INTERNATIONAL LABOUR OFFICE

Governing Body

334th Session, Geneva, 25 October–8 November 2018

GB.334/PFA/6

Programme, Financial and Administrative Section
Audit and Oversight Segment

PFA

Date: 2 October 2019
Original: English

SIXTH ITEM ON THE AGENDA

Annual evaluation report 2017–18

Purpose of the document

In Part I of the present document, the Governing Body is provided with a progress report on the ILO's evaluation work during 2017–18, as measured against its results-based evaluation strategy 2018–21. Part II touches upon ongoing internal efforts in the ILO aimed at strengthening the measurement of overall effectiveness of the Office. The Governing Body may wish to take note of the present report, endorse recommendations contained therein and confirm priorities for the 2019–21 programme of work for evaluations (see draft decision in paragraph 89).

Relevant strategic objective: Relevant to all strategic objectives.

Main relevant outcome/cross-cutting policy driver: Enabling Outcome B: Effective and efficient governance of the Organization.

Policy implications: The recommendations of this report will have policy implications.

Legal implications: None.

Financial implications: None.

Follow-up action required: Approved recommendations will be inserted in a rolling action plan for tracking follow-up.

Author unit: Evaluation Office (EVAL).

Related documents: GB.332/PFA/8; GB.331/PFA/8.

This GB document is printed in limited numbers to minimize the environmental impact of the ILO's activities and processes, contribute to climate neutrality and improve efficiency. GB members and observers are kindly requested to bring their copies to meetings and to avoid asking for additional ones. All GB documents are available on the Internet at www.ilo.org.

Contents

Page

Introduction 1

Part I. Implementation of the ILO's evaluation strategy 2

Progress made towards achieving key milestones 2

Outcome 1: Enhanced capacities and systems of evaluation for better practice and use 3

Sub-outcome 1.1 – Evaluation activities conducted in a timely fashion and in accordance with Evaluation Policy requirements 3

Sub-outcome 1.2 – Strengthened evaluation capacity of staff in regions and departments 6

Sub-outcome 1.3 – Constituents engaged in monitoring and evaluation of DWCPS and development cooperation activities in an SDG-responsive manner 8

Sub-outcome 1.4 – Evaluation integrated in DWCPs and development cooperation activities, including a focus on SDGs 9

Sub-outcome 1.5 – Established capacity of regions and departments to mainstream and use evaluation 10

Outcome 2: Enhanced value of evaluation through the use of more credible and higher quality evaluations (independence, credibility, usefulness) 11

Sub-outcome 2.1 – Use of strategic cluster evaluations to gather evaluative information more effectively 11

Sub-outcome 2.2 – Improved quality of internal, decentralized and centralized evaluations 12

Sub-outcome 2.3 – Credible impact evaluations conducted to build knowledge for effective policy interventions 14

Sub-outcome 2.4 – Evaluation framework further aligned with ILO mandate and context, including SDGs 14

Outcome 3: Stronger knowledge base of evaluation findings and recommendations 15

Sub-outcome 3.1 – Strengthened accessibility and visibility of evaluation information through i-eval Discovery 15

Sub-outcome 3.2 – More targeted communication of evaluation findings 16

Sub-outcome 3.3 – Improved use of evaluation findings and recommendations by constituents and management for governance and decision-making 17

Sub-outcome 3.4 – Evaluations used to meet strategic knowledge requirements through further analysis of findings and results of evaluations 21

Part II. Assessing the ILO's effectiveness and results 22

A challenging but opportune time to innovate the way we assess overall effectiveness and results 22

Support proper design, monitoring and reporting on programme and budget policy outcomes and DWCPs as the programmatic backbones of the ILO's Decent Work Agenda 24

Enhance the sufficiency of evaluations to inform a higher order of effectiveness 25

Page

Refining the meta-analysis approach and methodology to report on effectiveness 26

Taking a more transformative approach to developing a more comprehensive validation process of ILO performance at the country and global levels 27

Draft decision 27

Introduction

1. This annual evaluation report (AER) reflects the work undertaken by the Evaluation Office (EVAL) from August 2017 to August 2018. The reporting cycle is linked to the Governing Body's last session of the calendar year. During this period, the ILO's evaluation function underwent significant changes that were made in response to: the 2016 independent evaluation of the ILO's evaluation function (2016 IEE); implications of the Sustainable Development Goals (SDGs) for the ILO's results framework from an evaluation perspective; and ongoing reflections within the ILO to improve results-based management (RBM).

2. The 2016 IEE concluded that the ILO's evaluation function had made good progress since 2011 in delivering better evaluation products and reinforcing its structural independence, which had led to transformational changes in terms of evaluation practices and the uptake of findings. It also recommended revisiting the evaluation policy 2005 and the evaluation strategy's results framework (2011–17) in light of changing internal and external contexts. A new evaluation policy was subsequently formulated, following extensive consultations that involved over 250 staff and constituents. Following the adoption of the revised evaluation policy (November 2017), the Governing Body adopted the new results-based evaluation strategy 2018–21 (March 2018).

3. The results framework of the new evaluation strategy marks a departure from the previous strategy. It presents a clear theory of change based upon three new outcomes, 13 sub-outcomes and 19 indicators. It complements the evaluation policy 2017 and identifies implications and key areas of action that will result from it. It is also aligned with corresponding key ILO policy and programme documents, particularly the Strategic Plan for 2018–21. Reporting in the AER will be synchronized with information provided for the report on ILO programme implementation on outcome B: Effective and efficient governance of the Organization, indicator B.5: Adequacy of use of findings and recommendations from independent evaluations in decision-making by ILO management and the Governing Body.

4. Part I of this report describes the progress that the evaluation function made in rolling out the new evaluation strategy. It reports against each outcome, indicator and biennial milestone identified in the Strategy's results framework. While there is continuity in their goals and aims, the new evaluation policy and strategy also introduce bold principles that will guide the ILO's evaluation function to reach the highest level of maturity. [1] To address some of the "evaluation fatigue" that has set in, more strategic and clustered evaluations of project and programme activities are being considered. Moreover, evaluation approaches, methods and frameworks are being reviewed to ensure that they are more attuned to the ILO's specific mandate (for example, tripartism, social dialogue and normative work) without compromising the internationally recognized Development Assistance Committee of the Organisation for Economic Co-operation and Development (OECD/DAC) evaluation criteria.

5. Finally, evaluation-related capacity-development training programmes for ILO staff, tripartite constituents and other stakeholders are being developed and rolled out on demand. They focus on key evaluation competencies and linkages between the ILO's Decent Work Agenda and the SDGs from an evaluation perspective. Existing key challenges remain on EVAL's radar; for example, how gender issues could be better covered in evaluations, and how evaluations could generate more information on the impact and sustainability of the

[1] As set out by the Joint Inspection Unit (JIU) maturity matrix. JIU: *Analysis of the evaluation function in the United Nations system*, (United Nations, Geneva, 2014).

ILO's work on the ultimate beneficiaries. Similarly, evaluation-related knowledge products and services are of equally high priority as cornerstones in supporting an evaluation function that is credible, accountable and able to inform decision-making.

6. Part II of the report provides an overview of EVAL's perspective on ongoing internal efforts in the ILO to strengthen RBM in evolving internal and external contexts. This has increasingly been made clear by various external evaluations, ILO independent reviews and evaluations, [2] and recent International Labour Conference discussions on development cooperation. The latter specifically called for better use of RBM tools and evaluations, including impact evaluations, to demonstrate what works and to enhance the visibility of the Decent Work Agenda and its contribution to the SDGs. [3]

7. Building on earlier experience in measuring decent work results, operational opportunities provided in the new evaluation strategy and the ongoing work of the ILO's internal RBM task force, this part of the report also outlines ideas on a comprehensive validation process of ILO performance. The optimal scenario involves a transition from the "effectiveness assessment", based on meta-studies of multiple project evaluations, to a system that involves full coverage of all Decent Work Country Programmes (DWCPs) and programme and budget policy outcomes, using a four-year planning cycle. This would result in fewer dedicated project evaluations, with smaller ILO activities being evaluated under DWCP-specific or programme and budget policy outcome evaluations. The adoption of this ambitious scenario would require the consent of donors and key stakeholders, and updated protocols. It is expected that this would strengthen results-based monitoring and reporting, and enable better assessment of organization-wide effectiveness. This would be in line with the Programme and Budget for 2018–19 expectations that "flexibility in choosing evaluation approaches will be ensured to identify the 'best fit' methods and ensure that credible evidence is produced and critical evaluations questions and criteria covered".

Part I. Implementation of the ILO's evaluation strategy

Progress made towards achieving key milestones

8. Part I of the AER describes the performance of the ILO's evaluation function, as measured against the new results-based evaluation strategy 2018–21. The theory of change for the evaluation function has three dimensions: (i) the contribution of evaluation to an effective and efficient ILO delivering decent work policies and programmes; (ii) the enhancement of the credibility and leadership of the ILO through an embedded accountability, transparency and evaluation culture; and (iii) the advancement of decent work goals by leveraging national and international partnerships to measure the contribution to the SDGs.

[2] Over the years, external and independent reviews – namely the 2017 Multilateral Organisation Performance Assessment Network (MOPAN) exercise, the 2016 independent evaluation of the ILO's evaluation function 2011–16 (2016 IEE), the 2016 Report of the External Auditor on RBM, and various ILO AERs since 2011 – have identified the need to reinforce internal monitoring, reporting and evaluability to improve the link between performance and strategic decision-making.

[3] ILO: *Provisional Record* No. 7A, International Labour Conference, 107th Session, Geneva, 2018, and https://www.ilo.org/ilc/ILCSessions/107/reports/texts-adopted/WCMS_633138/lang--en/index.htm.

9. In brief, the evaluation function advances *effectiveness*, *credibility* and *partnerships* by realizing the following three outcomes: [4]

- outcome 1 – enhanced capacities and systems of evaluation for better practice and use;
- outcome 2 – enhanced value of evaluation through the use of more credible and higher quality evaluations (independence, credibility, usefulness); and
- outcome 3 – stronger knowledge base of evaluation findings and recommendations.

10. The strength of the evaluation function and the extent to which it can respond to all indicators outlined in the evaluation strategy depend on the ILO's enabling environment, particularly in terms of its evaluation culture, organizational learning and RBM.

11. While the new evaluation strategy was approved by the Governing Body in March 2018, reporting on the status of each indicator in 2018 takes into account selective data from 2017 as per the usual reporting cycle, which is linked to the timing of Governing Body sessions. The following section is organized by strategic outcome. For ease of reference, each sub-outcome indicates the status in respect of meeting the biennial milestone of the relevant indicator: achieved, on track, delayed or not yet started.

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

12. With respect to outcome 1, targets can be realized by: providing ILO staff and constituents with training, support and encouragement in building their skills; using the ILO quality standards for evaluation management and use; and developing and maintaining accurate evaluation systems that encourage integrated evaluation planning.

Sub-outcome 1.1 – Evaluation activities conducted in a timely fashion and in accordance with Evaluation Policy requirements

<table>
<tr><td>Indicator 1.1: All mandatory evaluations are completed in a timely manner for use by management, constituents and donors.</td><td>Baseline: 90% coverage for independent evaluations and 33% coverage for internal evaluations.</td><td>Status
On track.</td></tr>
<tr><td colspan="2">Biennial milestone (2018–19): 95% of independent evaluations and 50% of internal evaluations completed in a timely manner by the end of the biennium (target).</td><td></td></tr>
</table>

13. Independent and internal project evaluations are cornerstones for RBM and organizational learning. They complement the three high-level evaluations conducted annually by EVAL as mandated by the Governing Body. The results-based evaluation strategy 2018–21 includes an indicator on the completion rate of mandatory project evaluations.

14. The *i*-eval Discovery application illustrates all projects requiring evaluation, thus informing constituents, donors and ILO officials on the number of planned and completed evaluations (see biennial milestone 3.1). For the period under review, 40 ILO-managed independent evaluations and one joint independent evaluation were completed (representing a 98 per cent

[4] As described in the results-based evaluation strategy 2018–21.

completion rate), [5] in addition to 25 internal evaluations and one externally managed evaluation (figure 1). [6]

15. The 41 completed ILO-managed independent evaluations figure is slightly higher than during the previous reporting year (33) but is mostly on a par with the past eight-year average of 43 (2010–17). As of July 2018, 134 independent evaluations are required during the 2018–19 biennium. This figure is expected to fluctuate as new projects are approved and extended over the next 12 months.

16. Submission rates for internal and self-evaluations [7] have been more challenging. Self-evaluations (undertaken by project managers) are not overseen by EVAL, while internal evaluations (undertaken by consultants and managed by project managers) do not receive the same level of oversight as independent evaluations (undertaken by external evaluation consultants and managed by impartial evaluation managers overseen by EVAL). Although EVAL consistently requests compliance with evaluation requirements, the submission rates for self-evaluations and internal evaluations have been consistently low. During the period under review, 25 internal evaluations were completed, far from the 69 reports that were supposed to have been submitted. This represents a completion rate of 36 per cent (approaching the target of 50 per cent by the end of 2019).

17. In an effort to improve completion rates for internal and self-evaluations, EVAL launched the Internal Evaluation Certification Programme (IECP) in April 2017 to build organizational capacity, and has been encouraging project managers to incorporate self-evaluation components into final progress reports. The added benefit of the latter is that it reduces the reporting burden on project managers, as they then only have to prepare one report. This AER, therefore, includes a recommendation to formalize this practice.

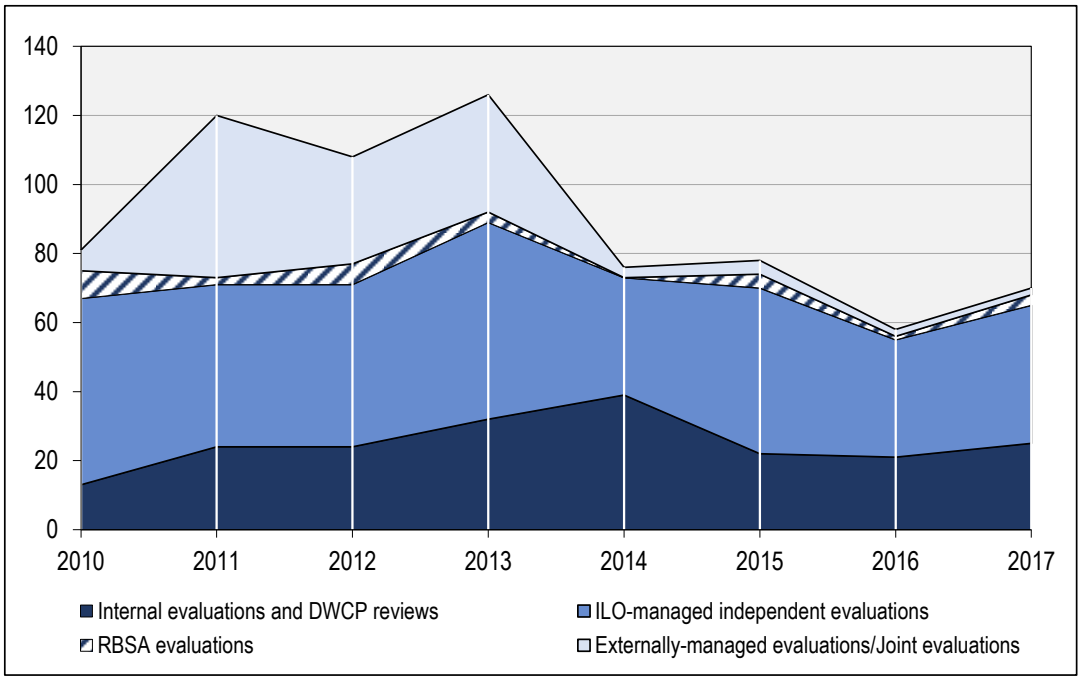
Recommendation 1: Formalize the good practice that final progress reports incorporate self-evaluation components in lieu of a separate formal self-evaluation report.

[5] EVAL is rigorous in ensuring that independent evaluations are completed as scheduled. However, in 2017 an independent evaluation was required to be conducted for a project that was administratively backstopped by the Social Finance Unit. The project was discontinued and the donor requested the return of all unused resources. The evaluation was not completed as required.

[6] See biennial milestone 3.3.2 on the status of management follow-up to recommendations from ILO-managed independent evaluations. The number of externally managed and joint evaluations is unpredictable. They are dependent on donor policies and joint funding trends.

[7] Self-evaluations are required for projects with budgets below US$500,000, internal evaluations for projects between $500,000 and $1 million, and independent evaluations for projects above $1 million.

Figure 1. Number of evaluations by type, 2010–17 *



* One joint evaluation and one external evaluation were completed in 2017.

Selecting high-level evaluation topics for strategic use

18. Annually, the Governing Body is presented with a draft rolling workplan for proposed high-level evaluations, which is based on inputs received from constituents. Table 1 indicates high-level evaluations to be reconfirmed (2019–20) and confirmed (2021–22) by the Governing Body. The rolling workplan is only indicative for planned evaluations beyond the ongoing biennium and can be adjusted according to changing requirements.

Table 1. Summary of high-level evaluation topics, 2018–22

		Institutional or outcome level	Outcome level	DWCP
		Topic	*Topic*	*Region/subregion*
New	**2022**	Development and use of labour statistics [1]	Strategies to promote the application of international labour standards [2]	Europe [3]
	2021	Action Plan for Gender Equality [4]	Promoting fair and effective labour migration policies [5]	Asia [6]
To be reconfirmed	**2020**	Research and knowledge management [7]	Promoting sustainable enterprises [8]	Americas [9]
	2019	ILO public–private partnerships [10]	Formalization of the informal economy [11]	Africa [12]
Completed	**2018**	Independent evaluation of the ILO's capacity development efforts – 2010–17 (all constituents)	Independent high-level evaluation of the ILO's strategy and actions for improved youth employment prospects, 2012–17	Arab States – High-level evaluation of the ILO's programme of work in Lebanon and Jordan in terms of decent work and the response to the Syrian refugee crisis

[1] Institutional: Not evaluated before; selected based on inputs received from consultations in 2018.

[2] Outcome: Not evaluated as an outcome in at least two biennia; selected based on inputs received from consultations in 2018.

[3] DWCPs: Rotationally due in 2022.

[4] Institutional: Due as part of approved Action Plan.

[5] Outcome: The International Labour Conference resolution concerning fair and effective labour migration governance adopted in June 2017 called for a high-level evaluation of work to promote fair recruitment. This evaluation was initially planned for 2019, but as this was considered too early by the concerned department it is now proposed for 2021.

[6] DWCPs: Rotationally due in 2021.

[7] Institutional: Not evaluated before; selection based on inputs received from consultations in 2017.

[8] Pre-approved in AER 2016–17.

[9] DWCPs: Rotationally due in 2020.

[10] Institutional: GB decision – GB.326/POL/7.

[11] Pre-approved in AER 2016–17.

[12] DWCPs: Rotationally due in 2019.

Sub-outcome 1.2 – Strengthened evaluation capacity of staff in regions and departments

Indicator 1.2.1: ILO staff evaluation capacities are upgraded.	**Baseline:** By end-2017, 77 staff members had been certified as evaluation managers and two were certified as part of IECP. Eight of the EMCP certifications were awarded in 2016–17.	**Status**
		On track.
Biennial milestone (2018–19): At least 30 additional ILO staff members are certified as evaluation managers and internal evaluators.		

19. The ILO's previous results-based strategy (2011–17) contained a combined indicator calling for constituents and ILO staff to be trained in evaluation. The new results-based evaluation strategy 2018–21 separates the two, and includes specific sub-outcomes dedicated to upgrading ILO staff (1.2) and constituents (1.3) in respect of evaluation-related capacities.

20. To meet the demands associated with undertaking the large number of internal and independent project evaluations every year, EVAL developed training programmes to improve organizational capacity. For independent evaluations, it developed the Evaluation

Manager Certification Programme (EMCP) in 2012 to upgrade the quality of evaluation management and to expand the pool of qualified candidates. To date, eight rounds of the EMCP have been held in partnership with the Human Resources Department and the International Training Centre of the ILO (ITC-ILO), [8] resulting in 81 certified ILO staff.

21. Furthermore, in 2017 EVAL launched the IECP to improve the quality, frequency and utility of internal evaluations. As with the EMCP, the IECP is delivered through a blended format that includes a distance-learning module, a workshop and a coaching phase during which trainees conduct an internal evaluation under EVAL's supervision.

22. In 2018, seven EMCP trainees and 11 IECP trainees completed all of the requirements for certification. Therefore, EVAL is nearly two-thirds of the way to achieving the biennial milestone of having 30 additional certified evaluation staff. Table 2 shows the number of staff members who received training (certification also requires trainees to pass a practicum). As the data shows, there are regional variations in participation that need to be addressed to ensure a homogenous evaluation culture throughout the Office. The approach taken towards professionalizing staff to participate in and contribute to evaluation in the ILO is considered a best practice in the United Nations evaluation community: various members of the United Nations Evaluation Group (UNEG) have plans or are planning to emulate it.

Table 2. ILO officials trained in evaluation in 2018

Type of training offered	Africa	Americas	Arab States	Asia and the Pacific	Europe	HQ	Total
General monitoring and evaluation (M&E) training	35	5	3	7	99 *	0	149
Evaluation management skills	5	0	0	13	0	25 *	43
Internal evaluation skills	2	5	0	5	1	4	17
Total	**42**	**10**	**3**	**25**	**100**	**29**	**209**

* Projected figures because additional training will take place in the fourth quarter of 2018.

Note: Any training that is less than one day is counted as sensitization and not included in this table.

Indicator 1.2.2: The ILO evaluation network is functioning based on clearly established roles and job descriptions.	**Baseline:** Currently, evaluation network functions (departmental level and evaluation managers) are performed on a voluntary basis.	**Status**
		On track.
Biennial milestone (2018–19): Evaluation responsibilities are included in job descriptions of departmental focal points for evaluation and certified evaluation managers receive standardized assessments in their performance appraisals.		

23. As identified in the 2016 IEE, there is a need to strengthen incentive structures for departmental evaluation focal points (DEFPs), evaluation managers and internal evaluators to manage and conduct evaluations. This was recognized by the new evaluation strategy, which, for the first time, includes an indicator on the extent to which the ILO's evaluation network is functioning, based on clearly established roles and job descriptions. Achieving this indicator would improve organizational incentives to engage in evaluation-related tasks, as work performed voluntarily would be formally recognized.

24. EVAL seeks to build an incentive structure whereby the work of DEFPs, evaluation managers and internal evaluators is systematically recognized in job descriptions and in annual performance appraisals. The EMCP and IECP certification processes currently

[8] The ninth EMCP will be held in October 2018.

provide them with professional recognition (see biennial milestone 1.2.1), which is considered to be an incentive. However, there is scope for improvement, particularly as these tasks are in addition to their regular duties.

25. EVAL and the Human Resources Department have agreed to cooperate on this issue and are discussing the way forward. Areas for consideration focus on the inclusion in performance appraisals of generic outputs and performance indicators for evaluation tasks being carried out by DEFPs, evaluation managers and internal evaluators (ensuring regular and consistent recognition of their weight); and developing standardized texts on evaluation tasks to be included in job descriptions of DEFPs who have longer term roles in the evaluation network.

Sub-outcome 1.3 – Constituents engaged in monitoring and evaluation of DWCPS and development cooperation activities in an SDG-responsive manner

<table>
<tr><td>Indicator 1.3: Relevant monitoring and evaluation training is mainstreamed into training and capacity-building programmes for constituents in order to enhance their participation in evaluations.</td><td>Baseline: During 2010–17, 1,052 constituents were trained, 124 of them in 2016.</td><td>Status
Achieved.</td></tr>
<tr><td colspan="2">Biennial milestone (2018–19): Evaluation training and capacity-building modules responsive to SDG issues developed for mainstreaming into programmes, covering all three constituent groups.</td></tr>
</table>

26. Over the last strategic evaluation period (2011–17), EVAL has strengthened the capacities of over 1,000 constituents in respect of evaluation and related topics. The 2030 Agenda for Sustainable Development (2030 Agenda) calls for wider engagement by national stakeholders in country reviews on progress towards the achievement of SDGs and the required capacities for such involvement. To this end, the ILO's results-based evaluation strategy 2018–21 includes an indicator to measure EVAL's efforts to provide constituents with access to relevant capacity-building initiatives so that they can engage effectively in decent work and national SDG reviews.

27. Following a thorough preparatory process (an in-depth review of the professional literature and a validity assessment in the form of a survey), [9] EVAL recently completed the design of a training programme on the required monitoring and evaluation (M&E) capacities for assessing progress towards decent work and the SDGs. The results of the needs assessment showed that ILO constituents are keen to both play an active role in the review processes and to acquire the necessary capacity.

28. The training curriculum will be implemented in partnership with the Bureau for Workers' Activities (ACTRAV), the Bureau for Employers' Activities (ACT/EMP) and other key departments, in conjunction with needs identified through the DWCP evaluability diagnostic instrument (see biennial milestone 1.4). A collaborative initiative with the Multilateral Cooperation Department (MULTILATERALS), incorporating modules on evaluability and evaluation concepts in their training programme on the SDGs took place in September 2018. Throughout the remainder of 2018 and during 2019, EVAL will seek to expand this collaborative network to national training institutions interested in joining forces in building national evaluation capacity.

[9] The survey reached out to 2,273 constituents, with an overall response rate of close to 15 per cent.

Sub-outcome 1.4 – Evaluation integrated in DWCPs and development cooperation activities, including a focus on SDGs

Indicator 1.4: Number of DWCPs and development cooperation projects that have well-established evaluation processes and mechanisms in place and that regularly engage with constituents in meeting monitoring and evaluation requirements.	**Baseline**: No baseline yet established.	**Status** Achieved.
Biennial milestone (2018–19): Process is developed and piloted to a sample of DWCPs for ensuring that DWCPs and projects have mechanisms (diagnostic instruments) to assess their evaluability, SDG-responsiveness and level of participation of constituents in monitoring and evaluation.		

29. Like every United Nations agency, the ILO is required to consider how the Organization will measure its contribution to the SDGs. Reflecting on what the 2030 Agenda means for the ILO at the operational level, the evaluation policy 2017 and its complementary results-based evaluation strategy 2018–21 incorporated new and specific milestones and targets to assess the ILO's preparedness for measuring the contribution of the Organization to the SDGs. [10]

30. Evaluability assessments are RBM tools that can serve as an early warning mechanism to identify the strengths and weaknesses of the design and M&E of projects and programmes. Applied during the design of DWCPs and programme activities, the evaluability diagnostic instrument supports the ILO in analysing the linkages between DWCPs and SDGs, and their monitoring and reporting capacity. In February 2018, EVAL released an evaluability diagnostic instrument [11] to enable the rapid assessment of SDG integration into DWCPs at country level.

31. This has been pilot tested in Argentina and Sri Lanka with the help of the regional and country offices concerned. Further testing is being considered in Iraq and Suriname to align the tool in the most efficient manner with the ILO's ongoing initiative to support the development of a new generation of DWCPs that reflect the aspiration of ILO constituents to achieve social justice through decent work, and position them as effective vehicles for implementing the 2030 Agenda. Concurrently, EVAL is collaborating with the Strategic Programming and Management Department (PROGRAM) to integrate the instrument, or parts of it, into existing guidance material for use in drafting DWCPs. This will contribute to reinforcing the evaluability components of DWCPs to enhance responsiveness to SDGs and the level of participation of constituents in M&E.

[10] Earlier studies by EVAL had confirmed that the extent to which a project or programme can be evaluated in terms of its contribution to the SDGs (evaluability) and constituents' participation in related M&E exercises would be crucial to assess SDG responsiveness. ILO: *The implications of the SDGs on ILO's results framework – remarks from an evaluative perspective* (Geneva, Evaluation and SDG Series, 2016).

[11] The diagnostic instrument provides country-level analysis in three dimensions: (i) the extent to which a DWCP is designed, implemented, monitored and reported on to illustrate ILO country-level contribution to the SDGs; (ii) the degree to which country-level M&E is ready to implement, monitor and evaluate the SDGs (focusing on decent work); and (iii) constituents' capacity needs, and the challenges to effectively advocate for and engage in SDG/decent work M&E processes.

Sub-outcome 1.5 – Established capacity of regions and departments to mainstream and use evaluation

Indicator 1.5: Evaluation-related initiatives taken by regions and departments other than mandatory requirements systematized.	**Baseline**: Examples of such initiatives and their use have not been systematically documented since the AER 2015.	**Status**
		Not yet started.
Biennial milestone (2018–19): Systematic documentation of such initiatives, establishing good practices based on the experience of large or flagship programmes.		

32. Recent AERs reported regularly on the uptake of evaluation-related initiatives in regions and departments. The new evaluation strategy includes an indicator making the measurement of evaluation-related activities and initiatives undertaken by departments and regions compulsory. Such activities illustrate the role of evaluation in enhancing projects, programmes and other activities in an effort to support RBM and organizational learning. The list below illustrates some evaluation-related initiatives undertaken by regions and departments during the reporting period:

- meta-analyses, synthesis reviews and stocktaking in specific thematic areas, globally or within regions; [12]
- use of comprehensive M&E strategies and systems across projects and within programmes; [13]
- impact evaluations and impact assessments; [14]
- specific regional thematic evaluations informing thematic plans and strategies; [15]
- specific initiatives to strengthen M&E evaluation capacity in the ILO in the regions, including theory of change, M&E and learning frameworks; [16]
- evaluative reviews of programming and implementation modalities. [17]

33. Systematic reporting on such initiatives will demonstrate the enhanced role of evaluation throughout the Organization, provide examples for other departments and regions, and

[12] For example, ILO: *Meta-analysis of evaluations on ILO child labour programmes and projects in Africa 2009–2014* (Geneva, 2017); ILO: *Stocktaking of ILO support to extend the coverage and effectiveness of social security systems in Asia and Pacific 2006–2015* (Geneva, 2016); and ILO: *Stocktaking of ILO technical assistance and development projects on labour reform in Pacific Island countries 2012–2017* (Geneva, 2018).

[13] For example, the unpublished social protection impact management tool, and the unpublished better work monitoring and evaluation framework for strategic use of evaluation.

[14] For example, ILO: *Impact evaluation of pilot fair recruitment in the garment section in Nepal-Jordan* (ongoing).

[15] For example, ILO: *Asia regional thematic evaluation of social protection 2012–2017* (Geneva, 2017).

[16] For example, the initiative by the ILO Regional Office for Asia and the Pacific on the theory of change in DWCPs starting in three pilot countries.

[17] For example, the review by the ILO Regional Office for Asia and the Pacific of the DWCP tripartite governance mechanisms in Asia and the Pacific.

facilitate cross-organizational learning. The next AER will report on the first results of the systematic documentation.

Outcome 2: Enhanced value of evaluation through the use of more credible and higher quality evaluations (independence, credibility, usefulness)

34. In an effort to realize targets listed under outcome 2, it is necessary to develop innovative methodologies indicative of the ILO's mandate, and monitor and evaluate the ILO's contribution to the SDGs, which can be achieved by applying a more strategic focus and by clustering evaluations. This would give more time and resources for higher quality evaluations. Successfully addressing the targets would help to ensure that the evaluation function continues its upward trajectory as a maturing and high-quality operation in the United Nations system.

Sub-outcome 2.1 – Use of strategic cluster evaluations to gather evaluative information more effectively

<table>
<tr><td>Indicator 2.1: Strategic cluster evaluations established as a modality in a substantial proportion of programmes and projects.</td><td>Baseline: Currently, no documented processes or procedures are in place to conduct strategic cluster evaluations for development cooperation projects.</td><td>Status
On track.</td></tr>
<tr><td colspan="2">Biennial milestone (2018–19): Methodology and procedure developed for strategic cluster evaluations, including a modality for pooling evaluation funds, and piloted in at least five projects.</td></tr>
</table>

35. The new evaluation strategy introduces an exclusive indicator calling for strategic cluster evaluations to be established as a modality in a substantial proportion of programmes and projects. It thus underscores the need for the evaluation function to develop an appropriate methodology and procedures responsive to the ILO's strategic focus. The 2016 IEE [18] also reflected on the need to use cluster evaluations to allow for more strategic focus and more effective use of evaluation funds.

36. Cluster evaluations have been mostly applied in instances when global programmes or projects have similar approaches and common outcomes. EVAL is undertaking a consultative study to review the options, methodologies and procedures needed to cluster evaluations in an effective, efficient and appropriate manner. The study will address the following issues: principles and criteria for clustering; purpose and scope of various types of clustering (for example, how they might provide for more comprehensive coverage of DWCPs); complementarity with the donor and ILO mandatory evaluation requirements; appropriate evaluation approaches/modalities; and implementation issues, including, where possible, the pooling of evaluation funds.

37. The outcome of the study will be pilot tested. The final aim is to potentially establish a procedure for clustering evaluations which, given the right conditions, could be increasingly used as the main modality. EVAL has requested the Partnerships and Field Support Department (PARDEV) to promote consistency in the way evaluation is covered in donor agreements and to advocate for clustered evaluations as an option. This will allow the initial

[18] ILO: *Independent evaluation of ILO's evaluation function, 2011–2016: Final report* (Geneva, 2016).

pilot testing and the eventual rollout of this modality, provided mechanisms and support for pooling of resources can be secured and implemented.

Sub-outcome 2.2 – Improved quality of internal, decentralized and centralized evaluations

Indicator 2.2.1: All evaluations of development cooperation projects comply with OECD and UNEG norms and standards and are tailored to the ILO's specific mandate and learning needs.	**Baseline**: External quality assessment for 2015–17 shows that about 90% of development cooperation project evaluations meet the required quality standards.	**Status** On track.
Biennial milestone (2018–19): Guidelines will be updated to incorporate new evaluation models that reflect the ILO's specific mandate while maintaining quality.		

38. The ILO's new evaluation policy and the evaluation strategy identify quality control as a continued important task for EVAL. At the decentralized level, evaluations involve various ILO officials, many of whom are volunteers, taking on a variety of evaluation-related duties (see biennial milestone 1.2.2). EVAL assures the quality of independent project evaluations by creating detailed evaluation guidelines, conducting training (EMCP and IECP), and employing a layered evaluation approval process. EVAL is accountable for the independence and quality of all independent evaluations.

39. EVAL's internal quality control system is complemented by ex post quality appraisal (QA) mechanisms that are conducted by external appraisers on a regular basis. Since 2007, a total of eight QAs have been conducted for independent project evaluations. The results of the last QA (2017) found that more than 90 per cent of the reports assessed met the minimum level of overall quality. Despite this positive result, analysis showed that the evaluation methodology and evaluators' efforts to mainstream gender into evaluations still left ample scope for improvement.

40. EVAL intends to begin conducting QAs on a rolling basis once a decentralized evaluation is completed and endorsed by the senior evaluation officers. The external QA consultants will use a rating protocol and assign a QA score for each report, which will in turn validate EVAL's review. In addition, the QA consultants will send a short questionnaire to the evaluation manager concerned to determine compliance with evaluation policy guidelines. Such an approach would enable EVAL to quickly detect quality issues, take prompt action to address them, and aggregate results on a regular basis for reporting purposes.

41. Recognizing that the new evaluation policy and strategy call for new evaluation models that reflect the ILO's mandate, EVAL is developing guidance on this matter (see biennial milestone 2.4.1). The guidance will also be informed by previous results from QA, particularly in terms of their lessons, in an effort to strengthen the quality of evaluation reports.

Indicator 2.2.2: Additional capacity released in EVAL at headquarters to focus on new evaluation models by reducing oversight of regional evaluations of development cooperation projects.	**Baseline**: The 2016 IEE identified the issue of independence at the regional level as a priority and recommended the integration of regional evaluation officers (REOs) as full staff members of EVAL.	**Status** On track.
Biennial milestone (2018–19): Preparation of a detailed report that analyses reporting lines for REOs and includes a presentation of possible scenarios, with the aim of ensuring the highest level of independence.		

42. The 2016 IEE reported that the highest priority regarding independence was to review the reporting lines of REOs. [19] The new evaluation strategy also recognizes its importance by including an indicator on the topic.

43. The IEE attributed chronic under-investment in monitoring and reporting to a situation in which REOs have had to take on these tasks, thus substantially increasing their workload. Moreover, this dual role also potentially affects their independence and decreases the amount of time they can appropriately dedicate to supporting the evaluation function.

44. To address this concern, the IEE recommends enhanced structural independence for REOs. The responsibility for final approval of decentralized independent evaluations could then be delegated to REOs, giving senior evaluation officers more time to oversee a greater volume and diversity of strategic evaluations. [20] EVAL has formally communicated with management to cooperate on a comparative study that would examine REOs current reporting lines and tasks, and to explore alternative scenarios available within the ILO Staff Regulations that could be applied to enhance their independence.

<table>
<tr><td>Indicator 2.2.3: Corporate governance-level evaluations incorporate UNEG norms and standards and are tailored to the ILO's specific mandate and learning needs.</td><td>Baseline: Independent review in 2013 confirmed quality met required standards as reconfirmed by the 2016 IEE.</td><td>Status
Not yet started.</td></tr>
<tr><td colspan="2">Biennial milestone (2018–19): Protocols will be updated to incorporate new evaluation models that reflect the ILO's specific mandate while maintaining quality.</td></tr>
</table>

45. High-level evaluations of strategies, programme and budget policy outcomes, institutional issues and DWCPs have always been based on the ILO's evaluation policies and EVAL's policy guidelines, including protocols developed to ensure consistency in approach and performance grading. These protocols have been based on years of experience, and respect for United Nations Evaluation Group (UNEG) norms and standards. As confirmed by the 2016 IEE, evaluation management procedures and a thorough quality assurance process have ensured that ILO evaluations meet international standards. Annual stocktaking of experience in high-level evaluations has led to an incremental improvement in the approach.

46. As new evaluation models reflecting the ILO's mandate are developed and as more strategic clustering of evaluations takes place, briefing and procedures for the management of these evaluations will be updated. The protocols will also need to reflect the evolving nature of the ILO results framework, the strategic planning and programming process, and the proposed pathway. This will ensure that evaluative information, including high-level evaluations, provides the basis for assessing the ILO's development effectiveness and provides key information for the ILO's organizational learning in support of constituents.

[19] REOs are designated evaluation officers in each ILO regional office. They are responsible for overseeing evaluations within their given regions. They also oversee and advise on the process of planning, managing and follow-up of DWCP reviews and project evaluations. They are evaluation professionals dedicated to supporting evaluation work in the ILO.

[20] Currently, one of the responsibilities of senior evaluation officers located in EVAL is to assess quality and independence standards of decentralized evaluations.

Sub-outcome 2.3 – Credible impact evaluations conducted to build knowledge for effective policy interventions

Indicator 2.3.1: Impact evaluations are considered credible and used for documenting effective policy interventions.	**Baseline:** Quality of impact evaluations not optimal or uniform, as indicated in EVAL stocktaking report of 2014. A new ex post quality analysis of a sample of impact evaluations, to be carried out in 2018, will establish a new baseline.	**STATUS** On track.
Biennial milestone (2018–19): Improved impact evaluations by technical departments and ILO offices as a result of improved technical support by EVAL and increased conformity with EVAL's guidance for 50% of impact evaluations.		

47. The current evaluation strategy recognizes the importance of impact evaluations to provide evaluative evidence on effective policy solutions that the ILO can promote and support. As part of knowledge building, various technical departments and regions carry out evaluations and assessments focusing on level of impact. EVAL provides technical support and ensures quality assurance in so far as its capacity permits. Departments and regions voluntarily submit evaluation proposals, designs and reports for EVAL's technical experts to review through its impact evaluation review facility.

48. An estimated 60 impact evaluations or studies have been carried out since 2009, but only a few have been reviewed through the facility established in 2016. An ex post quality review of a sample of these impact evaluations has been commissioned and is expected to provide a framework for more regular quality reviews. Using this framework, impact evaluations will be subjected to mandatory reviews and their documented quality can be used to further enhance the credibility of the ILO's knowledge base and technical support role.

Sub-outcome 2.4 – Evaluation framework further aligned with ILO mandate and context, including SDGs

Indicator 2.4.1: ILO-specific evaluation approaches, models and methods used for evaluations at various levels.	**Baseline:** Currently, minimal ILO-specific approaches and models are used in ILO evaluations.	**Status** On track.
Biennial milestone (2018–19): Pilot evaluation framework developed and used in five pilot evaluations; evaluation policy guidelines updated.		

49. Standards, social dialogue, equality and non-discrimination are central to the ILO's work. EVAL's existing protocols and methodologies provide the basic tools needed to evaluate interventions, taking into account the unique normative and social dialogue context of the ILO, as well as issues relating to gender equality and the inclusion of persons with disabilities.

50. Currently, EVAL is undertaking a study to assess ways and means of further aligning evaluation frameworks with the ILO's specific mandate. Initial findings have revealed that approaches, methodologies and frameworks need to be improved to enable evaluations to capture and assess these features. In terms of OECD/DAC evaluation criteria, social dialogue practices can clearly influence the relevance, effectiveness, efficiency, impact and sustainability of ILO projects. While normative foundations are indeed evident in all ILO projects, the degree of focus differs. Sometimes, norms are central to the intervention(s)

being evaluated, while in others they are peripheral. The research found that approaches, methodologies and frameworks need improving to enable evaluations to capture and assess these important contextual features, regardless of whether they are crucial or peripheral to an intervention, or whether they are intentional or unintentional.

51. EVAL QAs have also offered insights into improving evaluations (see biennial milestone 2.2) related to gender equality and the empowerment of women (GEEW). The overall GEEW ratings have remained low over the years: less than 10 per cent of evaluations have fully met the United Nations System-wide Action Plan on Gender Equality and Women's Empowerment (UN-SWAP) requirements. Whereas design-related issues need to be addressed on different fronts, the QAs have also provided EVAL with ideas on how to better mainstream gender concerns in evaluations (see sub-outcome 1.2 and biennial milestone 2.2.1). Taking a cue from the recommendations of 2016 IEE and the several duplications in the QA on GEEW indicators, EVAL is exploring other ways of strengthening GEEW aspects in evaluations. Revision of evaluation guidance currently under way places greater emphasis on integrating gender, disability and other discrimination-related issues.

Outcome 3: Stronger knowledge base of evaluation findings and recommendations

52. Building a stronger knowledge base of evaluation findings and recommendations requires the expansion and dissemination of targeted communication products and knowledge management tools, including, in particular launching an automated system to track management responses to evaluation recommendations and ensuring that *i*-eval Discovery captures all evaluation information. Combined, these initiatives will establish the ILO as a credible, transparent and evidence-driven Organization.

Sub-outcome 3.1 – Strengthened accessibility and visibility of evaluation information through i-eval Discovery

<table>
<tr><td rowspan="2">Indicator 3.1: i-eval Discovery contains all planned and completed evaluations, including recommendations, lessons learned and good practices; is consistently accessed by internal and external users; and is considered the gateway to ILO evaluation information.</td><td rowspan="2">Baseline: An initial baseline of nil.</td><td>Status</td></tr>
<tr><td rowspan="2">On track.</td></tr>
<tr><td colspan="2">Biennial milestone (2018–19): Further development and use of i-eval Discovery and the i-Track database to support targeted communication and use of evaluation information (target: 25% increase over baseline level).</td></tr>
</table>

53. EVAL places significant value on improving knowledge systems in order to promote the use of evaluation information. The new strategy recognizes this and an exclusive biennial milestone on the topic has been added, thus ensuring systematic reporting.

54. Launched in December 2016 and further refined in August 2017 to include a function for "planned evaluations", *i*-eval Discovery is an interactive and publicly accessible web-based application that displays on a map all planned and completed evaluations, in addition to their related recommendations, lessons learned and good practices. Data analyses show that there has been an increase in the number of users accessing the application. The number of users grew by an average of 61 per cent between August and December 2017, and January and

June 2018. [21] Therefore, the target has already been exceeded and will be revisited in the next reporting period.

55. By the end of 2018, the application will have been further developed to enable evaluations to be searched by donor and cluster. The underlying database for the *i*-eval Discovery application will be updated in 2019 to more efficiently capture evaluation information. New features will also be added, including information on evaluation clustering options and details on Regular Budget Supplementary Account (RBSA) evaluations, as requested by constituents. Furthermore, provided that the ILO's information technology infrastructure permits, *i*-eval Discovery will be linked to other ILO platforms in order to facilitate seamless access to Organization-wide results reporting.

Sub-outcome 3.2 – More targeted communication of evaluation findings

<table>
<tr><td>Indicator 3.2: Revised communications strategy leads to better targeting of evaluation findings to management, constituents and other users.</td><td>Baseline: The 2016 IEE recognized progress made (newsletter, think pieces, i-eval Discovery) but called for better presentation of evaluation findings to improve use.</td><td>Status
On track.</td></tr>
<tr><td colspan="2">Biennial milestone (2018–19): A communication and rebranding strategy is designed (target: 2018) and rolled out (target: 2019) in collaboration with the Department of Communication.</td></tr>
</table>

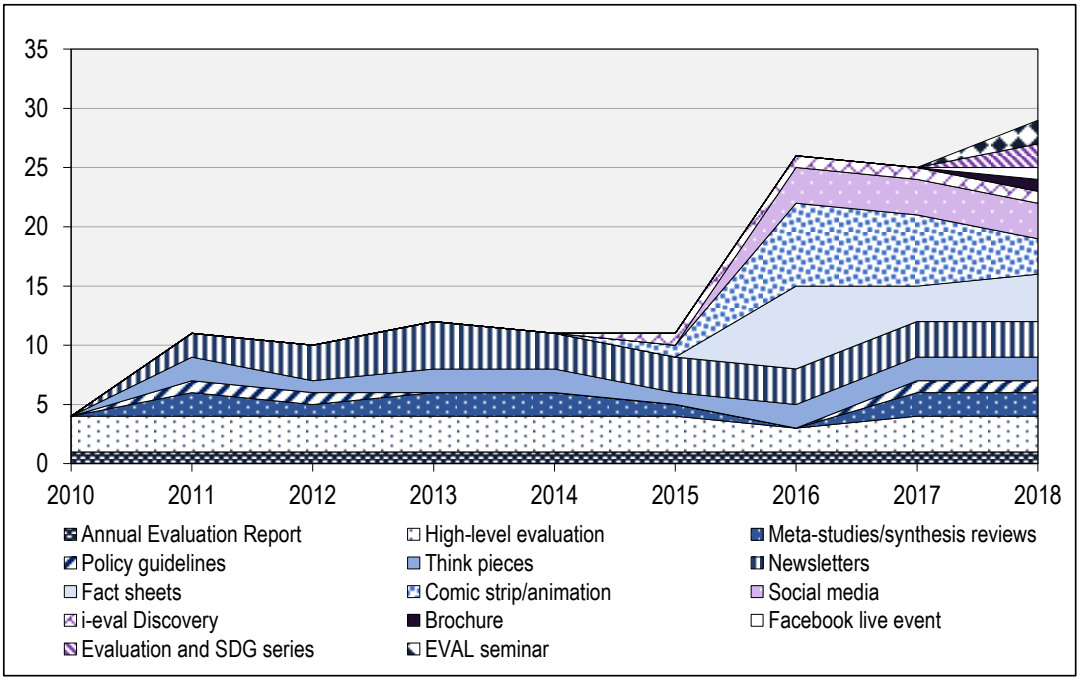
56. The development of targeted communication products on evaluation is essential to improving the uptake of evaluation findings, in addition to improving the ILO's evaluation culture. EVAL has regularly reported on such progress to the Governing Body through previous AERs. Recognizing the value of the link between communications and evaluation, the current strategy includes a specific biennial milestone on the topic, thereby systematizing reporting to the Governing Body for the first time.

57. EVAL is developing a communication plan to complement the evaluation strategy 2018–21. It will be informed by: suggestions from evaluation network members, a meta-analysis of recommendations from a sample of independent evaluations, and a survey covering a sample of ILO officials located in the field and at headquarters. The new communication plan will outline priorities and a detailed timeline for its rollout. This includes the identification of relevant communication tools to address each priority, in consultation with the Department of Communication and Public Information (DCOMM).

58. Over the past year, EVAL applied new tools to better communicate evaluation-related information, in addition to previous methods. Figure 2 illustrates all products and services that were developed by EVAL in the period 2011–18. The results show a positive trend that EVAL expects to grow as the new communication plan is rolled out.

[21] From an average of 2,540 users to 4,175 users.

Figure 2. EVAL products and services, 2011–18



Sub-outcome 3.3 – Improved use of evaluation findings and recommendations by constituents and management for governance and decision-making

Indicator 3.3.1: EAC advice on timing and use of evaluation prompts more robust uptake of evaluation findings for policy and strategic decisions at the global and regional levels.	**Baseline:** The EAC met on average four times per year and qualitative analysis showed it held strategic debates on about 40% of the corporate governance-level evaluations. Although the regions participate in the EAC, there are no regional evaluation advisory committees (REACs).	**Status** On track.
Biennial milestone (2018–19): The EAC continues to meet on a consistent basis (four times annually) and has strategic discussions on 50% of the corporate governance-level evaluations. By early 2019, a report on the added value of REACs will be produced. Subject to the outcome of that review, by end-2019 two regions will have piloted an REAC.		

59. The Evaluation Advisory Committee (EAC) was established in 2006 to oversee the use and implementation of, and follow up to, lessons learned and recommendations resulting from ILO evaluation activities. The previous evaluation strategy contained a biennial milestone calling for the EAC to hold at least four well-documented meetings per year to discuss the follow-up to high-level evaluations mandated by the Governing Body. The milestone in the new evaluation strategy is more specific and calls for four meetings per year and a qualitative analysis of the EAC's strategic discussions.

60. The percentage of strategic discussions used as a baseline in the results framework was determined by analysing the proceedings of the EAC meetings. When follow-up workplans or progress reports of governance-level evaluations are considered inadequate and returned, or when concerned departments or regional directors are requested by the EAC to provide additional verbal briefings, it is an indication that the EAC has engaged in a strategic discussion.

61. Table 3 shows the scope of the EAC's decisions during the current reporting year; 75 per cent of the high-level evaluations reviewed during the reporting period were the subject of strategic discussions. This figure, and the four meetings that took place, suggest that the EAC is making good progress towards accomplishing the biennial milestone.

62. In early 2018, the EAC agreed to conduct a review into expanding its composition to include representation from all regions. The new evaluation strategy includes a biennial milestone to review the creation of the Regional Evaluation Advisory Committees (REACs) suggested in the 2016 IEE. Thus, EVAL will undertake a study into the added value of the proposed REACs and, if considered appropriate, two regions will be asked to pilot an REAC for review by the end of the biennium.

Table 3. EAC decisions on high-level evaluations in 2018

High-level evaluation	Status of follow-up workplan	Review of actual follow-up
Independent evaluation of the ILO's DWCPs, strategies and actions in the Western Balkans (2012–15).	Completed.	In February 2017, the EAC returned the Western Balkans follow-up workplan for revision because not all recommendations had been addressed. The workplan was approved in May 2017 and a progress report was requested. The progress report was approved in May 2018 with no additional follow-up reporting required.
Independent evaluation of the ILO's DWCP strategies and actions in the Mekong subregion, 2012–17. *	Approved.	The EAC approved the follow-up workplan after initial review in February 2018. A progress report was scheduled for six months later.
Independent evaluation of the ILO's strategy and actions for creating and extending social protection floors, 2012–17.	Approved.	In February 2018, the EAC returned the follow-up workplan for revision because not all recommendations were addressed. It was subsequently approved in May 2018 with a request that a progress report on action taken be presented to the EAC in six months.
Independent evaluation of the ILO's field operations and structure 2010–16.	Not yet approved.	In February 2018, the EAC returned the follow-up workplan for revision. In May 2018, the Deputy Director-General for Field Operations and Partnerships (DDG/FOP) was again invited to revise, resubmit and include action steps to ensure subsequent implementation. DDG/FOP has been invited to resubmit the revised workplan to the EAC during its next meeting.

* Because the Mekong workplan was approved upon initial review, it did not count as a strategic discussion.

Indicator 3.3.2: Enhanced follow-up to evaluation recommendations through systematic monitoring.	**Baseline:** Follow-up to management response stood at 83% in 2016 (partially addressed and completed).	**Status**
		On track.
Biennial milestone (2018–19): An automated online application for management to follow-up evaluation recommendations will have been established, improving overall efficiency and maintaining a high follow-up response rate (target: 85%).		

63. The evaluation policy 2017 requires line management to submit a response to evaluation recommendations, irrespective of whether they are independent or internal. EVAL has historically tracked and recorded management responses in previous annual reporting to the

Governing Body, in addition to indicating the extent to which recommendations were completed, partially addressed or on which no action was taken. This is the first strategic milestone that captures the level of management follow-up of evaluation recommendations, marking a significant departure from the previous strategy.

64. During the first quarter of 2018, EVAL launched an automated online application to track and report on the status of management responses, with support from the Information and Technology Management Department (INFOTEC). It incorporates data visualization tools and clarifies reporting lines, in an effort to support RBM across the ILO.

65. For the period under review, all management responses to the 427 recommendations stemming from 41 independent evaluations were received. Altogether, 88 per cent of recommendations were completed or partially completed, thus surpassing the biennial target of 85 per cent (table 4 and figure 3), in addition to representing a modest increase from the previous reporting period (82.7 per cent). Such data reflects a strong commitment from project management to RBM and organizational learning. A review of recommendations for the reporting period reveals that the majority are thematically linked to organizational issues (67 per cent), followed by working conditions and equality (11 per cent), as outlined in figure 4.

Table 4. Management response to evaluation recommendations, 2017 [22]

Administrative region or sector	Evaluation reports requiring management response (41)		Recommendations		Completed	Partially addressed	No action taken
	Management responses received		In reports	With responses			
	No	Yes					
Africa	0	5	64	64	38	20	6
Americas	0	6	56	56	27	26	3
Arab States	0	1	9	9	5	3	1
Europe and Central Asia	0	4	42	42	18	16	8
Asia and the Pacific	0	18	195	195	97	66	32
Subtotal	**0**	**34**	**366**	**366**	**185**	**131**	**50**
BETTER WORK	0	1	15	15	11	1	3
DEVINVEST	0	1	7	7	2	5	0
EMP-SEED	0	1	7	7	7	0	0
FUNDAMENTALS	0	1	10	10	9	1	0
MIGRANT	0	1	8	8	7	1	0
WORKQUALITY	0	2	14	14	4	9	1
Subtotal	**0**	**7**	**61**	**61**	**40**	**17**	**4**
Total	**0**	**41**	**427**	**427**	**225**	**148**	**54**
Percentage (%)	**0**	**100**	**0**	**100**	**53**	**35**	**12**
						88%	

[22] Some departments/offices do not appear in the table, either because their administrative responsibility was decentralized to the regions or because the project was below the budget threshold for management follow-up. In addition, a response of "no action taken" can refer to the recommendation being out of the ILO's scope of action.

Figure 3. Evolution of the follow-up to evaluation recommendations, 2011–17

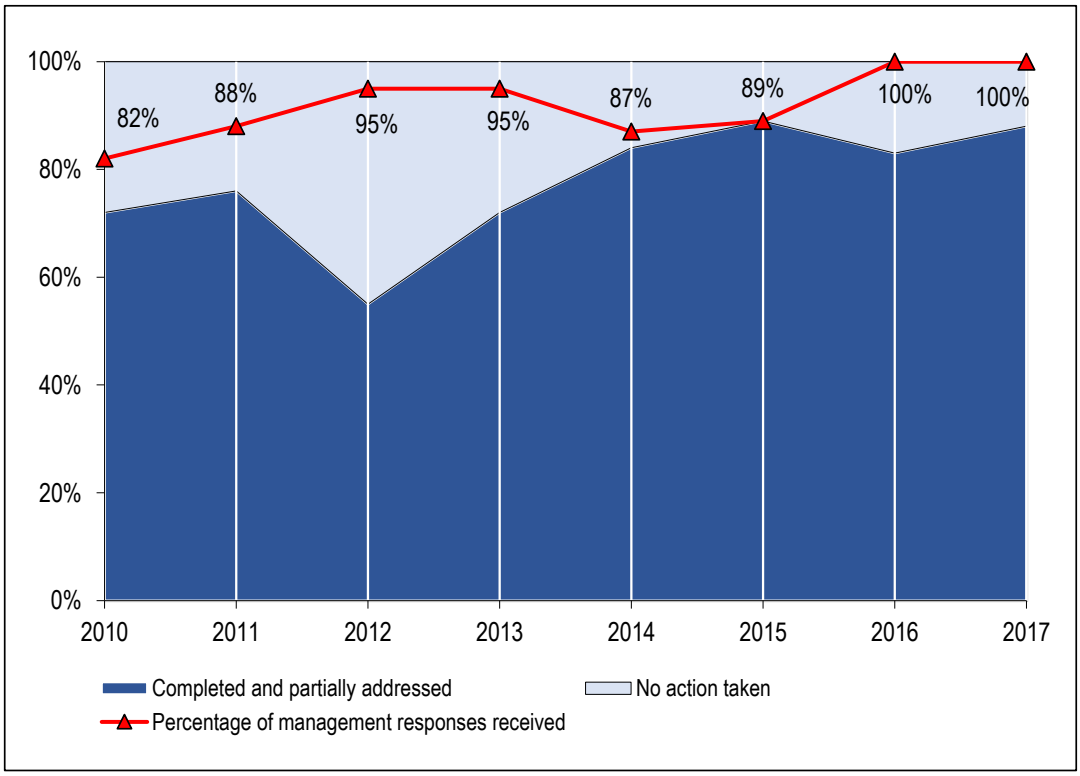
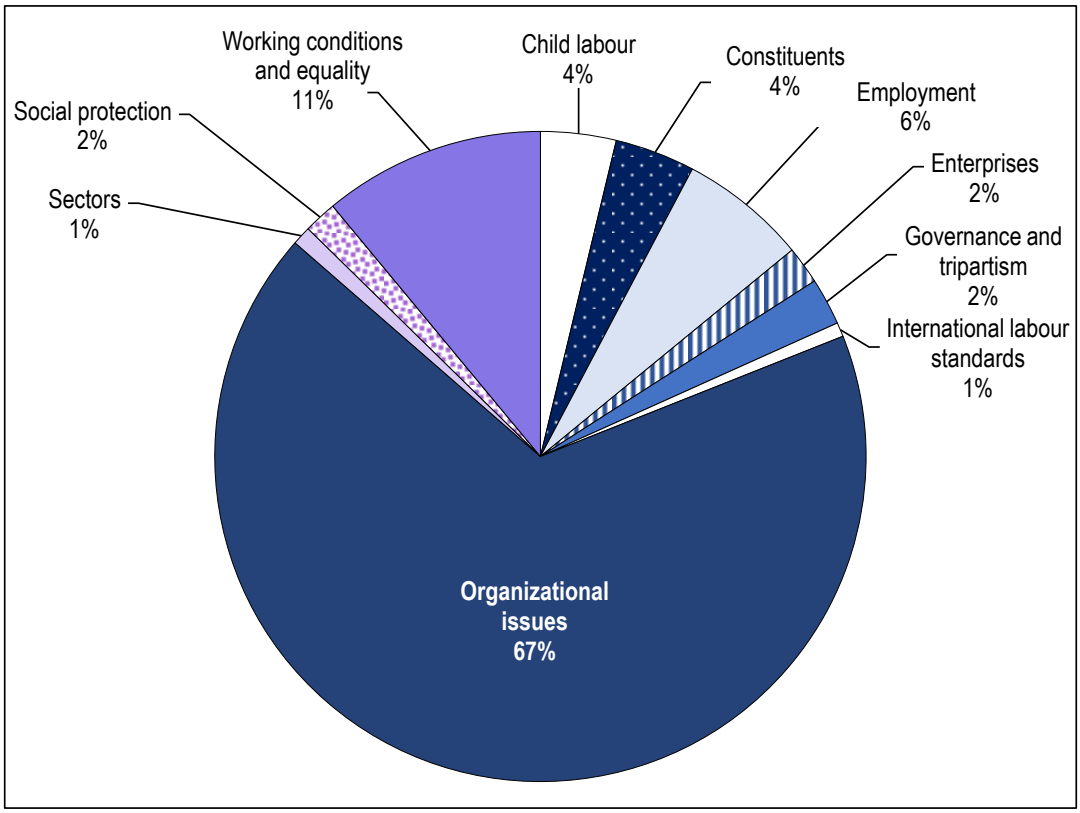


Figure 4. Recommendations by theme/technical subject, 2017



Indicator 3.3.3: Enhanced use of evaluations in strategic guidance, reviews and reporting for strategic plans, programme and budget records and other high-level plans and strategies.	**Baseline:** In AER documents, the use of recommendations and lessons learned from evaluations (40–50% for 2010–15, based on stocktaking exercise).	**Status** Not yet started.
Biennial milestone (2018–19): By end-2019, 75% of evaluation recommendations and findings are fully or partially reflected in relevant strategic guidance and reporting (for example, implementation reports, 2020–21 programme and budget reports and other strategic and programmatic documents).		

66. It is important that evaluation findings and recommendations are used to inform strategic guidance and reporting. The indicator and its milestone in the new strategy require more complex quantitative annual reporting than in previous reports. The measurement methodology under development will include a combination of quantitative and qualitative methods to measure use. This will require the analysis of multiple Governing Body documents reflecting organizational decision-making, and surveys of, and consultations with, departments and regions on the use and follow-up of evaluations.

67. The following examples are illustrative: (i) there were twice the number of references to high-level evaluations in ILO programme implementation report 2016–17 than in 2014–15; (ii) the Knowledge Strategy 2018–21, approved in 2017, specifically referred to a high-level evaluation; (iii) preliminary analysis of references to evaluation in 2014–17 strategic documents shows that 23 per cent refer to the explicit use of evaluation findings, 13 per cent refer to the importance of evaluations in informing decision-making, 10 per cent refer to the need for evaluations on new topics, and the remainder refer to evaluation as a topic.

68. A more concrete example is the independent evaluation of the ILO's Action Plan for Gender Equality (2010–15) and how it fed into the report on the implementation of the ILO Action Plan for Gender Equality 2016–17, and subsequently informed the new ILO Action Plan for Gender Equality 2018–21.

Sub-outcome 3.4. – Evaluations used to meet strategic knowledge requirements through further analysis of findings and results of evaluations

Indicator 3.4.1: Evaluation findings analysed, synthesized and documented in knowledge products in support of planning and knowledge building.	**Baseline:** In the previous strategy period, 22 think pieces, meta-studies and synthesis reviews were carried out.	**Status** Not yet started.
Biennial milestone (2018–19): Process established to determine topics in line with strategic knowledge requirements, maintaining an average of at least three studies per year.		

69. Evaluation reports and their findings are inputs to organizational learning and knowledge building. Knowledge products, such as meta-studies and synthesis reviews, analyse evaluation findings within specific themes and areas of work to identify systemic issues, lessons learned and good or bad practices, or to assess overall performance. They are meant to address the strategic knowledge requirements for higher-level decisions and actions. Since 2011, EVAL has produced 22 meta-studies, synthesis reviews and think pieces.

70. The new evaluation strategy's focus is to initiate a more systematic process for scanning and mapping knowledge requirements that can be addressed through these studies by: reviewing and mapping relevant governance, strategy and programming documents; creating an inventory of knowledge gaps; and reviewing the ILO's governance and policy agenda. Thus,

EVAL will develop a more demand-driven process to develop knowledge products from the many evaluations undertaken annually.

Review of overall effectiveness of the ILO

Indicator 3.4.2: The AER provides annual overview of overall effectiveness of the ILO.	**Baseline:** Analysis of decent work results and effectiveness of ILO development cooperation completed, covering 2009–16 with ongoing revision of methodology.	**Status**
		Not yet started.
Biennial milestone (2018–19): Analysis conducted for 2017 and 2018, providing a synthesis on the ILO's effectiveness; methodology further revised to facilitate regular analysis and reporting in the AER.		

71. Since the introduction of an evaluation strategy in 2011, a performance indicator calling for an overview of findings and issues related to the overall effectiveness of the ILO has been part of EVAL's annual reporting.

72. The new strategy is specific in its requirement for a regular analysis of ILO effectiveness using improved measurement methodologies. Decent work results assessments are normally undertaken every two years to provide for a sufficient sample of evaluation data and the continued evolution of the ILO's results framework. These initiatives and processes are expected to converge into a new methodology that will facilitate regular analysis and reporting to the Governing Body. This year's traditional chapter in the AER on effectiveness (Part II) therefore reflects on new approaches that could be explored for a more comprehensive validation process of ILO performance at the country and global levels.

Part II. Assessing the ILO's effectiveness and results

73. Part II of the report provides an overview of EVAL's contributions to initiatives aimed at improving RBM in the ILO. Since 2011, EVAL has provided annual assessments of selected areas of the ILO's effectiveness, focusing on strategic relevance, effectiveness, impact, sustainability and efficiency. The new evaluation strategy 2018–21 calls for the continuation of these annual overviews and points to the need for a methodological review to enable more comprehensive reporting and validation of the ILO's decent work results and effectiveness. [23]

A challenging but opportune time to innovate the way we assess overall effectiveness and results

74. The new evaluation strategy is being introduced at a time that is both challenging and opportune. It can bring originality to the way in which EVAL functions. The 2030 Agenda is already in its third year of implementation and there is growing emphasis on the need to streamline the monitoring, evaluation and reporting of SDGs at the country and global levels. These concerns are reflected in the ILO's policy and governance documents. For instance, the resolution concerning effective ILO development cooperation in support of the SDGs, [24] adopted at the 107th Session (2018) of the International Labour Conference, reiterates the centrality of DWCPs as the primary vehicle for supporting member States in achieving the

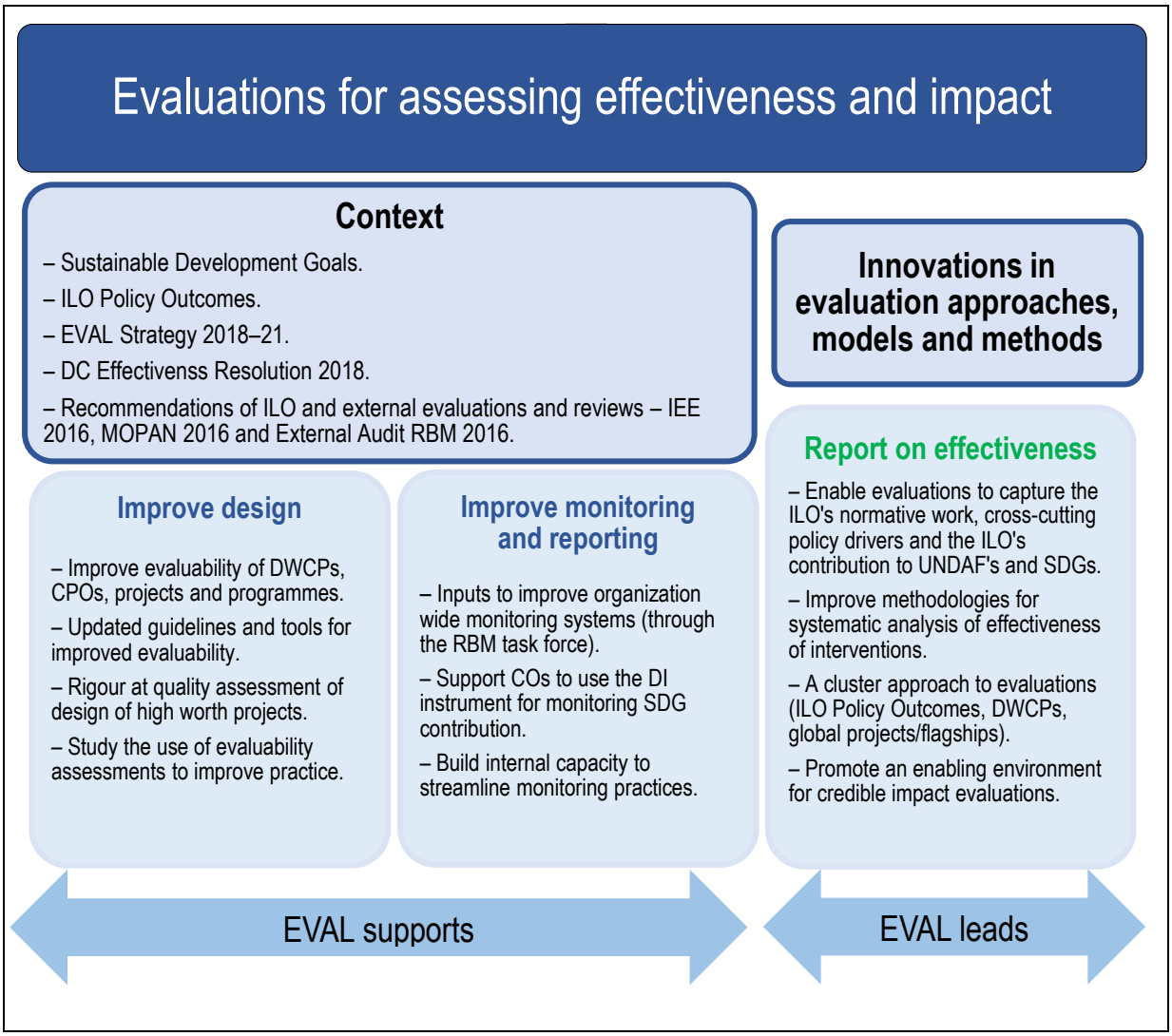
[23] Biennial milestone 3.4.2 for 2018–19.

[24] Adopted by the International Labour Conference at its 107th Session, Geneva, 2018, https://www.ilo.org/ilc/ILCSessions/107/reports/texts-adopted/WCMS_633138/lang--en/index.htm.

SDGs. It also calls for a "better use of data collection, results-based management tools and evaluation, including impact evaluations, to demonstrate what works, support the scaling up of interventions and enhance the visibility of the Decent Work Agenda". These developments, combined with the ongoing work of the Office's internal RBM task force, provide a unique opportunity to innovate.

75. EVAL has a leading role in reporting on the effectiveness of interventions. This has been accomplished by applying a continuous cycle of feedback across evaluations and by providing more direct solutions through specific products. These include evaluability assessment tools, the SDG diagnostic instrument and inputs to project design and M&E frameworks at the appraisal stage of project documents. EVAL also works on improving the quality, comprehensiveness, coverage and analysis of evaluations so as to provide source material when reporting on effectiveness. Figure 5 illustrates the interconnectedness of these elements and the efforts made thus far towards strengthening design, monitoring and effectiveness.

Figure 5. Assessing effectiveness and impact: Key elements and EVAL's role



Support proper design, monitoring and reporting on programme and budget policy outcomes and DWCPs as the programmatic backbones of the ILO's Decent Work Agenda

76. Programme and budget policy outcomes and DWCPs demand effective linkages to the SDGs and should be well equipped to allow for robust monitoring, reporting and evaluation of their contribution to the SDGs and their indicators. Building capacities to monitor and evaluate the collective contributions to the Decent Work Agenda and the SDGs is as important as leveraging national and international partnerships to advance decent work goals. A diagnostic instrument, developed by EVAL, is already available to help assess and improve the evaluability of DWCPs in the context of the SDGs. Parallel to this, EVAL has developed a training programme for ILO constituents to build their capacity to monitor and evaluate DWCPs and SDGs (see biennial milestone 1.3). EVAL is also supporting capacity-building partnerships to enhance national capacity to conduct systematic country-led follow-up and reviews.

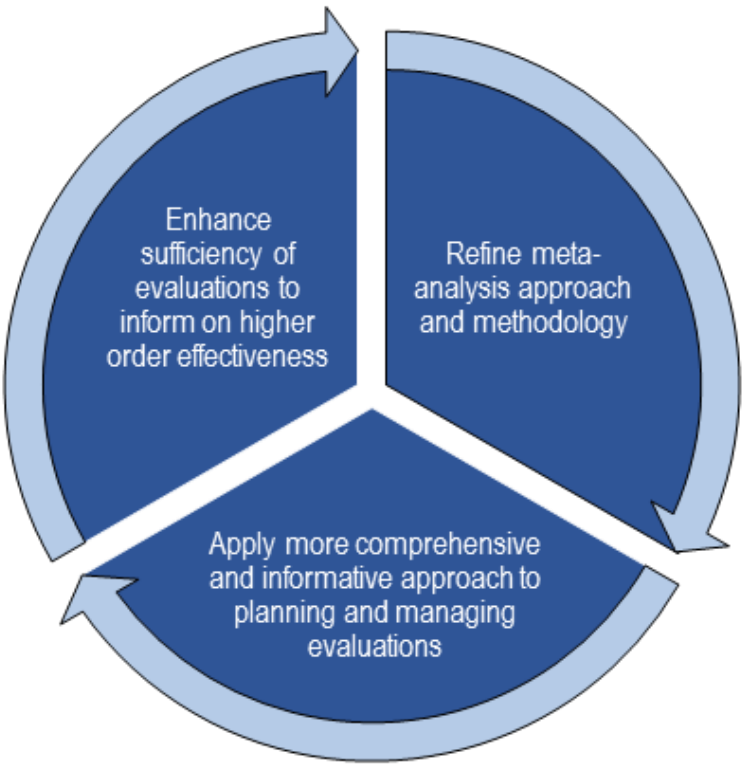
77. These efforts, however, have to be backed by improved monitoring and reporting practices to ensure effective evaluation of the contribution of policy outcomes and DWCPs to the Decent Work Agenda and SDGs. Past AERs have stressed that weak monitoring and reporting capacity within projects and programmes limit the availability of data as well as the range of methods and questions that can be considered to evaluate development effectiveness at the level of delivering decent work results, as well as to measure the impact of ILO interventions. The establishment of an RBM task force in the ILO and an increased emphasis on developing clear theories of change are important improvements to address these gaps. EVAL has been contributing to this agenda, but evaluation reports continue to highlight two key issues that need to be addressed to achieve the desired results: (i) strengthening of monitoring and reporting practices for extra-budgetary-funded activities; and (ii) establishing clear accountabilities for overseeing Office-wide compliance.

> **Recommendation 2:** Improve the Office-wide monitoring and reporting framework and practices for extra-budgetary-funded activities and assign clear accountability.

78. In recent years there have been increasing calls to document and report on the ILO's development effectiveness and the multilateral system in general. Previous AERs have presented findings on effectiveness by synthesizing the numerous evaluation reports and distilling recurrent and systemic findings. This is primarily done through a meta-analysis of project-level evaluations using a set of performance criteria and relating the findings to decent work results. While the current methodology of meta-analysis provides a valid indication of areas of stronger and weaker performance, it is not equipped to fully measure or validate effectiveness. Some of the main challenges are: the insufficient reflection of higher-level effectiveness in project evaluations; the disproportionate time and resource distribution involved in conducting project evaluations as compared to strategic evaluations; and the lack of resources for ex post evaluations and impact assessments.

79. Indeed, any comprehensive narrative on the ILO's effectiveness requires improvements at three levels, as illustrated in figure 6.

Figure 6. Reporting on the ILO's effectiveness: Potential improvements



80. The following subsections provide a narrative on the improvements required at each level and the ongoing efforts in this direction.

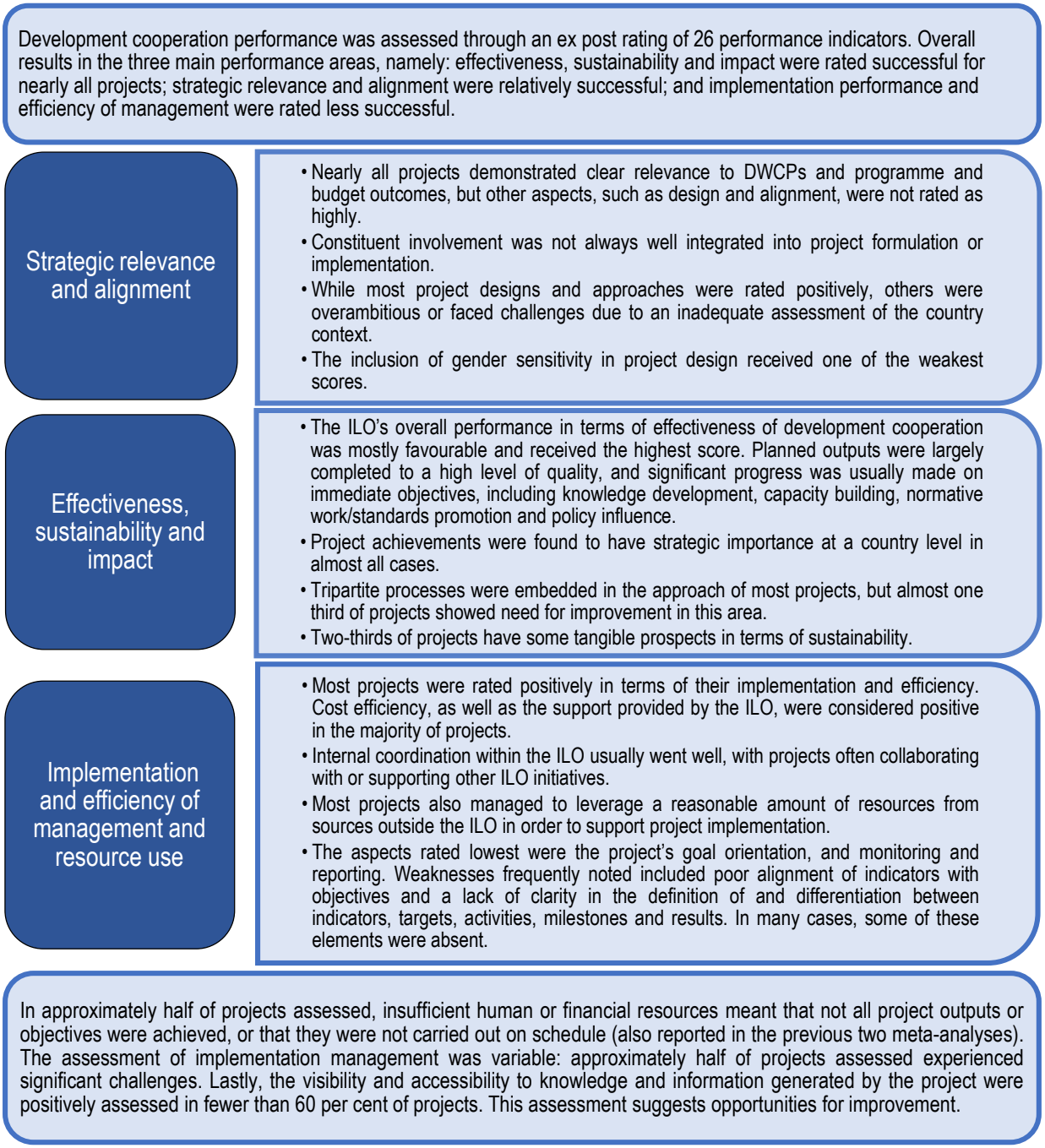
Enhance the sufficiency of evaluations to inform a higher order of effectiveness

81. Some challenges in using project evaluations for reporting on effectiveness relate to inconsistencies in the manner in which they address evaluation criteria and the ILO's mandates. This includes the need for a better and more consistent reflection of cost-efficiency and of the contributions made by projects to relevant programme and budget policy outcomes. Policy outcomes have already been mapped to the SDGs, and could be useful in identifying information to enable reporting on results for specific SDG targets and indicators.

82. Gaps can only be addressed by ensuring that evaluations follow standard and non-negotiable guidance regarding what they capture and present in the reports. It is particularly important to indicate which aspects evaluations should prioritize, such as an explicit analysis of tripartism, pro-poor coverage, inclusion and normative work. EVAL is currently conducting a study to have more practical guidance on evaluation models in order to better capture norms and social dialogue across a variety of projects (Part I, sub-outcome 2.4).

83. Figure 7 provides an overview of findings of the most recent study done on decent work results, covering the period 2013–16. The performance areas assessed include: (i) strategic relevance and alignment; (ii) effectiveness, sustainability and impact; and (iii) efficiency of management and resource use.

Figure 7. Snapshot of decent work results and effectiveness of ILO operations: A meta-analysis of development cooperation evaluations, 2013–16



Refining the meta-analysis approach and methodology to report on effectiveness

84. The current meta-analysis approach requires changes to what it measures and which basic reference documents are used. One emerging alternative is to be more selective in deciding the topics/themes so that the analysis is more focused (rather than analysing project evaluations across all themes). The other complementary option is to identify criteria that can be used to comment on organization-wide concerns and on decent work results. Currently, the methodology applied in the meta-analysis approach to measure project effectiveness uses 26 criteria under three broad performance areas. Smaller subsets can be

derived to comment on Organization-wide concerns and decent work results. When combined, these options are likely to give a more strategic and comprehensive overview of programme and budget outcomes, related SDGs and the cross-cutting policy drivers. As part of refining the approach, EVAL will explore the use of selected impact evaluations to demonstrate the ILO's development effectiveness. This would complement EVAL technical support to impact assessments in departments (Part I, sub-outcome 2.3).

Taking a more transformative approach to developing a more comprehensive validation process of ILO performance at the country and global levels

85. A more transformative approach is needed when evaluations are carried out in order to: better inform the ILO's understanding of its effectiveness in delivering on its policy outcomes; allow comprehensive coverage and analysis; and reduce "evaluation fatigue" among Office officials and constituents. The guiding principle of having a longer term programmatic approach to the results and impact of development cooperation, as suggested in the resolution concerning effective ILO development cooperation in support of the SDGs, adopted at the 107th Session (2018) of the International Labour Conference, also supposes that evaluations should take a longer term programmatic view of the ILO's work. This requires a transformative approach in terms of scope, coverage, management, resourcing, use of evaluations and evaluation culture in general.

86. The clustering of evaluations, recommended by the 2016 IEE and proposed in the new evaluation strategy, is a step in this direction. A significant advantage is that clustering projects into a single thematic evaluation (programme and budget policy outcomes) or geographical area (DWCP countries) will put the focus on the interconnectedness of the achievements of projects, the achievements of the organization and impact, including contributions to the Decent Work Agenda and the SDGs.

87. DWCPs and programme and budget policy outcomes are the two major anchors around which clustering can be planned. A cyclical approach to their evaluation will ensure that most programme and budget policy outcomes, DWCPs and selected global projects are covered. Thus, greater validation of reported results will be achieved compared to the relatively smaller coverage of the current approach to high-level evaluations prepared for and discussed in the Governing Body.

88. The clustering approach to evaluations requires closer collaboration with management to elicit donor consent to integrated evaluation processes. It also requires the re-conceptualization of evaluation resources management, and a review of the reporting lines of REOs to ensure the highest level of independence. Additional studies are under way to further assess these options and inform a more comprehensive validation process of ILO performance at the country and global levels.

Recommendation 3: Endorse the principle of working towards a more comprehensive validation process of ILO performance at the country and global levels in line with the concepts outlined above.

Draft decision

89. ***The Governing Body took note of the present report and endorsed the recommendations (paragraphs 17, 77 and 88) for implementation by the ILO, to be reported on in subsequent annual evaluation reports. It also confirmed the priorities identified in the report for the 2019–21 programme of work for evaluations.***