly to different levels of decision making according to the subject of the evaluation
The evaluation function is provided with adequate resources for conducting its work
✓ The evaluation function of the ILO is financially independent
✓ The Evaluation Office is funded through the regular operating budget
✓ In projects, resources are earmarked for monitoring and evaluation with a minimum of 2% of total project resources reserved for evaluations and an additional 3% reserved for monitoring and reporting (at a total minimum of 5%), according to the ILO Finance Manual and the ILO Policy Guidelines for Evaluation[48]
Behavioral independence and impartiality in connection with evaluations must not have negative repercussions including for career advancement
✓ All evaluations for programmes over $500k are conducted by someone with no ties or COI with management of the intervention. Over $1m recommended use of consultant external to the ILO
✓ Evaluators are able to conduct their work during the evaluation without undue interference by those involved in implementing the unit of analysis being evaluated.

---

[46] 4th edition, 2020

[47] See Annex M UNEG norms and standards in full. This assessment also informs the IEE team's assessment of ILO Evaluation Maturity section 3.4.1

[48] In addition to the evaluation policy on allocation of 2% for evaluation and 3% monitoring and reporting, the ILO Development cooperation internal governance manual 2021 (ILO 2021) also encourages the option of real cost estimates for M&E budgets rather than a flat percentage.

✓	Director of EVAL provides assurance on compliance with norms and standards for decentralized evaluations.

<table>
<tr><td colspan="2" style="background:#00B0A6;color:white">UNEG NORM 5 IMPARTIALITY: OBJECTIVITY, PROFESSIONAL INTEGRITY AND ABSENCE OF BIAS</td></tr>
<tr><td colspan="2">Evaluators need to be impartial and must not have been (or expect to be in the near future) directly responsible for setting policy in respect of the subject of the evaluation or its design or management</td></tr>
<tr><td>✓</td><td>Evaluation managers are located in departments and regional/ country offices and are independent from the programme evaluated.</td></tr>
<tr><td>✓</td><td>All consultants sign a code of conduct</td></tr>
<tr><td>✓</td><td>The roles and responsibilities are clearly defined in the ILO's evaluation policy and related policy guidelines</td></tr>
<tr><td colspan="2">Impartiality exists at all stages of the evaluation process</td></tr>
<tr><td>✓</td><td>Analytical transparency is built into the evaluation design, conduct and reporting</td></tr>
<tr><td>✓</td><td>System of REOs, DEFPs, Evaluation Managers support implementation of decentralized evaluations</td></tr>
<tr><td>✓</td><td>Mechanism for resolving COI lies with Director of EVAL</td></tr>
<tr><td>✓</td><td>Approval of all evaluation reports remains with EVAL</td></tr>
</table>

Opportunities for strengthening

Ensuring independence and impartiality through a range of mechanisms: The 2016 IEE recommended that the ILO transitioned REOs to become full staff members of EVAL to enhance the independence of decentralized evaluations. This to date has not happened, however the REOs have a technical reporting line to the Director of EVAL. The self-evaluation conducted by EVAL states that the lack of structural independence of the REOs requires EVAL to quality assure all evaluation products delivered by regional offices: *"Overseeing the quality and independence of all decentralized evaluations puts a heavy toll on the small Evaluation Office team based in HQ."* The OECD/DAC (OECD/DAC 2010) perspective on independence stipulates that the evaluators are independent from the development intervention, including its policy, operations, and management functions, as well as intended beneficiaries. The IEE evaluation team suggest that independence and impartiality can be upheld and protected by the use alternative mechanisms and processes such as independent evaluation governance oversight bodies or reference groups as per UNEG norms and standards or through quality assurance helpdesk services. The volume of evaluations, combined with limited resourcing within EVAL, constrains EVAL's capacity for strategic support across departments. The IEE team suggest that for EVAL-led evaluations (centralized evaluations) the use of real time quality assurance undertaken by persons or organisations independent of EVAL would represent the 'gold standard' in quality assurance and maximize impartiality.

Rigid application of the norm of Independence risks undermining the utility of evaluation products and skills and creates inefficiencies: There was frustration expressed by some ILO staff that they would benefit from EVAL's expertise in key stages in the design of programmes. This would not compromise its independence: some UN and multi/bi lateral agencies are actively promoting more participatory approaches. For example, UNICEF promotes participatory approaches in its impact evaluations - agreeing lessons learned and recommendations (to improve accuracy and relevance), joint generation of recommendations to build on emerging impacts (or lack thereof) to improve project performance through active, adaptive implementation of the intervention by project staff and programme participants or those living with policy changes (Guijt 2014).

3.3.4 Resourcing dedicated to the evaluation function (both human and financial) to support the coverage, quality, and uptake of evaluations

Financial resources

The ILO evaluation policy sets a goal of approaching a combined evaluation expenditure of 1.5–2 per cent of total expenditures, as recommended in international evaluation standards. EVAL secures regular budget funding for the core ILO Evaluation Function, to ensure that the Evaluation Policy and strategy can be implemented, as required. Most ILO evaluations are financed from programme or project budgets. As per the ILO evaluation policy, a minimum of 2 per cent of total project funds should be reserved for all mandatory evaluations. This is not compulsory for smaller projects below $500,000 that only require a final progress report with self-evaluation components. ILO policies also recommend that resources be set aside for monitoring, collecting baseline data, and reporting and conducting evaluability assessments – a minimum of 3 per cent is recommended (EVAL 2020). Funds of the Regular Budget Supplementary Account (RBSA) complement evaluation activities at both centralized and decentralized levels.

The budget allocation for both centralized and regional evaluation accounted for just over 0.8 per cent of the total ILO budget each year of the strategy period. In terms of evaluation budget as a proportion of programme budget for 2020-21 this represents 1.28 per cent (US$496.6 million of regular budget was earmarked for the delivery of policy outcomes).

This allocation to evaluation has not changed since the previous IEE, which also found that based on the biennial budget for ILO 2014–2015 ($400.6 million per year), evaluation represented 0.8 per cent of total budgeted expenditure (i.e., excluding extra budgetary resources such as TC projects), table 12 shows the budget allocation for evaluation for the period of 2017-21. The 2021 MOPAN assessment rated the adequacy of resources for evaluation as "Highly satisfactory"[49]. Yet the ILO has not fully met its policy ambition of 1.5–2 per cent of total ILO expenditure.

Total allocation for Evaluation (RB, PSI, RBSA) $US				As % of total ILO budget
	EVAL	Regional	Combined allocation	
2016-17	3,084,032		3,084,032	
2018-19	4,498,319	1,756,889	6,255,208	0.80
2020-21	4,308,191	2,038,760	6,346,951	0.81
2022-23	4,277,250	2,124,290	6,401,540	0.82

---

[49] This refers to 8.1.4 Element 4: The central evaluation programme is fully funded by core funds which received a 'highly satisfactory' score. (MOPAN 2021)

ILO's evaluation allocation is on par with comparable UN organizations such as UNFPA who recently reported evaluation spends as 0.83 per cent of total UNFPA programme expenses (UNFPA, 2022). UNFPA aims to allocate a minimum of 1.4 per cent of its total programme expenditure to the evaluation function, up to a maximum of 3 per cent. For UNESCO, there is a nominal allocation of 3 per cent of programme budget for evaluation, 3 per cent of regular activity budget. Extra budgetary projects also 3 per cent of budget, providing room to advance especially on decentralized evaluations[50].

Evidence from other studies of funding for evaluation functions across UN entities reveal a range of budgets from **$60,400 to $14.1 million** and as a percentage of programme expenditure they range from **0.02 per cent to 4.58** per cent (taking a sample of 33 UN entities, not including the ILO) (UNGA\OIOS 2013). The 2014 Joint Inspection Unit (JIU 2014, 6, para. 77) reported a range from 0.5 to 3 per cent of **organizational** expenditure for evaluation to be considered as a benchmark. Variation is expected based on differences in the purpose of the evaluation function, the type of evaluation undertaken, and the economies of scale achieved, including as influenced by an entity's size.

Development Cooperation Project Evaluation: The IEE team estimate that a total of $6.46 million was spent on evaluation activities of DC projects over the period 2017-21. As the ILO does not have a dedicated finance expenditure code for evaluation, this data should be interpreted with care but as the only available data provided to the IEE team offers some insights.

Table 13: DC Evaluation Expenditure by region[51]

	Evaluation expenditure DC projects						All DC project exp 2017-21 $m	% evaluation spend as total project spend (indicative)
	2017	2018	2019	2020	2021	Total 2017-2021		
Global	370,989	492,644	286,611	275,152	362,008	1,787,404	248.84	0.7
Asia and the Pacific	265,391	355,659	375,361	350,254	374,193	1,720,858	389.86	0.4
Africa	294,344	264,227	286,283	339,001	379,452	1,563,307	345.53	0.5
Americas	36,475	118,923	120,318	75,207	185,032	535,955	144.12	0.4
Arab States	1,507	100,446	159,360	77,532	138,904	477,749	206.21	0.2
Europe and Central Asia	104,142	1,559	128,368	86,169	55,560	375,798	111,47	0.3
	1,072,847	1,333,458	1,356,301	1,203,315	1,495,149	6,461,070		
Average DC project evaluation expenditure	19,506	15,688	13,168	11,570	15,257			

---

[50] Interview with UNESCO
[51] Expenditure data provided by ILO finance cleaned by EVAL, analysed by the IEE team

There has been a growth in DC evaluation expenditure in Africa, Americas and Arab States and a decline in Europe and Central Asia over the evaluation period, with Asia and the Pacific accounting for 26.6 per cent of all DC evaluation spend.

Figure 12 DC Evaluation Expenditure by Region
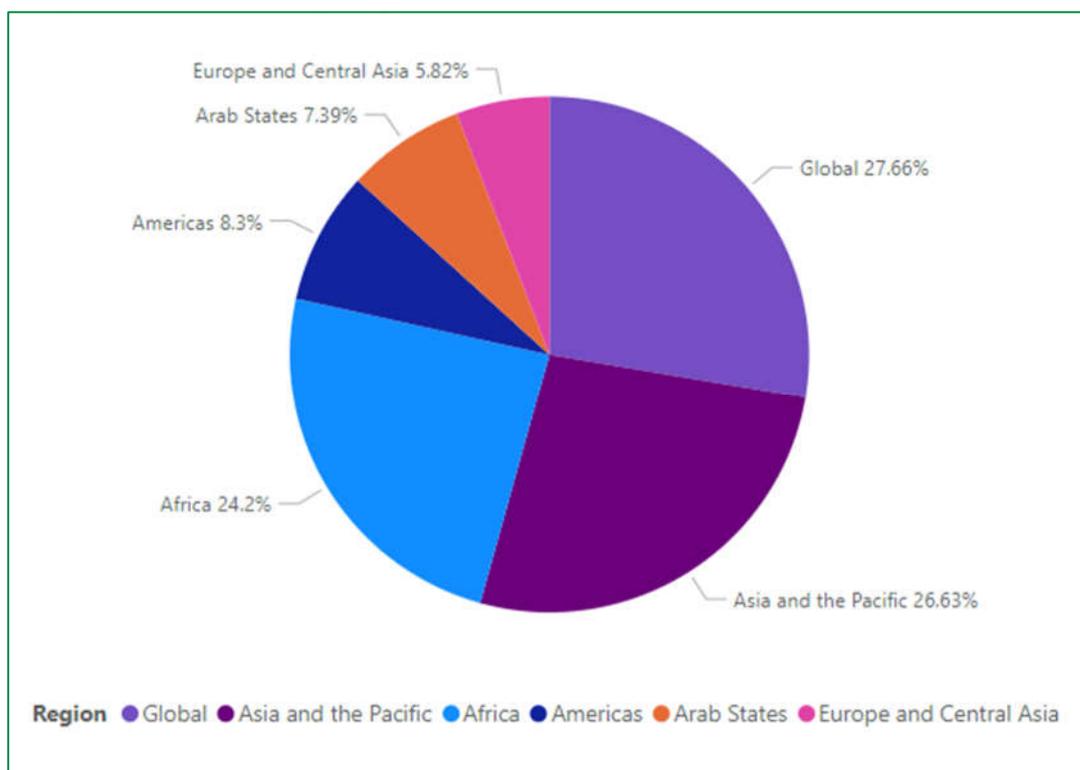
The median expenditure per DC project on evaluation varied between $11,570 to $19, 506 with a decrease in 2020 that is likely attributable to impacts of the Covid 19 pandemic. These figures illustrate the prevalence of small project evaluations taking place across the ILO. Such small evaluation budgets will restrict the ability to measure impact, identify ILO's specific value add, or deal with complex long-term change. The current evaluation policy requires a more efficient and strategic use of extra-budgetary evaluation funds, such as integrating budget sources and clustering evaluations whenever strategically and procedurally possible. In addition, the policy aims for the establishment of an evaluation trust fund to pool evaluation funds for smaller projects, but this has not yet been achieved.

Human resources

The ILO has a hybrid (centralized/decentralized – dedicated evaluation roles/voluntary evaluation roles) system (see Section 2.3).

What is working well

The ring-fenced funding for programme evaluations is a strong feature of ILO's evaluation function, securing guaranteed resources. The resources for evaluation as a proportion of total programming spend is on par with other UN entities, along with evaluation resources as a proportion of total organisational budget. The ILO meet the JIU 2014 target of between 0.5 per cent and 3 per cent of organizational expenditure for evaluation.

EVAL has exceeded its target set for 120 certified Evaluation Managers by the end of 2021 and created a database that contains a list of all EMCP trainees, access to the evaluation

report(s) that they managed, and the ex-post quality appraisals associated with them. This facilitates the identification of evaluation managers.

Each HQ department has a person assigned as DEFP, responsible for coordinating mandated evaluation activities with EVAL. However, there is no official training for DEFPs.

The REOs are the **keystones** of the decentralized evaluation system. The value of the REOs in supporting the decentralized system was expressed strongly by many respondents; their hard work and heavy workload is recognized.

Opportunities for strengthening

Organizationally inflexible evaluation funding mechanisms: The dedicated resources for project evaluations are a strong feature of the ILO's evaluation approach that should be protected to ensure adequate evaluation coverage. However, there is room to introduce more flexibility into the resourcing models for evaluation. Over the strategic period EVAL has presented a strong case to the senior management team, EAC and the GB for more cluster evaluations as a way of generating evaluative evidence in a more strategic and efficient manner. However, challenges remain on the procedural access to these cluster funds -with hurdles reported in the finance mechanisms (e.g., harmonizing multiple funding partners, securing funding). A cost-recovery scenario to compensate for staff time devoted by evaluation managers to evaluation-related tasks has been developed and is due to be discussed with the ILO management as part of the actions following this IEE. Additionally, despite EVAL's policy drive to the creation of pooled resources, resources for ex-post or longitudinal evaluations this remains a struggle with a lack of organisational appetite or financing structures that inhibit such flexible use of funds.

Lack of real-time tracking of evaluation expenditure: There is a need for an improved online real time system for the systematic tracking of the ILO evaluation budget and expenditure. EVAL does not have ready access to evaluation expenditure to keep a track of policy adherence or to aid evaluation planning.

Weak support structure and incentives for evaluation managers and DEFPs: The decentralized evaluation function continues to face challenges in matching over 50 independent evaluations per year with evaluation managers. Although there is a community of practice for evaluation managers, there is little evidence of an active peer-to-peer community among the network's members. A survey of attendees on an IECP training course found that 94 per cent of interviewees agreed that incentives should be linked to the participation in the programme, these incentives could cover: a recognition in the individual's Performance Appraisal, including a monetary bonus (i.e.: step increment); foster an active community of practice with webinars and coaching from EVAL. There are few incentives or capacity for evaluation managers to manage more than one evaluation. The previous IEE also identified the need to strengthen incentive structures for departmental evaluation focal points (DEFPs), evaluation managers and internal evaluators to manage and conduct evaluations.

A stretched decentralized staffing structure: Whilst the regional architecture is well established, with REOs as a critical piece of the evaluation commissioning and quality assurance machinery, dual reporting lines have not been fully smoothed out. REOs are overwhelmed; the Director of EVAL recognized the "crushing workloads" of some of the REOs[52] with both evaluation and regional responsibilities. For the REOs, the majority of their evaluation management skills are directed to overseeing evaluations for projects less than $5 million – accounting for nearly 56 per cent of their evaluation portfolio (see fig 9).

---

[52] Regional conference call minutes June 21

Figure 13 Type of evaluations REOs are managing
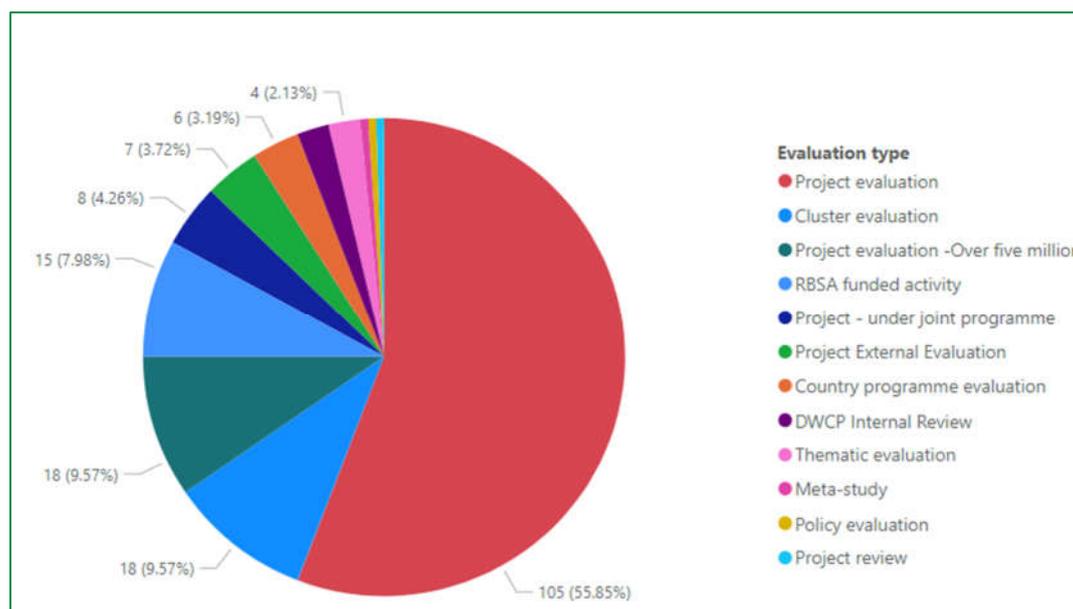
Table 14: Evaluations managed by Regional Evaluation Officers (2017-2021)

	Total number evaluations 2017-2021	Total managed by REO	%	Average number of evaluations managed by REOs/year
Arab States	25	20	80%	4
Africa	122	40	42%	8
Asia and the Pacific	123	53	43%	10.6
Latin America and the Caribbean	57	51	89%	10.2
Europe and Central Asia	38	21	55%	4.2

Heavy workload for EVAL staff – the policy requirements and quality assurance mechanisms place a strain on the staff of EVAL in terms of volume and pressure to deliver on time. Suggestions for reducing this workload could include reducing the number of low value project evaluations undertaken, allow REOs to undertake final quality assurance on evaluations (this would be more manageable with a lower volume of evaluations being conducted) and where necessary bringing in peer REO quality assurance, evaluation reference or advisory groups, or the use of real-time external quality assurance services on a 'call down' basis. It is worth noting that WFP, UNICEF and UNHCR use these types of demand-led services. The reliance on voluntary evaluation managers and a small number of REOs adds a level of fragility to the system, which put the quality, credibility, and utility of evaluations at risk.

REOs and ILO staff reflect that it can be challenging to find evaluation consultants with the appropriate balance of evaluation expertise and relevant technical skills. Particularly in regions with a higher volume of evaluations (i.e., Africa, Asia and the Pacific) there is a shortage of "good" evaluators to conduct ILO evaluations. This, combined with finite budgets, requires that the evaluation manager (as well as the REO and in some cases Senior

Evaluation Officers in EVAL), invests additional time for oversight and support of evaluators to ensure a quality product.

3.3.5 Capacity to conduct and use evaluations

What is working well

EVAL has specifically focused on enhancing mechanisms for knowledge management and communications. Communications for ILO's evaluation function are supported by one dedicated staff person. Guided by its 2018-21 Communications Plan, EVAL aims for enhanced outreach, more targeted, useful communication products. With this focus, EVAL provides regular communication of its progress against outcomes through a wide range of media which are designed to facilitate interaction with evaluation products, provide targeted information, and to offer relevant updates on evaluation activities. Communications are tailored to specific user groups and actively seek to engage ILO's evaluation network.

Knowledge and communications: A hallmark of EVAL's knowledge management is the *i-eval* Discovery knowledge platform which provides public access to all planned and completed evaluation reports from 2011 onward. The platform provides a repository for all evaluations and collates recommendations, management responses, lessons learned and good practice which has enabled EVAL to enhance the accessibility and visibility of evaluation products through the database.

EVAL's knowledge and communications are prolific. EVAL has significantly expanded the diversity and quantity of knowledge products and processes it supports. In addition to its evaluation guidance and evaluation reports, currently EVAL produces regular newsletters, synthesis of results, factsheets from evaluations, synthesis reviews and meta studies, and has conducted a series of webinars. Guided by a focus on enhancing the use and dissemination of evaluation information, EVAL tailors communications through targeted listservs.

Stakeholders surveyed are generally positive about the information that they receive about evaluations with 79.3 per cent of all survey respondents reporting that they receive evaluation products with the right frequency. Notwithstanding the frequency of communications, the range of products provided and the multiple opportunities to engage with evaluation, only 25.3 per cent of survey respondents reported that evaluation knowledge is "largely" shared and mainstreamed throughout the organization (fig 12). This differed significantly for those staff members that are actively involved in evaluations (evaluation managers, REOs and DEFPs). Of this group, 40.63 per cent responded that evaluation knowledge is "largely' shared and mainstreamed throughout the organization. Stakeholders feel that they receive the right amount of information from ILO's evaluation function. Nevertheless, mainstreaming evaluation as a source of knowledge across the ILO will take a long-term

Figure 14 Evaluation Knowledge (IEE survey results)

To what extent is evaluation knowledge shared and mainstreamed throughout the organization?

Largely   25.3%
   63.0%
Not at all   11.6%

effort that appropriately balances the interests and absorption capacities of target audiences.
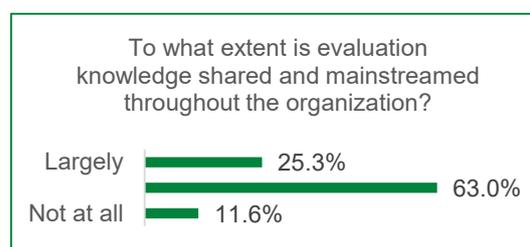
EAC: The Evaluation Advisory Committee (EAC), which has representation from across the ILO (Policy Portfolio, programme, regions) was established to oversee the use, implementation and follow up to lessons and recommendations from ILO evaluations. EAC minutes demonstrate a high level of engagement and interest in evaluations from across the ILO. The EAC offers an example of good governance for supporting transparency and uptake of evaluation recommendations to maintain an independent, credible evaluation function.

According to ILO staff, the need for the EAC is an indicator of the degree to which the actions and recommendations from evaluations have been satisfactory.

Prioritization of evaluations: The question of whether to conduct evaluations is not asked as part of programming, although evaluability assessments are required for projects above USD 5 million. Instead, the basis for project evaluations is triggered by the project budget, determined by the evaluation policy. EVAL has developed guidance to determine when cluster evaluations can be conducted, including the need for administrative clearance by both PARDEV (donor agreement) and EVAL (technical agreement).

Joint evaluations: Given the emphasis on joint programming across the UN and the push toward harmonized approaches to meet shared global commitments, joint evaluations are increasingly important. When ILO is the lead agency in a joint project or programme, it is an EVAL requirement that ILO also takes the lead in the joint evaluation process.

Evaluability of ILO projects and programmes: good evidence relies on evaluable projects and programmes. Responding to challenges in the evaluability of ILO's interventions, EVAL has piloted tools to enhance evaluability of projects - including an evaluability diagnostic instrument for DWCPs and co-developed UNEG guidelines for the evaluation of UNSDCF to support consistent monitoring and evaluation across UN agencies. EVAL updated its guidance note on the Procedure and Tools for Evaluability in 2020. The tools developed by EVAL which articulate the purpose, process, and guidelines for evaluability assessments are clear and user-friendly. They hold great promise if they will be effectively disseminated and used across the organization. Currently evaluability assessments are supported by REOs upon request.

Opportunities for strengthening

Communications and products: Although evaluation guidance suggests that customized products may be targeted for specific constituent groups, the evaluation did not find evidence that knowledge products have been specifically tailored for workers, employers, or governments.

Knowledge management: The i-eval discovery platform is a well-established and comprehensive repository for evaluations attracting 8,893 users during 2021[53], however, some ILO staff interviewed said that while the repository of evaluation is comprehensive, they felt it wasn't used as much as it could be. Project evaluations are most often consulted for the second phase of an existing project but are not systematically used for the design of a new project.

EAC: Given the importance of the EAC, the 2016 IEE recommended that regional EACs be developed. So far, the Americas region has taken steps to establish a Regional Evaluation Advisory Committee. Given the large numbers of evaluations coming through the pipeline each quarter, the ability to fully absorb the findings and implications of the evaluations is challenging for EAC. This limits the degree of strategic engagement from the EAC.

Funding for cluster evaluations: The approach to evaluation does not include a provision for a reallocation of resources from project evaluations toward cluster evaluations. EVAL must approve all decisions for cluster evaluations: this may inadvertently create bottlenecks for a more efficient prioritization process.

Joint evaluations: Where ILO is not the lead agency, the decision to conduct an evaluation jointly or solely by the ILO is made on a case-by-case basis. Often this comes when the ILO has a particular accountability need for visibility of ILO's role in a project. In these cases, ILO conducts an evaluation of its contribution to the project which can be challenging to

---

[53] EVAL data provided by ITCOM

coordinate across the demands from the wider joint evaluation and places additional demands on REOs.

Evaluability: To date there has been limited use of the evaluability diagnostic instrument developed by EVAL. In addition, monitoring data is frequently incomplete from ILO interventions, which compromises the value and credibility of evaluations to generate new knowledge and insights that can be used across the organization. There would be significant value in positioning EVAL to support programming, for example by offering their expertise in evaluability assessments.

3.4 To what extent does the evaluation function respond to international, organizational, and tripartite policies, interests, and information needs?

> **Key finding 4:** The ILO's evaluation function is meeting the expected norms and standards for credibility, quality and independence but has not yet attained fully the utility expectations. EVAL is an active and valued member of the UN evaluation community and needs to respond to UN calls for more joint evaluations and harmonized working. The ILO would benefit from a sharper strategy (including in the choices of what to evaluate) to meet the varied tripartite evidence needs, harness their interest in evaluation evidence, and reap the gains from fuller engagement from its constituents.
> Strength of evidence: Level 3

3.4.1 Alignment with international norms and standards and contribution to UN evaluation practice

The IEE examined the extent to which the evaluation function aligns with international norms and standards and contributed to UN evaluation practice. It looked specifically at implementation of the UNEG norms and standards 2016 (see list provided in Annex M) and the extent to which the OECD/DAC and UN SWAP norms and standards have been taken into consideration.

What is working well

As presented earlier in this report, the ILO is meeting the expected norms and standards for credibility quality and independence but has not yet attained fully the utility expectations. The IEE undertook an assessment of the maturity of the ILO's evaluation drawing on a tool developed by UNEG (UNEG 2020)[54].

Figure 11 presents an assessment of the maturity of the ILO's evaluation function against a set of defined performance criteria and maturity benchmarks graded 1-4. The IEE team has conducted this assessment drawing on evidence presented in EVAL's Self-Assessment, MOPAN 2020, as well as other documentation, interviews, surveys and focus group discussions. The results are presented graphically, and a full breakdown of the assessment is provided in Annex J. The purple-coloured circles demonstrate where full maturity has been

---

[54] UNEG has drafted a set of 48 organizational and performance criteria to help assess the maturity of UN evaluation functions against the established norms for evaluation agreed in the UNEG Norms and Standards for evaluation. It draws on the assessment criteria and benchmarks used by UN entities such as the JIU and on earlier peer reviews of UN evaluation entities carried out under the auspices of UNEG.

reached (level four), the blue show level 3, green show level 2. The ILO scored level two or above in all indicators.

Figure 15 Results of Evaluation Maturity Assessment (key provided in Annex J)
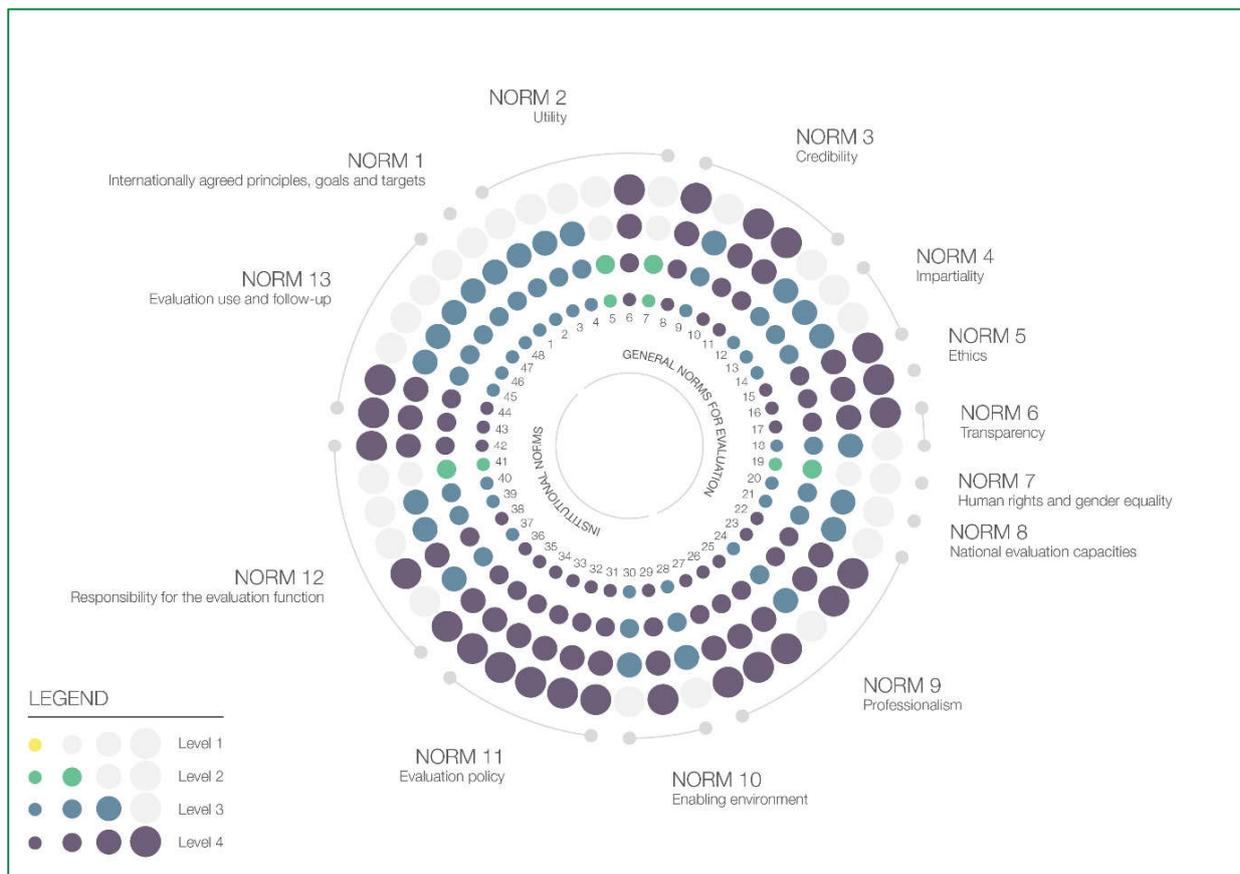
The strengths demonstrated by the ILO evaluation function are particularly in the evaluation policy, professionalism, credibility, and recommendation tracking. The full assessment is provided in Annex J, and the wider recommendations presented in this report have been informed by this assessment. The IEE team concluded that the ILO is almost reaching the highest level of evaluation maturity, reaching level 4 in 23 of the indicators (47 per cent) and level 3 in 21 indicators.

The self-assessment shows a contribution to several UNEG-wide evaluation related initiatives and working groups for example:

- working group on UNDIS indicator on evaluation and guidance for evaluations.

- Gender, Disability and Human Rights: EVAL actively contributed to the work done on the technical notes for the UN-SWAP assessments and the UNDIS indicator framework.

- *Evaluating Policy Influence and Normative work subgroup,* Evaluation Practice Exchange Organizing Committee.

- UN Cooperation Framework (UNCF) Working Group, various UN Evaluation Initiatives for e.g., COVID-19 Global Evaluation Coalition.

- EVAL served as co-convener for the initial sub-working group on Guidance and Knowledge sharing that worked on a repository for Covid19 evaluation documentation as part of the UNEG website. ILO presented at two Evaluation Practice Exchanges on the

planning and management of evaluation during Covid19 and on current experience with evaluation of Covid19 response.

- In addition, the ILO has conducted 9 joint evaluations including with the UNDP, UN Educational, Scientific and Cultural Organization (UNESCO) and other UN agencies.

- In line with the United Nations Funding Compact commitment to increase accessibility of strategic evaluation results, the ILO made 4 per cent of centralized evaluations available on the UNEG website.

- In the ILO region offices, collaboration has taken place with local/regional UN entities. For example, in Asia the ILO is part of the UN Development Group for Asia and the Pacific (UNEDAP). It jointly conducts training for all UN staff in the region once a year, supports UNRC in the UNDAF/UNSDCF evaluations in collaboration with DCO, and undertakes joint studies with fellow UNEDAP members. In country and subregional offices, joint programs help to strengthen strategic relationships with other UN agencies, and evaluation exercises themselves fosters reflection and exchange with others.

Opportunities for strengthening

Areas to strengthen for evaluation maturity are:

- To reach higher evaluation maturity, the ILO would need to ensure that gender equality values are respected, addressed, and promoted with gender and diversity considered in a systematic way across all evaluation work (including evaluation process, findings, conclusions, recommendations). This could achieve a rating of 'exceeds requirements' in the UN System-Wide Action Plan. However, the ILO was rated as "approaching requirements" in the last assessment. Section 3.2.2 sets out how the external quality assessment of 2020 evaluations found that only 8.7% of evaluations that year scored as "meets requirements", with failures primarily in integrating gender equality and the empowerment of women (GEEW) in the scope of the evaluation investigation, inclusion of gender responsive methodologies, and presentation of gender considerations in the findings and recommendations. Section 3.4.3 presents a further analysis on the extent to which the evaluation function has been responsive to gender as a policy driver.

- In order to achieve the highest rating in the maturity assessment the ILO would need to be systematically sharing evaluation results and lessons learnt with other UN organizations and external stakeholders (including other evaluators, Members States, beneficiaries, etc.). This was recognized in the EVAL self-assessment as an area for improvement: "*only on rare occasions are evaluation results and lessons shared with other UN agencies and multilateral partners (most recently as part of the OECD-DAC COVID-19 coalition group)*". The IEE team found that only ILO evaluations up to Aug 2019 are included on the UNEG evaluation database[55]. That said, EVAL has made regular presentations about its work to the UNEG Evaluation Practice Exchange.

- For the maturity criteria of "effect of evaluation use on organizational effectiveness and evidence of impact", the IEE team assessed the ILO as scoring level two. To reach level four the ILO would require a comprehensive set of evidence that shows significant impact of the effect of use of evaluations on organizational effectiveness.

- Joint evaluations: MOPAN indicated that although ILO is actively undertaking joint evaluations, the visibility of this activity is limited amongst the partners surveyed. In addition, in contrast to other agencies, ILO does not have an explicit outcome area for joint evaluations. In its most recent evaluation report, UNFPA stated that 57 per cent (8

---

[55] http://www.uneval.org/evaluation/reports. EVAL provided a full list of all 2019, 2020 and 2021 evaluations to UNEG secretariat in May 2022

out 14) of centralized evaluations are either joint or system wide. They have set an aspirational target of 58 per cent of the proposed evaluations in the next four years to be either joint or system-wide, and deliver an estimated cost-saving of almost $2.5 million (UNFPA 2020)[56]

3.4.2 The extent to which evaluation is inclusive, relevant, useful, and responsive to the needs of the ILO member states, constituents and key stakeholders

The degree to which evaluations can assess the value and relevance of ILO's interventions, as well as reflect the needs of ILO member states, employers', workers' organizations, and other key stakeholders is highly dependent on the strength of design, delivery, and inclusion of constituents in the interventions themselves. This is evident in the records of high-level strategic evaluations discussed at the Governing Body, in which the importance of evaluations is discussed to build the capacity of ILO's constituents to engage in social dialogue and strengthen tripartism. They offer opportunities for discussion based on objective, independent and strategic reviews of ILO's work, rather than negotiating on sensitive issues in which consensus needs to be built between divergent views. Where there is an evaluative culture that prioritizes learning from challenges as well as successes, there is space to discuss in a more open, transparent way on what works and what doesn't in achieving goals and involving national and local constituent groups in this kind of dialogue.

What is working well

Constituents are consulted on ToR and draft reports. Questions related to tripartism, and constituents are always considered in evaluations. There has been a good uptake of capacity building opportunities amongst constituents. High-level evaluations are discussed at the highest level of decision-making in the ILO – both at Governing Body, and in the EAC, representing senior decision-makers in the ILO. The findings, recommendations and lessons learnt draw on the views of constituents - especially evaluations of large and complex programmes across multiple countries and policy areas (including meta-studies, surveys, interviews, studies, and reports). Overall, tripartite constituents[57] that took part in the survey for this evaluation found products and services of the evaluation function to be inclusive, relevant, useful, and responsive to their needs.

The evaluation of the ILO's flagship social protection programme (2012-2017) is an example of a programme that is recognized as highly relevant at global, country, and regional levels in delivering SDGs and the objectives of the Global Partnership on Universal Social Protection (USP 2030), and the opportunities presented to strengthen tripartism and the capacity of constituents. The HLE made recommendations focused on enhancing the capacity of tripartite constituents through comprehensive long-term capacity-building approaches that respond to constituents' needs at the country level; the need to enhance internal coherence of the ILO's social protection work by "fostering a dialogue to develop a common understanding and vision of implementation of the social protection agenda across headquarters, regional offices and field offices, and the need to deepen the use of the ILO's specific strengths, such as tripartism, social dialogue, rights-based approaches and gender equality as the core of its brand". During COVID-19, social protection rose considerably

---

[56] In contrast, a recent UNEG peer review UNESCO's evaluation function (2020) noted a lack of progress on joint evaluations and the panel recommended that the new policy and strategy "Reflect the emerging UN reform context and spell out the implications for evaluation at corporate and country level of the new UN Sustainable Development Cooperation Framework, including in joint and system-wide evaluations".

[57] 89.6 per cent of all survey respondents considered the products and services of the evaluation function to be 'largely' or 'moderately' inclusive, relevant, useful, and responsive to the needs of tripartite constituents. Among respondents to the survey specific to GB members and tripartite constituents that number reaches 83 per cent.

higher on the global and national policy agenda, as measures needed to be put in place quickly to mitigate the impacts of lockdown. The ILO was well placed to step in benefiting governments, workers, and employers equally.

Opportunities for strengthening

Timely engagement: The most appropriate and efficient ways to ensure consultation and inclusion of the views of constituents and key stakeholders is at the design stage of evaluations as well as when initial findings and recommendations are shared for comment. Planning of evaluations need to ensure there is enough time to discuss the key questions for evaluations that are of interest to key stakeholders and allow sufficient time for discussing findings and integrating constituent comments. Some interviews reflected that while constituents are consulted in evaluation processes, there were occasions when their inputs were not integrated into evaluations due to lack of time (e.g., HLE on gender). When this has happened, they have found it more difficult to engage their constituents in future evaluation exercises. Frustration about 'gate-keeping' (having to observe protocol and working through intermediaries whose own time and capacities were extremely stretched), was expressed. This resulted in delays and missed opportunities to respond in a timely way to concerns which could offer opportunities for course corrections – e.g., in mid-term evaluations.

Innovative and appropriate ways to involve constituents and beneficiaries could increase ownership and learning from evaluations – for instance by innovating beyond standard forms of consultation – text-based comments on ToR and draft reports. For example, there could be more use of in-person or virtual workshops than paper-based consultations.

3.4.3 The extent to which evaluation is responsive to ILO's strategic priorities and cross-cutting issues

The evaluation policy highlights gender equality and non-discrimination as one of the six core principles that need to be considered in evaluation design, analyses, and reporting in addition to UNEG gender-related norms and standards. The ILO policy guidelines for results-based evaluation further highlight disability as another cross-cutting driver for ILO's work. This aligns with the ILO Decent Work Agenda and Sustainable Development Goals (SDGs). As highlighted above, the ILO has an obligation to report results on indicators related to the inclusion of gender issues in evaluation reports in the UN SWAP.

What is working well

Figure 16 Percentage of reports that received a score of "highly satisfactory" and "satisfactory" in 2020, and comparison with results from 2019. Source: EVAL Self-Assessment Report 2022
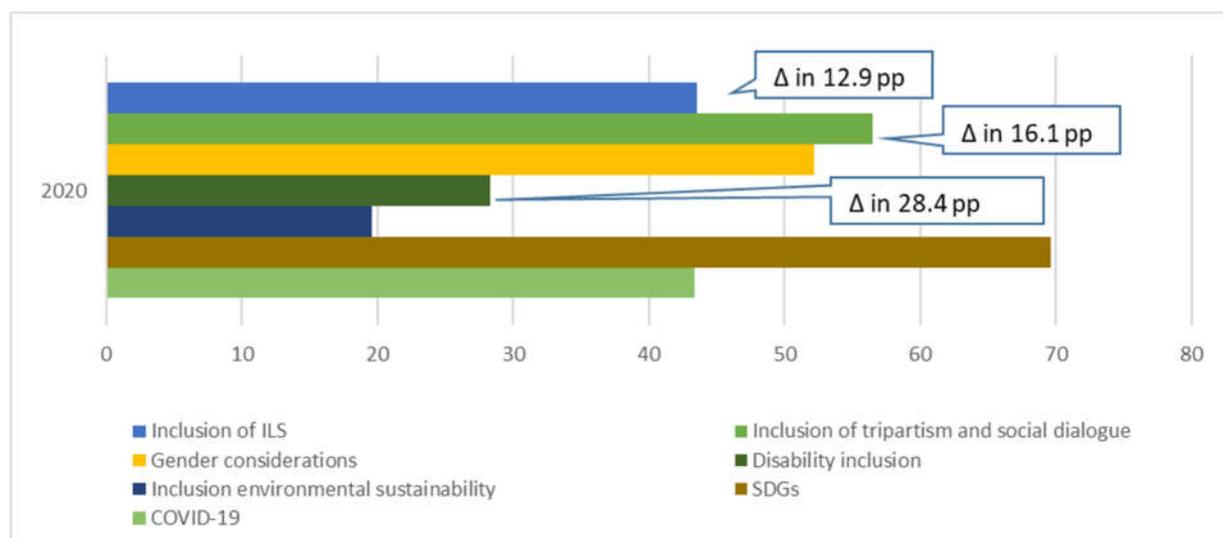
To ensure that evaluations are more responsive to the ILO's normative and tripartite mandate, EVAL updated its guidance note on integrating gender equality in monitoring and evaluation in 2020 in addition to revising existing guidelines to better integrate gender equality and non-discrimination issues (including disability), the SDGs and environmental concerns (including in quality control and checklists) (MOPAN 2021).

The 2020 ex-post quality appraisal showed progress in the integration of the SDGs in evaluation questions with 70 per cent of reports obtaining a "satisfactory" or "highly satisfactory" rating. Progress has also been made in the integration of gender considerations with 85 per cent of reports reached "somewhat satisfactory" or above ratings, which represents a 28-percentage point increase in satisfactory ratings between 2019 and 2020. A survey conducted with evaluation managers showed that the guidance produced on gender mainstreaming in evaluations was considered as extremely helpful by 53 per cent of respondents with 60 per cent considering that the relevance and level of clarity of the information it provided was satisfactory or extremely satisfactory and 53 per cent considering its applicability as satisfactory.

Opportunities for strengthening

Despite policy commitments and the development of guidance and more attention paid to gender in quality appraisals of evaluation TORs and reports, it is not yet fully mainstreamed throughout evaluation practice and ILO's evaluation function falls short of meeting UN-SWAP

requirements (UN WOMEN, 2020). Progress has been made towards the UN-SWAP targets for evaluations and in 2020 the ILO evaluation function approached requirements in a positive trend developing since 2016 (UN Women 2021). The integration of gender in evaluations is facilitated by the gender-

Figure 17 Trend in UN-SWAP ratings for ILO decentralized evaluations. Source: Annual Evaluation Report 2018-19.
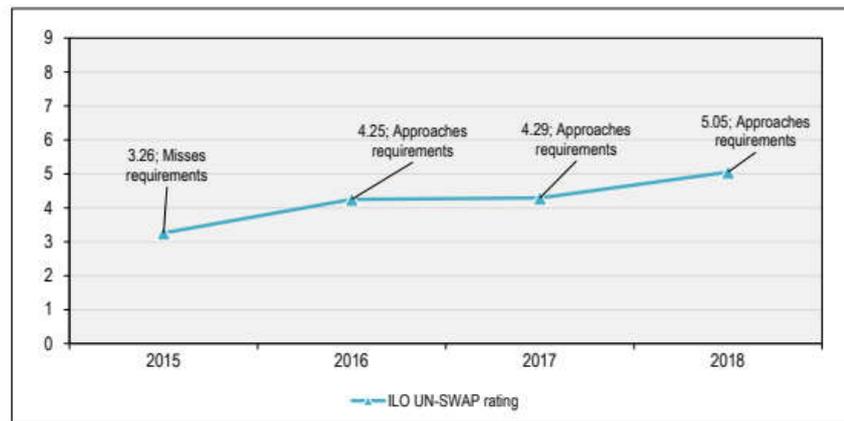
responsiveness of programming. There is room for improvement by encouraging more exchange of gender-sensitive data, results and learning across portfolios and departments and regions.

The results from the monitoring of the implementation of the ILO Action Plan for Gender Equality 2018-21 show that only two out of five targets for the evaluation function were met (ILO, 2021). The review of GEEW mainstreaming in evaluation reports from quality assessments of evaluations undertaken in 2020 show that integration of gender is uneven with regards to evaluation questions, methodologies and that gender is not always considered in findings, conclusions, and recommendations. Gender mainstreaming at design and proposal stage of projects and programs, gender-sensitive budgeting and coordination with specialized agencies would strengthen the inclusion of gender in evaluation processes.

Table 15: Monitoring results of the Implementation of the ILO Action Plan for Gender Equality 2018-21

Targets met in the evaluation function: 2 out of 5
• Four initiatives integrated gender into the Evaluation Office's networks (result: 10 which shows that gender is increasingly taken into consideration) • One evaluation every five to eight years on organisational performance on gender mainstreaming (The HLE on Gender Mainstreaming was conducted in 2021).
Targets not met in the evaluation function: 3 out of 5
• 65 per cent of sampled evaluations whose scope of analysis, and criteria and questions integrated gender equality (result: 51 per cent) • 50 per cent of sampled evaluations integrated gender-responsive methodology, methods, tools and data analysis techniques (result 31 per cent) • 65 per cent of sampled evaluations integrated findings, conclusions and recommendations reflecting a gender analysis (result: 51 per cent)

Disability inclusion: Although, EVAL actively contributed to the development of UNDIS Indicator Framework, and pioneered guidance on disability inclusion, nevertheless, this remains an area with scope for improvement. There is no specific guidance on inclusion of disability concerns in evaluation processes and report. The MOPAN Assessment reported that "The ILO increasingly produces results that target disability and non-discrimination, but these are not yet fully visible across evaluation reports. The organization has made progress in incorporating disability into policy and programming, but results are less evident at the country level" (MOPAN 2021). The Annual Evaluation Report 2019- 20, and the ex post meta-

analysis of development co-operation in decent work results for 2019- 20 demonstrate that less than 30 per cent of projects were rated "successful" or "highly successful" in linkages with disability inclusion. The Evaluation Office has reported that disability inclusion as well as the SDGs are areas that need to be strengthened in the evaluation methodological framework through the development of guidance and review of evaluation products (ILO, 2020)

3.5 To what extent is the evaluation function able to respond and flex to emerging priorities and stressors?

Key finding 5: The evaluation strategy provides a clear strategic vision and outcome areas that have potential to support a transformative approach. The policy provides for a consistent but less nimble approach to evaluation coverage and prioritization.
Strength of evidence: Level 3

3.5.1 COVID-19

The coronavirus pandemic created unprecedented constraints on the conduct of evaluations in ILO with the implementation of strict travel restrictions and lockdowns leading to the introduction of remote working on a wide scale. The ILO's evaluation framework and activities as well as results targets were adapted to support the changes in the operating model and the reallocation of resources and priorities (MOPAN 2021)

What is working well

The evaluation office was able to adapt and adopt flexible approaches to evaluation to support the delivery of ILO's Policy framework for tackling the economic and social impact of the COVID-19 crisis (ILO 2020). Building on existing reflections around reducing travel for evaluation based on budgetary and environmental concerns, EVAL quickly developed operating procedures on implications of COVID-19 on evaluations in the ILO at the start of the pandemic in 2020(ILO 2020).This guidance was completed by a protocol on collecting evaluative evidence on the ILO's COVID-19 response measures through project and programme evaluations published by EVAL in October 2020 (ILO, 2020).

In compliance with ILO's policy framework and the adapted protocols, evaluations have been required to systematically assess ILO's response to the pandemic through the consideration of COVID-19 in the development of terms of reference[58], while adapting methods and mitigating risks to produce credible evaluations to inform decisions.

To adapt to the restrictions of international missions, engagement of national consultants for country-level evaluation activities was encouraged as well as an increased use of online communication tools and platforms, including for simultaneous translation services. The respondents acknowledged the efforts made to include issues related to the pandemic response in evaluation questions. Evaluation managers found useful guidance on protocols and adapted ToR to include the context of the pandemic, including extending timeframes due to challenges in data collection. In addition, EVAL produced new communication products, such as podcasts for training events and material adapted to the context of COVID-19 (ILO, 2020). One example is the special issue of the i-eval In-Focus internal learning series of ILO's response to the impact of COVID-19 on the world of work (ILO, 2021).

---

[58] Checklist 4.6: Writing the Evaluation Terms of Reference was adapted to include the COVID-19 response as an area for assessment and made reference to OP No.2

Opportunities for strengthening

The protocols and guidance developed by EVAL at the start of the pandemic were circulated to evaluation managers and evaluators. 62 per cent of respondents in the survey to evaluation managers, departmental focal points and regional evaluation officers showed awareness of these tools. However, 11 individuals (33 per cent) in this respondent group (mainly evaluation managers and departmental focal points) did not know about the new evaluation guidance to response to COVID-19. Amongst the wider ILO staff surveyed, 67.4 per cent did not know about the new guidance and 61 per cent of EMs, DEFPs and REOs were interested in getting more support and advice on adapting evaluation in light of COVID-19. Although the new guidance was not aimed at a larger audience, evidence shows there is scope and demand to circulate it among a wider audience to encourage good practice in contexts of crisis and reduced direct access to the field.–This would contribute to an increase in inclusion in evaluation questions.[59]

3.5.2 The extent to which the evaluation policy and strategy are fit for purpose and flexible

The evaluation policy sets out principles for evaluation, reinforced through the strategy which identifies "concrete outcomes, milestones and targets". It describes the necessary enabling environment for evaluation to contribute to an effective ILO. The evaluation policy (2017) calls for more strategic evaluations, participatory evaluation approaches and methods, a framework for the evaluation of capacity development, strengthened independence for decentralized evaluations, and strengthened follow-up and use of evaluations.

What is working well

The evaluation strategy is broadly fit for purpose and sets out a strategic vision for the role of the evaluation function within the ILO. The ILO's contribution to realizing Decent Work is enhanced by evaluative evidence of high quality, using independent findings of the effectiveness and efficiency of programmes, and following up recommendations through strong accountability processes and partnerships that measure its contribution to the SDGs.

The Evaluation Policy guidelines establishes clear guidelines for conducting project evaluations. In 2011, the financial threshold for conducting independent evaluations was increased from $500,000 to $1 million. The increase ILO's threshold for project level evaluations is lower (for smaller projects) than other agencies (table 16), and the guidelines are budget-based rather than demand-driven. According to EVAL, setting financial thresholds is part of a risk management approach. Where less money is involved, there is lower risk if an independent evaluation is not conducted.

Table 16: Financial threshold for independent evaluation

ILO	UNESCO	UNFPA
All projects with budgets of USD 1 million and higher [60]	All projects with budgets of USD 1.5 million and higher	No financial threshold

---

[59] ILO is currently conducting a high-level evaluation on COVID-19. The IEE team in agreement with TC and EVAL made the decision not to replicate the aims of that evaluation for cost-effectiveness. More insights will be provided on the inclusion of COVID-19 in the high-level evaluation.

[60] ILO: Over 5 million: M&E plan appraisal tool, Independent Mid-term and Final independent. Mandatory: evaluability assessment; 1-5 million Final independent evaluation (and mid-term if over 18 months); 500-1 million Final Internal evaluation); Under 500000 Final self-evaluation (part of final project progress report)

Evaluation coverage: The clarity and extensive coverage that the guidelines offer to project evaluations is commendable. As noted in the MOPAN assessment (2020), the clarity of the policy and coverage criteria are a key strength of the evaluation function. The evaluation policy guidelines are applied in full. It is rarely possible to opt out of any aspects of the process.

Opportunities for strengthening

Coverage: There are drawbacks to taking a budget driven approach to coverage criteria. The project-based approach does not easily lend itself to discretionary evaluations or adaptation to emerging stressors. This means there is less scope for the timing and scope of evaluations to be adjusted and prioritized in response to emerging interests, strategic themes, country or regional issues. For projects over 5 million, the policy requires both a mid-term and final evaluation, even if the project has a shorter timeframe (e.g., 18–30-month duration), or where there have been delays in project start-up, This can, and has, created a logjam between project implementation and the conduct of the evaluation. In few cases, a final evaluation will take place when a project has ended and staff, whose posts are tied to project funding, are no longer in place. In some cases, reported by regions, problems were caused by project delays and no-cost extensions which affected planning, required multiple interviews of the same person, and caused delays in receiving documents. This created heavy burdens on the evaluation process and the stakeholders involved.

Volume: Analysis of the Development Cooperation (DC) projects revealed a body of 1242 projects with a start date between 2017-2021. Given the current evaluation policy requirements and guidelines, this would mean that 333 projects would have required at least one independent evaluation over the period (27 per cent of the portfolio), with a further 238 projects requiring a formal internal evaluation, and 671 projects requiring a self-evaluation. This policy requirement would produce a considerable amount of evaluative evidence, but arguably little capacity to make full use of it.

Taking the portfolio for the period 2017-2021 - if the threshold trigger of when an evaluation is mandatory is raised to either $2 million or $3 million, there would be a shift from the current 27 per cent of the portfolio independently evaluated to either 16 per cent or 10 per cent (equivalent to 206 or 129 projects) over the strategic period. Fewer independent project evaluations would mean that learning and accountability would need to be be supplemented by internal evaluation, self-evaluation, reviews, and strong monitoring. Freeing up staff to engage in more strategic evaluations, accompanied by a bolstered monitoring system, could move the ILO into the territory of a more mature evaluation function. It could demonstrate the impact of its evaluation efforts in helping the organization achieve its strategic mandate. (See Annex H). Alternative mechanisms (beyond project budget as trigger) to set the criteria on whether or not to evaluate are presented in the recommendations section of this report.

UNESCO: All extrabudgetary projects are subject to a form of evaluation; and spells out the form an evaluation should take. Self-evaluation is required for projects with budgets under USD 500,000; self-evaluation with external validation for projects with budgets ranging from USD 500,000 to USD 1.5 million; and external evaluation is required for all projects with budgets of USD 1.5 million and higher; UNFPA: Evaluations are identified based on strategic relevance, risk of the subject, potential for joint or UN system-wide evaluation, significant investment, feasibility, potential for replication or scaling up, knowledge gaps, and stakeholder commitments

Table 17: Evaluations per policy requirement mapped to DC Portfolio 2017-21 (analysis undertaken by IEE based on DC dashboard and i-EVAL discovery data)

Project budget US$	% Of DC project portfolio covered	Number projects	Average project budget	Evaluation required[61]
Over 5 million	6	72	$10.1m	Appraisal, EA, (+Mid>18 months), Final Independent
1-5 million	21	261	$2.18m	Independent evaluation (+mid when >18months)
500,000-1 million	19	238	$710,245	Internal evaluation (+mid>30mnths)
under 500,000	54	671	$211,250	Self-evaluation (+mid>30mnths)
	100	1242		

The ILO Evaluation Theory of change: The inclusion of a TOC in the evaluation strategy is welcome, however, its limitations could be overcome by including:

- detail about how the evaluation function will create a strong evaluative culture or how it will maximize the impact of evaluative evidence.

- a clear articulation of the evaluation policy aspiration of "more strategic evaluations of projects and programme activities under identical or similar themes, programme frameworks and locations by means of clustering and integrated funding".

- a clear articulation of the policy aspiration of "evaluation approaches, methods and frameworks that are participatory and people centered, are inclusive of disadvantaged workers, human rights and gender equality and are adapted to the ILO's specific mandate and context (for example tripartism, social dialogue, normative work".

- a clearer visualization and organization of the roles and contributions made by each element of the evaluation function with an ability to see EVAL's unique and enabling role.

- inclusion of lower-level activities, outputs and intermediate outcomes required to deliver on the evaluation strategy outcomes and higher-level impact,

- clearer articulation of what are the main features of 'evaluation culture'

- a reorganization of the enablers box – the current structure implies a logical pathway

- presentation of how challenges or barriers within the enabling environment may constrain the achievement of outcomes (see table 18 below)

- Ideally the ToC should be adapted so that it stipulates how it will contribute to the organization's strategic objectives (articulated in the ILO Strategic Plan 2022-2025) and higher-level outcomes.

Table 18 offers an illustration of the conditions that would need to be in place in the enabling environment to facilitate achievement of strategic outcomes.

---

[61] Definitions of internal: Formalized eval process managed by project management, can be conducted by ILO (no ties of COI with management of project) as well as external. >$1m recommended use of evaluation consultant. Independent: managed and carried out by someone free of control of responsibility for design and implementation of development intervention (process overseen by EVAL and its network)

Current Evaluation Strategy Outcomes	Conditions for the enabling environment
Outcome 1. Enhanced capacities and systems of evaluation for better practice and use	Evaluation is built into policy, program and project design
	Stakeholders are ready, engaged, and understand evaluation
	There are resources available to support evaluation at HQ and regional level
	There is a blend of incentives and requirements for use of evaluation
	Governance processes support use of evaluations
Outcome 2. Enhanced value of evaluation through the use of more credible and higher-quality evaluations (independence, credibility and usefulness)	Evaluations are chosen based upon to the strategic and operational needs a
	ILO has the systems and capacity to absorb evaluations
	Resources proportionate to evaluation demands
	Evaluations are based upon a credible knowledge base and monitoring data
Outcome 3. Stronger knowledge base of evaluation findings and recommendations	There is an interest and a demand to use evaluations

3.6 To what extent does the evaluation function align with other accountability and learning processes

Key finding 6: Evaluations broadly meet the accountability and information requirements of key stakeholders, but struggle to tell the story of ILO's specific unique contributions to change and to impact. The evaluation function needs to move beyond the solid foundation it has built on accountability to embrace the unique learning and improvement potential that evaluation can bring for the ILO.
Strength of evidence: Level 2

3.6.1 Alignment with accountability and reporting mechanisms

ILO's oversight and accountability functions comprise: the Office of Internal Audit and Oversight, External audit, the Ethics function, and Evaluation. Of these departments, the evaluation function and the audit and investigation functions receive comparable proportions of the total ILO budget (0.4 per cent). Operationally, the evaluation function does not interact with the audit function and is functionally separate.

Nevertheless, stakeholder interviews suggest that due diligence processes including risk management and audits, operations audits, which include programme reviews that assess project implementation, delivery, and process review, can be perceived to overlap with the evaluation function. There is a tension between the role that evaluations play in both learning and accountability, alongside other reporting demands. While audits and investigations are intended to serve assurance and accountability purposes, evaluation is intended to be both a learning and accountability tool. Stakeholders point to opportunities to rebalance the role of audit for accountability purposes, while evaluations could focus more on insights generated for planning and learning purposes.

---

[62] Views of the IEE team

3.6.2 Funding partners and evaluation

During the review period 2017-21 $144.12 million USD funding was derived from UN partners and $890 million USD from multi-bilateral funding partners (table 19 sets out the top ten partners). Review of funding, monitoring, evaluation, and reporting requirements of ILO's key bilateral and multi-lateral funding partners (see Annex F) as well as joint programmes funded through pooled financing arrangements reveal a mix of requirements related to the size of budget. They are also based upon funding partners' own timeframes for reporting, budgeting, and spending decisions. In some cases, for joint programme that uses pooled funding, there are no requirements for the ILO to conduct its own separate evaluation of its component. . For some cases of bilateral and multilateral funding, such as USDOL (and others, such as SIDA and the EC where they feel it appropriate), they commission and manage their own evaluations of ILO programmes to meet their own accountability requirements but do not expect the ILO to conduct separate evaluations of those programmes for reasons of efficiency. In some cases, where the timeframe for a programme was short (e.g. £1.8m over 21 months for UK funding of Better Work Ethiopia), it was felt there was an unreasonable expectation for both a mid-term and end-term evaluation. It put considerable pressure on programme delivery as well as evaluation staff,

Table 19 ILO's top ten funding partners 2017-21 (data ILO website)

Partner	US $m
European Union	208.47
United States	175.97
Germany	125.10
Multi Donor	106.21
Netherlands	77.07
Core Voluntary funds (RBSA)	70.10
Sweden	64.89
MPTF/UNDP	61.80
Private sector	59.77
Norway	49.24

What is working well

The ILO has an effective system for ensuring that evaluations are delivered in a timely way and in line with the Evaluation Policy and Strategy. The system is designed to ensure independence, professionalism and there are multiple stages of quality assurance stages. ILO constituents and all key stakeholders - including funding partners - are consulted on Terms of Reference and are invited to provide comments on draft findings and recommendations before evaluations are finalized. In developing a clear plan for evaluations that will be conducted, the ILO, through EVAL and Regional Evaluation Officers, seek to ensure that the expectations of funding partners are met.

There is an increasing trend towards clustering decentralized evaluations to achieve greater complementarity, reduce duplication of effort and streamline resources (time, expertise, budgets). This enables opportunities to derive greater learning and value from multiple initiatives and programmes where feasible and relevant. Evaluation of Decent Work Country Programme are a good example for clustering, with centralized evaluations of several DWCPs

carried out to evaluate regional trends, complementarities as well as differences in the work that the ILO does in different contexts and policy areas. Meta-analyses and synthesis reviews are another way that the ILO enables links to be made across multiple sources of evidence and lessons to be drawn from complementary efforts. Joint UN programming using pooled funding and joint evaluations provide opportunities for the ILO to demonstrate its expertise and added value.

Opportunities for strengthening

Duplication and harmonization: Some funding partners are concerned about duplication of evaluations, especially in cases where they conduct their own evaluations of ILO programmes to meet their own accountability requirements. This can generate challenges for the ILO to apply its own criteria for independent evaluations and result in duplication as well as additional costs. The donors that conduct their own evaluations also highlight concerns about the expectation to contribute two percent of a programme budget for ILO's evaluation function, regardless. Some also complained that the two percent budget requirement makes little sense for some large infrastructure programmes where the evaluation investment is no different from a smaller programme. The concern is that these scenarios diminished capacity that could be better used for learning and uptake of recommendations. There is a concern also about the perception of competition between the ILO and funding partners although both are likely to draw on the same pool of evaluation consultants. In all cases, there is a commitment for the ILO and funding partners to contribute inputs into the Terms of Reference and draft reports and each may even consult the other about the selection of consultants. There was also some frustration expressed by some staff about duplication of evaluations where the ILO was one of several UN agencies delivering joint programmes but not leading evaluations. In some cases, EVAL conducted a separate evaluation of its own contribution to meet its own policy requirements, which was felt to be an additional use of time, staff, and financial resources.

Evaluation Content: Concern was raised about a long list of questions used by the ILO for evaluations, regardless of whether they were relevant or useful. This detracted from the strategic value of evaluations, particularly for extracting insights and learning that could inform global public knowledge on the world of work and improve coherence. Where there are too many findings and recommendations on too many issues, there are challenges for prioritizing the most important ones for follow-up and learning. There is a recognition of the diverse (and often divergent) views of different stakeholders and constituents about what matters most. Several funding partners felt that the ILO was sometimes defensive when they challenged recommendations from evaluations. Several said they felt the findings and recommendations of evaluations were not actionable or were too high-level - lacking in practical or technical specificity about what needed to be done differently.

Strategic prioritization: Where funding partners had significant evaluation departments and in-house expertise, they tended to be more critical of the ILO's evaluation approach to determining when an evaluation should be conducted. They were keen to see the ILO adopting an evaluation culture that prized learning and continuous improvement - including being open about what was not working. Most funding partners trusted the ILO's evaluation expertise and expected the ILO to set its own policy - including when evaluations were needed. Most of those interviewed said they did not insist that the ILO conducts separate evaluations of their own funded projects and were open to alternative suggestions for clustered evaluations if there was a clear rationale for this and if it satisfied their own accountability requirements. They wanted to see fewer, more strategic evaluations that had a clear purpose and would add new insights. Some funding partners asked that the ILO proactively shares information about all strategic evaluations conducted by the ILO - not only

those they are involved in funding. They welcomed joint donor meetings to discuss programmes and evaluations, but often lacked the time to discuss issues in depth.

4. Progress made since last IEE 2016

Of the 13 recommendations made in the last IEE 2016, good progress has been made on nine of the recommendations (1,3,4,5,7,9,10,12,13), including strengthening use of evaluability assessments, inclusion of M&E specialists in decent work teams and projects, expansion of the quality assurance system, HLEs and cluster evaluations, introduction of automated recommendations tracker, strengthened knowledge management and dissemination of evaluation results. These achievements have been primarily driven by EVAL.

However, less progress has been achieved on four recommendations (2,6,8,11) including the transitioning of REOs to EVAL, promotion of gender responsive and participatory evaluation, and a more flexible and diverse portfolio of evaluations. Barriers to full implementation mostly arise in the lack of enabling environment concerning evaluation culture (lack of organisational incentives for participation in evaluation, over reliance on voluntary evaluation management), financial barriers (preventing more flexible funding or control of funding for evaluations), but some are linked to evaluation process and procedures (lack of progress on gender responsive evaluation). Table 20 below sets out the overall progress and Annex K sets out the justification behind the assessment.

Table 20: Progress made on IEE 2016 Recommendations

Legend: Little progress achieved | Some progress but not fully achieved | Fully achieved

2016 Recommendations	Progress made
1: Recognizing ILO's comparative leadership in evaluation in the UN system, the necessity for evaluative thinking to meet the 2030	Some progress but not fully achieved
2: Transition REOs into full staff members of EVAL.	Little progress achieved
3: Incentivize and strengthen the evaluation manager and focal person system.	Some progress but not fully achieved
4: Establish an integrated evaluation planning system.	Some progress but not fully achieved
5: Further develop collaboration with other agencies of the UN system to advocate for and support a diverse community of evaluators and national constituents with expertise in evaluating decent work and promoting social dialogue.	Some progress but not fully achieved
6: Enhance evaluation value added and relevance by promoting participatory, gender-responsive and mixed-methods evaluation.	Little progress achieved
7: Expand the quality assurance system to include internal evaluations, and switch to an annual or real-time independent quality assurance system.	Some progress but not fully achieved
8: Diversify and elevate the overall portfolio of evaluations to include more DWCP evaluations and thematic evaluations.	Little progress achieved
9: Strengthen the decentralized evaluation management response mechanism.	Some progress but not fully achieved
10: Prioritize EVAL's communications capacity and coaching function.	Some progress but not fully achieved
11: Strengthen RBM and M&E systems to promote DWCP, programme and project evaluability.	Little progress achieved
12: Update and align the evaluation policy to IEE recommendations and current organizational structure and processes.	Some progress but not fully achieved
13: Develop the new evaluation strategy in a participatory manner to promote ownership and visibility	Some progress but not fully achieved

5. Conclusions

The ILO has a well-developed evaluation function, embedded, and highly regarded with great potential to ensure the organization's decision making, programming design and implementation are underpinned by useful robust evidence. There are some barriers, challenges, and missed opportunities that need addressing in the next strategy period to enhance the value of evaluation, prioritize the evaluation resources, improve the culture of evaluation and its critical role in driving learning and continuous improvement. There are opportunities to strengthen and take the evaluation function to the next level of impact and maturity.

The evidence gathered for this evaluation reveals a system built on a solid strategy clear mandate, and strong operational processes. It has a degree of fragility given the weighty burden on EVAL staff, Regional Evaluation Officers, Evaluation Managers, external evaluation consultants, and Departmental Evaluation Focal Points to deliver its ambitions.

1. **The ILO has a mature, independent evaluation function, with embedded (institutionalized) systems and processes with recognized quality and credibility.** EVAL and the REOs strengthened communications and knowledge management efforts are critical and widely recognized.

2. **Although EVAL has supported extensive learning exercises, the accountability role of evaluations is more visible than learning and improvement.** This has influenced the evaluation culture with many participants of this IEE viewing evaluation as a compliance exercise. There is frustration about the limited time and opportunity to use evaluations and engage in evaluative practice as a vehicle for critical analysis and continuous improvement. Evaluations are not always asking the right questions that could provide valuable insights on what is working well, less well, and why. A strengthened evaluation function focused on strategic lesson-learning will enhance the delivery and effectiveness of ILO's strategy and mandate.

3. **There is an expressed demand from the ILO and amongst constituents for evaluation evidence on impact. This could be better harnessed, particularly as there is an** appetite for evidence on the story of impact, understanding pathways to impact, and seeing what ILO's unique contribution is to attain that impact. Assessment of impact requires robust baseline, integration with intervention design, monitoring, and follow-up and adequate funding (sometimes beyond the life of the intervention).

4. **Despite the perseverance of EVAL and the evaluation network to forge ahead with a transformative ambition (e.g., through the delivery of 30 cluster evaluations during the period of review) this is constrained by barriers in the enabling environment and its own policy and operational guidelines.** The volume of evaluations jeopardizes the usefulness, learning generated, and capacity to engage in evaluation results that can be used to improve design of interventions and decision making. The policy promotes a 'production line' of evaluations which presents risks to the capacity of ILO to ensure high quality useful evaluations. A proportionate approach to evaluation at all levels will ensure optimal use of limited resources. It would ensure that evaluations are able to provide sufficiently robust answers to the questions that are most important to the ILO and can generate the most learning. The evaluation staffing resources are not adequate to meet the demands of the current evaluation strategy and policy or ILO's knowledge management ambition, furthermore the financial system and funding mechanism are too rigid to allow for pooled funding for smooth implementation of portfolio or cluster evaluations, for evaluation funds to be spent once a project has ended (ex-post evaluations, or longitudinal studies), or cost recovery mechanisms to '*payback*' or incentivise home departments of evaluation

managers who give up their time to oversee evaluations as called for in the evaluation policy.

5. **Independence and impartiality are a key strength and a success for the ILO evaluation function, but there are opportunities for a more practical and nuanced approach to upholding independence in the evaluation process.** This could involve greater participation of policy departments and field offices to ultimately enhance the relevance, utility and ownership of evaluation processes and results. Harnessing the evaluation skills and knowledge that exist across the entire evaluation function to shape the conversation on effectiveness and performance at the programme design stage can assist organisational learning. The reliance on EVAL staff for internal real-time quality assurance not only diverts evaluation resource away from delivery of strategic evaluations, but also risks timeliness of evaluation products.

6. **Evaluation is seen as a potentially powerful tool to foster tripartism, social dialogue and build capacity amongst all constituents and intended beneficiaries.** Opportunities for engagement of constituents in the identification of questions and to absorb and act on the evidence generated are not being maximized. The length of reports, number of standardized questions (not tailored to their needs and interests), and a high degree of reliance on eliciting written commentary on documents (Terms of Reference, Draft Reports etc.) does not encourage participation of constituents.

6. Recommendations

The following recommendations arise from the findings and conclusions developed in this IEE and have been informed by consultations with ILO staff through the validation and commentary processes.

Recommendation 1: Maintain EVAL's independence through its organizational location and make more strategic use of the EAC as an influential body

Ensure that EVAL maintains its independent location and access to decision making bodies of the DG, SMT, and GB. Use the time of EAC members more strategically with carefully set agendas, fewer evaluation reports to read and use of appropriate attendees to bring sharper discussion of evaluation insights and implications. Replicate the EAC on a regional basis but with a balanced strategic focus on identifying what evaluation evidence is needed, how evaluation recommendations are being implemented, and lessons learnt discussions informed by evaluative insights on necessary and sufficient factors for achieving impact, implementation blockers and evidence gaps. Review membership of the EAC in light of recent Governing Body discussions (for example, on the inclusion of ACTRAV and ACT/EMP).

Responsible unit	Priority	Time implication	Resource Implication
DG with EVAL support, EAC members	High	Can be achieved within six months	None

Recommendation 2: Use evaluation resources more strategically, stringently prioritize and plan whilst maintaining commitment to improve results-based management system

Reduce the number of evaluations conducted for smaller projects or where evidence is already strong and ensure ability and agility to direct resources to areas of least evidence,

risk, or innovation. Investigate alternative ways to set the criteria of whether to evaluate or not, guided by the requirement to generate evaluation evidence and learning for the ILO's strategic objectives and planned outcomes. Ensure ability and agility to direct resources to areas of least evidence, risk, or innovation. Consider organizational priorities and capacity to deliver and use evaluations. Continue to facilitate cluster evaluations where more appropriate, efficient or could generate better learning opportunities. Continue the HLE planning mechanism.

PROGRAM and PARDEV to collaborate to strengthen monitoring as key accountability and course correction mechanism: This will require less formal evaluation and ability to move threshold for internal and independent evaluation (or introduce more flexibility and demand-driven evaluation). Shift the internal self-evaluation to project monitoring and formal annual and project completion reviews.

Actions that can be taken:

i) Move to a model that incorporates mandatory evaluation and discretionary evaluation; the use of thresholds, targets and incentives can help drive demand, reduce work burdens and unnecessary formal independent evaluation exercises, whilst maintaining accountability and oversight.

ii) Investigate alternative methods for ensuring evaluation coverage (reducing volume but ensuring accountability and lesson learning). Options include:

a) Increase the threshold of project spend as the trigger to evaluate independently (for example moving to $3 million trigger would provide approximate cover of 10% of DC portfolio). Remove requirement for mid-term evaluations for interventions with budget of <$5m) but replace these with improved monitoring and review systems. Introduce a new criteria-based mechanism to decide on whether to evaluate (alongside the project spend trigger until such time that sufficient monitoring mechanisms are in place to provide accountability and performance review function). Criteria[63] could include strategic relevance to the ILO, evidence gap, innovative design or untested theory, formal commitment to funding partner or key stakeholder, likelihood of influencing policy, potential to leverage partnerships, evaluability / feasibility of undertaking evaluation, capacity to absorb findings and generate useful learning.

b) Devolve decision making on decentralized evaluations to regions and technical departments to a greater extent. Encourage use of criteria-based approach (as set out in ii (a) above) coverage targets determined by country/region/department for percentage of their portfolios and plans. These could be endorsed by EVAL and the EAC. Additional coverage and accountability can be supplemented by EVAL-mandated evaluations - especially for higher level strategic and thematic centralized evaluations. This would bring the benefit of ensuring coverage for regions or departments with lower project budgets or fewer evaluations.

iii) Allow space for demand driven evaluation planning to complement required project evaluations. EVAL should work with departments to identify their evaluation needs and evidence gaps– to allow for demand driven evaluations at the portfolio/thematic level, while maintaining its independence on final decision to evaluate. Incorporate participatory discussion with funding partners (facilitated by PARDEV). Use the EAC (and regional EACs once established) to identify evidence needs to inform the Evaluation Plan (see option b above).

---

[63] Criteria suggested in Annex H are derived from IEE team experience and a review of criteria used by other selected UN agencies.

Responsible unit	Priority	Time	Resource Implication
EVAL and EAC members	High	For next evaluation strategy and revision to policy/guidelines	None. No increase in funding required for evaluation
PARDEV and PROGRAM (in relation to strengthening monitoring)	High	Over the course of 2 years	Significant investment in monitoring required

Recommendation 3: Strengthen the enabling environment: Ensure adequate financial and human resourcing for evaluation to deliver a transformative evaluation ambition

The ring-fencing mechanism within the resource allocation system needs to be sustained to ensure decentralized evaluations are fully funded. In addition, the senior management team need to ensure that the financial process and mechanisms allow for ease of creation of pooled funds, ex-post evaluation funding, and cluster evaluations.

Actions that can be taken:

i) Reduce workload of REOs via change in policy guidelines on number of evaluations required. Provide call down technical support and on-demand quality assurance to relieve their workload.

ii) Enhance formal coordination between M&E officers and the evaluation function

iii) Ensure sustainable and high-quality consultant pool – ensure that the ToR are realistic and relevant to what is being evaluated (reduce number of evaluation questions), benchmark day rates with other UN entities, provide online system for approvals/processes. Consider long term agreement (on demand service) or procurement framework system with pre-qualified suppliers, balance evaluation experience with knowledge of normative mandate.

Responsible unit	Priority	Time	Resource Implication
FINANCE and SMT (for recommendation relating to resource allocation)	High	Over the course of one year	None
EVAL	Medium	Over the course of one year	Funding for a call down demand led service. Fewer evaluations but average cost per evaluation would increase.

Recommendation 4: Strengthen the enabling environment: Invest in building the capabilities and skills to undertake, manage, and use evaluations

Professionalize, incentivize, and recognize evaluation skills. Actions that can be taken:

i) Shift to more dynamic ongoing capacity building rather than static provision of training. Formally assess, train, accredit EMs, DEFPs, and M&E officers. Staff should be required to refresh their skills in order to maintain 'accreditation'. Capabilities and skills should include investigation, credible evidence, critical thinking, communication, and ways of disseminating / enabling learning – not just systems management. Introduce a programme of internal and external speakers and training opportunities to keep REO, DEFP, EVAL skills up to date.

ii) Formal recognition of evaluation in performance appraisals (for example five per cent of time /objectives formally dedicated to evaluation).

iii) Formal requirement of senior managers (in objectives) to adopt evidence-based decision making and maintain skills in evaluation and evidence.

iv) Invest time and effort in creating a well-supported, invigorated, active evaluation community of practice (EMs, DEFPs, REOs, EVAL) with appropriate opportunities for training and exchange (see above comment on what training should include).

Responsible unit	Priority	Time	Resource Implication
EVAL and HRD	Medium	Over the course of strategy period	Staff time for implementing and more importantly maintaining this recommendation, especially in enhancing training of REOs, DEFPs, and EVAL

Recommendation 5: Strengthen the enabling environment: Foster a stronger culture of learning

Promote a culture of collaboration, continuous learning, professional development through evaluation design, delivery and follow up. Enable discussions of failure – what is going wrong and help to understand why. Invest in forging stronger links between monitoring, learning and evaluation that enable evidence-based decision-making, incentivize honest and open discussion on what is working less well, allow for adaptation and continuous learning.

Rebalance from accountability and towards learning by reducing the number of evaluations conducted, ensuring that all evaluations are needed and have a clear purpose will give staff the space for more engagement in evaluations and evidence generated by evaluations.

i) Align evaluation plans and strategies with the planning and budgeting process: seek where possible to use evaluations to assess how ILO is meeting its strategy objectives– especially in areas where the ILO has a unique mandate (improving standards and opportunities for decent work, fostering social dialogue and collaboration through tripartism).

ii) EVAL,PROGRAM, Field Directors and Directors of Policy Departments can enhance coordination during design and quality control of DWCPs. PARDEV and EVAL can continue to join forces to encourage the use of evaluation findings and lesson into the appraisal stage of programming.

iii) Continue the excellent work on knowledge management and communications of evaluation evidence through webinars, podcasts, twitter, email, seminars.

Responsible unit	Priority	Time	Resource Implication
SMT. EVAL, PARDEV, Field and Policy Directors Knowledge Management Coordination Team / senior responsible for Knowledge Management Gateway	High	Over the course of new strategy period	None

Recommendation 6: Ensure quality, credibility, and utility of evaluations

Improve the design of evaluations to focus only on the most relevant issues. Improve the quality of evaluation recommendations with a greater utility focus. Reduce the potential risk

of highly visible HLE being of inadequate quality by ensuring the real-time quality assurance of HLEs, conducted independently from EVAL.

i) If resources permit, use an external agency to undertake real time quality assurance, this will ensure timeliness of evaluation products, reduce workload of REOs and EVAL staff and ensure a live gatekeeping mechanism for poor evaluation. Introduce formal quality assurance of HLEs outside of EVAL or introduce reference /governance group mechanism.

ii) A participatory evaluation approach should be embedded throughout evaluations' life cycle. Actively use governance panels (reference groups and technical advisory groups) to ensure independence. This can allow for a more participatory evaluation approach whilst not impacting on impartiality/independence. This also follows UNEG guidance.

iii) Recognizing that the world of work is complex, an adaptive management approach could be used that will enable course correction along the way – based on real-time factors and taking account of unexpected changes or influences.

iv) Enable co-creation of recommendations without jeopardizing independence. More participatory approaches to generation of recommendations to enhance likelihood of implementation and learning. Ensure adequate time for consultations with key stakeholders that are critical to the change.

v) Improve recommendations of evaluations: revise recommendations guidelines and ensure quality assurance at the reporting phase to check the recommendations are 'doable'.

vi) Revise ToR guidelines to stress the need for fewer evaluation questions – only including what is directly relevant and necessary. Revise Evaluation Policy Guidelines (incorporate useful advice from the OECD DAC publication on Applying Evaluation Criteria Thoughtfully) [64] to stress that not all DAC Criteria need to be covered on every evaluation. This will help to scale down the ambition and sharpen the focus of the evaluations.

Responsible unit	Priority	Time	Resource Implication
EVAL, PROGRAM, Field Directors, Policy Departments and PARDEV	Low	Over the course of new strategy period	Cost of external quality assurance – real time provision

Recommendation 7: Invest in knowledge capture and information management systems

Enable more synergy between evaluations and other learning products such as project reporting, research, and other knowledge products. Introduce an accounting code for evaluation to enable EVAL to track and analyze spend on evaluations, including of XBTC and RBSA funded interventions. Ensure all project evaluations are accessible on i-EVAL Discovery, allow for better tagging of evaluations to allow search by ILO strategic objectives and P&B outcomes. Building a culture of evaluation relies upon staff and stakeholders seeing evaluation as a useful, credible source of knowledge, to support this there must be more synergy between evaluations and other learning products such as project reporting, research, and other knowledge products.

---

[64] See for example OECD DAC Applying Evaluation Criteria Thoughtfully 2021

Responsible unit	Priority	Time	Resource Implication
Finance	Low	Over the course of strategy period	None
EVAL, Knowledge Management and Research, PARDEV, Regions	Low	Over the course of strategy period	Cost of integrating systems, making improvements to i-track

Recommendation 8: Continue and expand ILO's commitment to the UN evaluation system

Continue to share ILO's evaluation expertise and knowledge with the UN evaluation network and other UN organizations. To enhance coherence, minimize overlaps and avoid overburdening stakeholders, the ILO should seek opportunities for collaboration and coordination among UNEG members (through joint and system wide evaluations), when appropriate and feasible.

Responsible unit	Priority	Time	Resource Implication
EVAL	High	Over the course of the strategy period	None

Recommendation 9: Build on strong foundation to develop a future-fit evaluation strategy and policy

Update policy guidelines (to incorporate prioritization recommendations and revised quality assurance, and independence mechanisms). Develop the new Evaluation strategy in participatory manner to increase buy-in from key stakeholders. Develop a new refined Evaluation ToC as a participatory exercise to underpin the new Evaluation Strategy. This will be a vital tool for planning, communications, and management, and can be used as the basis for identifying the milestones, targets and indicators of success in the Evaluation Strategy results framework. The two products together will be key in articulating and communicating the purpose of the evaluation function - providing clarity, a sense of ownership, strategic direction, measuring and managing performance. As part of this, EVAL should develop its strategic approach to Impact Evaluation – EVAL & EAC should have a stronger voice in determining IEs undertaken as well as ILO responsible officials and funding partners.

Responsible unit	Priority	Time	Resource Implication
EVAL	High	By October 2022	None

7. Lessons learned

1. **Internal and external coherence:** The evaluation highlighted that the Evaluation Function is independent but is intrinsically part of a system. The enabling environment, culture of learning and of evaluation, and the quality of programming relies on strong coherence, dialogue, and collaboration at all levels. Coherence across the evaluation function and beyond can help to strengthen ownership and participation of stakeholders, evaluability of programmes, and the opportunity to ensure that evaluations provide useful findings, recommendations, and lessons. Ensuring opportunities to foster dialogue is key to delivering policy goals and to effective and efficient systems.
2. **Integration of gender and cross-cutting issues into all stages of programming and evaluation:** The evaluation identified opportunities to more systematically mainstream

gender and cross-cutting issues in all evaluation stages and deliverables. Greater attention, and shared understanding of how these issues can be embedded in programming at all stages can strengthen the integration of ILO policies on these issues.

3. Embedding evaluative thinking across the organization: Fostering a culture that rewards transparency about what is working and not to generate insights on what could be done differently. Integrating evaluative thinking at all levels provides an opportunity to strengthen RBM and thinking about how activities are linked to broader mid-term and long-term goals.

4. Financial and human resource requirements: Effective implementation of policies and programmes requires adequate human and financial resources. Gaps in resourcing, or overloaded staff members can compromise sustainability. It is critical that policy decisions be matched with adequate human and financial resources

5. Incentives and accountability or balancing "the carrot" and "the stick": The evaluation pointed to the challenge of achieving a balance between incentive structures and accountability mechanisms. Institutional systems and accountability frameworks need to continually strive to incorporate both rewards and clear accountability requirements

References cited

Boris Groysberg, Jeremiah Lee, Jesse Price, and J. Yo-Jud Cheng. Jan/Feb 2018. "The Leader's Guide to Corporate Culture How to manage the eight critical elements of organizational life." *Harvard Business Review.*

EVAL. June 2020. "Guidance Note 3.2:Adapting evaluation methods to the ILO's normative and tripartite mandate v1."

EVAL. 2020. "ILO policy guidelines for results-based evaluation Principles, rationale, planning and managing for evaluations, 4th edition."

EVAL. 2017. "ILO result-based Evaluation Strategy 2018-21."

EVAL. 2021. *Results and reflections from a quality appraisal of ILO evaluations, 2020 i-eval Think Piece No. 21 .* https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_830075.pdf.

EVAL. 2021. "Self assessment of the Evaluation Function of the ILO 2018-2021."

EVAL/Luc Franche, Juan-David Gonzales and Sophie. 2020. "Quality assessments of ILO decentralized evaluations: Key results of the quality appraisal 2019 and way forward for the integration of gender equality and empowerment of women considerations into evaluation." i-eval THINK Piece, No. 19.

Guijt, Irene. 2014. *Methodological Briefs Impact Evaluation No. 5 Participatory Approaches .* UNICEF.

ILO . 2021. "Development Cooperation Internal Governance Manual."

ILO. 2006. "Circular 245 Evaluation Advisory Committee 1.09.2006."

ILO. 2017. "ILO Evaluation Policy GB.331/PFA/8."

ILO. 2020. "ILO policy framework for tackling the economic and social impact of the COVID-19 crisis."

ILO. 2016. "Independent Evaluation of the ILO's Evaluation Function for the period 2011-2016."

ILO. 2009. "Office Directive. IGDS Number 112."

ILO. 2020. "Operating Procedures No.1. Implications of COVID-19 on evaluations in the ILO. Practical tips on adapting to the situation."

—. 2009. "Strategic Policy Framework 2010–15: Making Decent Work Happen." *Governing Body, 304th Session, GB/304/ PFA/2.* Geneva.

—. 2016. "The ILO's Strategic Plan for 2018–21 GB.328/PFA/1." *328th Session, Geneva, 27 October–10 November 2016.* Geneva.

ILO. 2020. "The ILO's Strategic Plan for 2022–25 GB.340/PFA/1(Rev.1)."

Jane Buckley, Thomas Archibald, Monica Hargraves,William Michael Trochim. 2015. "Defining and Teaching Evaluative Thinking: Insights From Research on Critical Thinking." *American Journal of Evaluation 36(3).*

JIU, Sukai Prom-Jackson and George A. Bartsiotas. 2014. *Analysis of the evaluation function. JIU/REP/2014/6.* Geneva: United Nations.

Mayne, John. 2010. "Building and evaluative culture." *The Canadian Journal of Program Evaluation* Vol. 24, No. 2, Pages   1–30.

MOPAN. 2021. "MOPAN Assessment Report: ILO 2020 Assessment Cycle, Paris."

OECD/DAC . 2019. *Better Criteria for Better Evaluation Revised Evaluation Criteria Definitions and Principles for Use.* OECD.

OECD/DAC. 2010. "Evaluating Development Co-operation –Summary of Key Norms and Standards."

OECD/DAC. 2002. "Glossary of KeyTerms in Evaluation and Results Based Management."

OECD/DAC. 2010. "Quality Standards for Development Evaluation."

Schwandt, Thomas A. July 2019, 25, no. 3. " Post-Normal Evaluation? ." *Evaluation* 317–29.

UNEG. 2020. "Ethical guidelines for evaluation."

UNEG. 2016. "Norms and Standards for Evaluations." New York.

UNEG. 2020. "Self-Assessment Maturity Matrix For UN Evaluation Functions Draft."

UNFPA. 2020. *Annual report on the evaluation function, .* (DP/FPA/2021/5).

UNGA\OIOS. 2013. "Strengthening the role of evaluation and the application of evaluation findings on programme design, delivery and policy directive. Report of the Office of Internal Oversight Services ."

World Bank Group/IEG. 2019. "Evaluation Principals."