



September 2020

UPDATED COMMON EVALUATION FRAMEWORK FOR IMF CAPACITY DEVELOPMENT AND GUIDANCE NOTE

IMF staff regularly produces papers proposing new IMF policies, exploring options for reform, or reviewing existing IMF policies and operations. The Report prepared by IMF staff and completed on September 2, 2020, has been released.

The staff report was issued to the Executive Board for information. The report was prepared by IMF staff. The views expressed in this paper are those of the IMF staff and do not necessarily represent the views of the IMF's Executive Board.

The IMF's transparency policy allows for the deletion of market-sensitive information and premature disclosure of the authorities' policy intentions in published staff reports and other documents.

Electronic copies of IMF Policy Papers
are available to the public from
<http://www.imf.org/external/pp/ppindex.aspx>

International Monetary Fund
Washington, D.C.



September 2, 2020

UPDATED COMMON EVALUATION FRAMEWORK FOR IMF CAPACITY DEVELOPMENT AND GUIDANCE NOTE

EXECUTIVE SUMMARY

This document updates the Common Evaluation Framework (CEF) for the Fund's capacity development (CD) activities and provides practical guidance on its implementation. Since its adoption, the Fund has made progress in implementing the CEF. However, areas for improvement remain. The document aims to address these areas, drawing on lessons from experience with evaluations since the CEF's adoption.

The updates to the CEF aim at emphasizing learning from evaluations and simplifying the evaluation approach. The updates shift the primary purpose of evaluations toward learning and de-emphasize the goal of comparability of evaluations. This highlights the role that evaluations play in scrutinizing the results of CD, particularly those that feature noteworthy successes and/or failures, in order to draw valuable lessons on design and implementation of future CD. The updates also simplify the requirements for the evaluation process by de-emphasizing the requirement to use a counterfactual approach while basing the framework on (1) use of Results-Based Management (RBM) log frames; and (2) consistent application of the OECD Development Assistance Committee (DAC) criteria.

The document provides guidance on evaluation methods, approaches, and process to support the full implementation of the updated framework. The guidance covers the four phases of an evaluation: (1) high-level selection of evaluation topics for inclusion in the strategic work plan; (2) evaluation design, implementation, and reporting/follow-up; (3) dissemination and publication; and (4) addressing findings and recommendations.

Finally, the document provides guidance on applying the OECD-DAC criteria to the IMF context and gives examples of evaluation questions.

Approved By
Sharmini Coorey

Prepared by the Institute for Capacity Development (ICD) in consultation with other departments, under the guidance of Michaela Erbenova.

CONTENTS

Acronyms	3
INTRODUCTION	4
A. Background	4
B. Scope	6
C. Roadmap	7
UPDATES TO THE COMMON EVALUATION FRAMEWORK	8
PHASE I. HIGH-LEVEL SELECTION	10
PHASE II. EVALUATION	11
A. Design	11
B. Implementation	16
C. Reporting and Follow-up	17
PHASE III. DISSEMINATION AND PUBLICATION	20
PHASE IV. ADDRESSING FINDINGS AND RECOMMENDATIONS	21
BOXES	
1. Procedures for Ex-Post Evaluations of Exceptional Access Arrangements	6
2. Questionnaire and Interview Questions Best Practice	15
FIGURES	
1. The IMF's CD Evaluation Framework	7
2. The Results Chain for the IMF's CD Activities	14
TABLES	
1. Progress of Implementation of the Common Evaluation Framework	4
2. Updates to the Common Evaluation Framework	9
ANNEX	
I. The OECD-DAC Criteria and Example Questions	23

Acronyms

CCB	Committee on Capacity Building
CD	Capacity Development
CDD	Capacity Development Department
CDMAP	Capacity Development Management and Administration Program
CEF	Common Evaluation Framework
DAC	Development Assistance Committee
EPE	Ex-Post Evaluation
ESC	Evaluation Subcommittee
ICD	Institute for Capacity Development
ITP	Institute Training Program
RBM	Results-Based Management
RCDC	Regional Capacity Development Center
SC	Steering Committee
SPR	Strategy, Policy, & Review Department
TA	Technical Assistance
ToR	Terms of Reference

INTRODUCTION

A. Background

1. **Since its adoption, the Fund has made progress in implementing the CEF for CD.** The CEF was intended to streamline practices and increase comparability and use of results to foster learning and enhance accountability in CD.¹ Among the four objectives that describe the elements of the CEF (Table 1), significant progress was made on improving the information supporting evaluations, particularly on the application of the RBM framework and the adoption of pre- and post-course tests for all ICD courses delivered under the IMF Institute Training Program (ITP). Steps have also been taken to implement other elements, including introducing a CD evaluation work plan with a rolling three-year horizon (to use evaluation resources more efficiently), and applying the OECD-DAC criteria in most internal evaluations as well as all donor-mandated evaluations, as stipulated in a terms of reference (ToR) (to support shorter, and more focused evaluations).

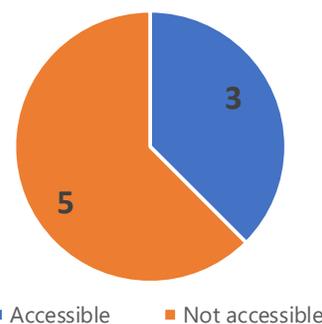
Table 1. Progress of Implementation of the Common Evaluation Framework	
CEF Elements	Implementation
Support shorter, more focused, and more comparable evaluations.	
Specify clearly the objectives of all CD activities to be evaluated.	In progress. All donor-mandated evaluations and some departmental evaluations use the objectives in the RBM log frames.
Apply a four-step common framework (thought process) to all evaluations.	In progress. All donor-mandated evaluations and some departmental evaluations use the RBM framework and apply the DAC criteria. Very few have employed a counterfactual.
Ensure clarity and consistency in definition and application of internationally accepted OECD-DAC evaluation criteria.	In progress. All donor-mandated evaluations and some departmental evaluations apply the DAC criteria.
Use standardized ToR templates to reduce workload and sharpen the focus of evaluations.	In progress. Standard ToRs are used for donor-mandated evaluations.
Improve the information supporting evaluations.	
Consistent with the new results focus of all IMF CD, ensure that all CD activities have a data collection plan in place before the start of the CD activity.	Implemented. Most CD activities have an RBM log frame.
Obtain ex-post information from technical assistance (TA) providers and users and other stakeholders.	In progress. End-of-project self-assessments are done for all IMF02-funded projects.
Extend the use of pre- and post-course tests to all IMF training courses.	In progress. Pre- and post-course tests are used for all ICD courses, delivered under the ITP.
Revise the end-of-course and follow-up survey questions for training to be more results-oriented.	Implemented. Surveys for training have been revised.
Maintain a channel for regular reviews of training course content.	In progress. Some CD departments (CDD) have ensured channels for regular review (e.g., ICD's topic networks) to keep courses up to date.

¹ [International Monetary Fund, *New Common Evaluation Framework for IMF Capacity Development*, April 2017.](#)

Table 1. Progress of Implementation of the Common Evaluation Framework (concluded)	
CEF Elements	Implementation
Seek information from IMF country teams on the extent to which they observe tools learned in IMF training being applied at the country level.	Partially implemented. The questionnaire for country teams was developed but not rolled out.
Maintain the current level of resources allocated to evaluation, but use these scarce evaluation resources more efficiently.	
Establish reasonable accountability standards without evaluating everything.	In progress. The medium-term evaluation work plan covers a representative sample of the Fund CD but relies mostly on external evaluators with varying degrees of familiarity with Fund outputs and CD processes.
Introduce a CD evaluation work plan with a rolling three-year horizon, revised annually.	Implemented.
Decide what and how to evaluate based on potential value of the information, cost, and achievement of accountability standards.	In progress.
Use the information from evaluations to alter practices and/or shift resources.	
Introduce quantitative scoring to complement qualitative information in evaluations to facilitate comparisons and aggregation.	In progress. Use of quantitative scoring has been mixed.
Ensure evaluations are easily accessible upon completion and findings are broadly shared among Fund staff.	In progress. Some departmental evaluations are strictly confidential/confidential.
Incorporate evaluation results when setting CD priorities at the November meetings of the Committee on Capacity Building (CCB). ¹	In progress. Evaluation results are discussed in the summer CCB, allowing CCB members to consider them in the fall.
¹ The CCB is a high-level committee with a mandate to organize and oversee the Fund's policy work in capacity-building and to implement the Managing Director's strategic directions.	

2. Despite progress, areas for improvement remain. Evaluation reports, particularly the donor-mandated ones, remain long and, although quality varies, not always focused on the achievement of CD objectives. At times, too many entity-level governance and operational issues in regional capacity development center (RCDC) evaluations divert resources and focus away from the evaluation of project performance. Thematic trust fund (TTF) evaluations tend to have more findings and recommendations on CD content and delivery. Internal evaluations, i.e., those funded by IMF01 funds, are often led by staff in the same department whose CD is the focus of the evaluation, which could potentially lead to conflicts of interest, unlike the practice in ex-post evaluations of lending programs (Box 1). The lack of thematic CD expertise among Fund staff outside of the CDD that delivers specific CD topics is a factor behind this practice. While the three-year strategic evaluation workplan is presented in the CCB, active discussion on the topics that yield the most value could be fostered further, to ensure better use of

Staff Access to Internal Evaluations
(All IMF01-funded evaluations, FY14-19)



Source: IMF staff estimates.

scarce evaluation resources. Finally, reports, especially from internal evaluations, could be disseminated more widely to staff, partners, and the Executive Board, and used more purposely in CD prioritization.

3. This document aims to address these areas for improvement. It updates the CEF and provides guidance on the evaluation process and approach. Drawing on lessons from experience since its adoption, the updates modify some elements of the CEF, emphasizing the learning objective of evaluations (over accountability) and simplifying the approach to enhance focus and usability. The document also provides guidance on evaluation methods, approach, and process to support full implementation of the updated framework.

Box 1. Procedures for Ex-Post Evaluations of Exceptional Access Arrangements

Team composition. An ex-post evaluation (EPE) is undertaken by an interdepartmental team led by a mission chief from a department other than the home area department. The team includes representatives from the home area department and one each from the Strategy, Policy, & Review Department (SPR) and at least one other functional department.

Timeline. An EPE is completed within one year of the end of the arrangement.

Internal review. The draft report is reviewed by departments. The report is sent to management prior to or concurrently with the brief for the mission during which the EPE will be discussed. The cover note conveying the report to management is signed by the head of the area department and the SPR review officer. It sets the main conclusions and any dissenting views.

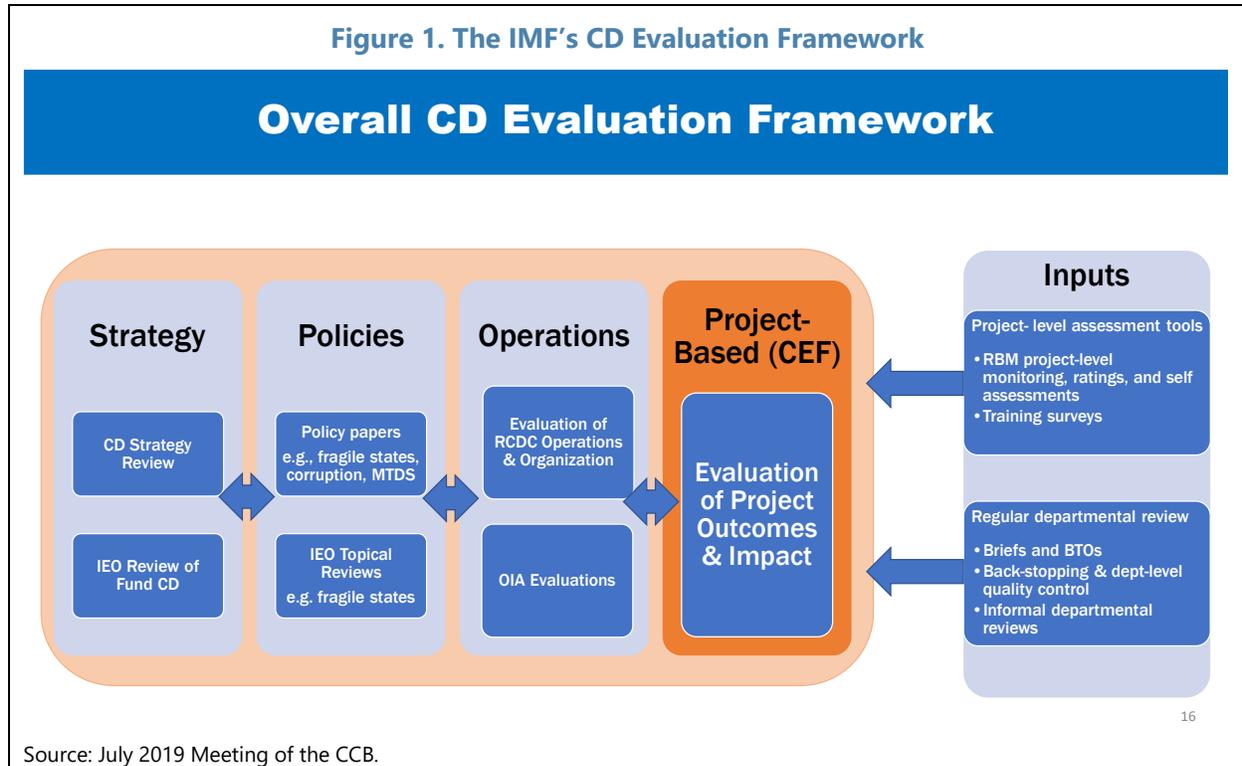
Dissemination and publication. A Board discussion is envisaged, preferably combined with Article IV consultation or post-program monitoring discussions. The publication of the report is voluntary but presumed, with the member's consent obtained on a non-objection basis.

B. Scope

4. The CEF applies to all project-based evaluations that are included in the three-year strategic work plan (Figure 1). Project-based evaluations aim at assessing progress toward the achievement of the objectives of CD delivery. The CEF does not apply to other types of evaluations regularly conducted by the Fund on CD strategy, policy, and operations. The CEF does not also apply to progress assessments made during and at the end of project implementation (e.g., RBM ratings, training surveys) and other departmental reviews of CD projects (e.g., internal papers on CD activities, briefing papers, back-to-office reports, backstopping). The latter are important inputs to formal evaluations, including project-based evaluations. Finally, the CEF—although not required—is the preferred approach for evaluations that are not endorsed by the CCB as part of the three-year strategic workplan.

5. A CD project is the unit of analysis for all evaluations under the CEF. As defined in the RBM Governance Framework, a country CD project is designed to achieve one or more objectives (and related outcomes) in a single country (or regional body) over a specified timeframe through a series of CD activities. A multi-country CD project is CD delivered to counterparts from multiple countries designed to achieve one or more objectives (and related outcomes). An evaluation would

typically cover a set of CD projects, that may be grouped by funding vehicle (e.g., a thematic fund, a regional technical assistance center (RTAC), a bilateral subaccount), country group/region/topic (e.g., tax policy CD in AFR), or other parameters.



C. Roadmap

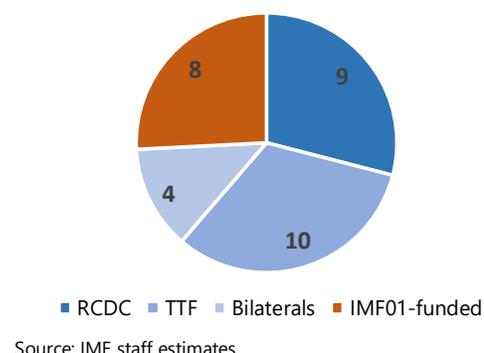
6. The remainder of the document is divided into two sections. The first section proposes updates to the CEF, based on experience with implementation since its adoption. The second section presents practical guidance on implementation and is organized by the key phases of a project-based evaluation. *Phase 1* focuses on the high-level selection process for what to evaluate and why. *Phase 2* covers the three stages in conducting an evaluation: design, implementation, and reporting and follow up. *Phase 3* provides guidance on the dissemination policies for evaluation-related documents. Finally, *Phase 4* covers how to distill, share, and follow up on evaluation findings, lessons learned, and recommendations across evaluations for maximum learning, use, and accountability.

UPDATES TO THE COMMON EVALUATION FRAMEWORK

This section covers broad updates to the CEF based on challenges encountered during implementation.

7. Experience from implementing the CEF brought to light two challenges. First, the CEF sought to produce comparable evaluations that would ensure some reasonable standard of accountability. However, in practice, most evaluations are donor-mandated and implemented by different external teams, with varying degrees of expertise on CD topics, as well as familiarity with Fund CD processes, leading to incomparable assessments. Fund-staff led evaluations, which would be more amenable to comparability, average only 1-2 every year. Second, the CEF specified a four-step process as an evaluation approach, that could be rigid, particularly the emphasis on identifying and assessing the counterfactual (or analysis of results if Fund CD was not provided). While good counterfactuals could be very helpful in assessing how much of the results can be attributed to Fund CD, they are not often available. As a result, this approach was very rarely used.

Reliance on Donor-Mandated Evaluations
(All evaluations, FY14-19)



8. In response to the first challenge, this update proposes to shift the primary purpose of evaluations toward learning. RBM log frames have close to full coverage of CD projects, a well-defined structure, that is also internally managed. The role of evaluations is to dig deep on what drives certain RBM results, particularly those that feature noteworthy successes and/or failures. Evaluation topics and scope should therefore be chosen, and the evaluation designed, mainly to draw lessons from specific CD projects for future design of country engagements. Evaluations would continue to contribute to the purpose of accountability on the set of projects that it covers.

9. In response to the second challenge, this update also proposes to simplify the four-step evaluation process, particularly de-emphasizing the counterfactual. When good counterfactuals are available, they should be used in evaluations. When relevant (such as presence of several CD providers on the ground, as well as obvious external factors that could have led to the outcome), the evaluator should make an effort to assess the extent to which the CD project contributed to achieving the objectives, but a counterfactual is not required. This change would recast the four-step process² in the CEF into two requirements: (1) use of RBM log frames as a basis for an evaluation; and (2) consistent application of the OECD-DAC criteria.

² The four-step process are (i) define the log frame or causal chain from inputs to outcome; (ii) to the extent possible, indicate what is likely to have happened if the IMF did not delivery the CD (e.g., assess the counterfactual) to help assess impact; (iii) assess outcomes using the OECD-DAC evaluation criteria; and (iv) discuss why achievement of the

10. Other elements of the CEF remain unchanged, except for updates to align the CEF with the new RBM Governance Framework (Table 2). The full implementation of these other elements will be supported by practical guidance on the evaluation process, as well as the approach and methodology.

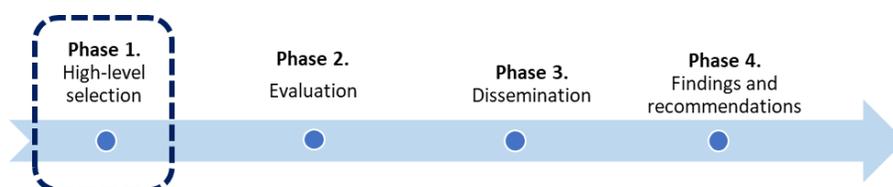
Table 2. Updates to the Common Evaluation Framework	
CEF Elements	Update
Support shorter, more focused, and more comparable evaluations.	Support shorter and more focused evaluations, aimed at learning.
Specify clearly the objectives of all CD activities to be evaluated.	Unchanged. Guidance provided in this note, requiring evaluations to use RBM log frames (Phase 2).
Apply a four-step common framework (thought process) to all evaluations.	Apply the following principles for all evaluations in the strategic workplan: (1) use of RBM log frames as a basis for the evaluation; and (2) use of the OECD-DAC criteria to frame the evaluations. Guidance provided in this note (Phase 2, Annex).
Ensure clarity and consistency in definition and application of internationally accepted OECD-DAC evaluation criteria.	
Use standardized ToR templates to reduce workload and sharpen the focus of evaluations.	Use a ToR to sharpen the focus of evaluation. Guidance provided in this note (Phase 2). ToR template to be provided, based on the guidance.
Improve the information supporting evaluations.	Unchanged.
Consistent with the new results focus of all IMF CD, ensure that all CD activities have a data collection plan in place before the start of the CD activity.	Unchanged. Guidance provided in the RBM Governance Framework.
Obtain ex-post information from TA providers and users and other stakeholders.	Unchanged. Guidance provided in the RBM Governance Framework.
Extend the use of pre- and post-course tests to all IMF training courses.	Extend the use of pre- and post-course tests to all ITP courses.
Revise the end-of-course and follow-up survey questions for training to be more results-oriented.	Use results-oriented follow-up survey questions for ITP courses.
Maintain a channel for regular reviews of training course content.	Unchanged. Departments are responsible for updating training course materials.
Seek information from IMF country teams on the extent to which they observe tools learned in IMF training being applied at the country level.	Dropped. Information will be gathered through tools such as CD Country Strategy Notes and Regional Strategy Notes, as well as Capacity Development Management and Administration Program (CDMAP).
Maintain the current level of resources allocated to evaluation, but use these scarce evaluation resources more efficiently.	Unchanged.
Establish reasonable accountability standards without evaluating everything.	Maximize learning and support accountability
Introduce a CD evaluation work plan with a rolling three-year horizon, revised annually.	Unchanged. Guidance provided on the process for establishing the work plan (Phase 1).
Decide what and how to evaluate based on potential value of the information, cost, and achievement of accountability standards.	Decide what and how to evaluate based on the potential value of the information, cost, and learning gains. Guidance provided in this note (Phase 1).
Use the information from evaluations to alter practices and/or shift resources.	Unchanged.

DAC criteria was low/high, what factors explain it, and whether alternative intervention might have provided better results.

CEF Elements	Update
Introduce quantitative scoring to complement qualitative information in evaluations to facilitate comparisons and aggregation.	Introduce project-level scoring to ensure that lessons are identified from specific interventions . Guidance provided in this note (Phase 2).
Ensure evaluations are easily accessible upon completion and findings are broadly shared among Fund staff.	Unchanged. Guidance provided in this note on the dissemination and publication of evaluations, consistent with the TA dissemination policy (Phase 3).
Incorporate evaluation results when setting CD priorities at the November meetings of the CCB.	Unchanged. Guidance provided in this note on the use of relevant summer CCB output for decision-making, particularly on priority-setting in the fall CCB (Phase 4).

PHASE I. HIGH-LEVEL SELECTION

This section covers the process for the high-level selection of evaluation topics for inclusion in the strategic work plan. The topics proposed by departments and endorsed by the CCB, along with donor-mandated evaluations, will form the three-year strategic workplan. The CEF applies to all evaluations in the workplan.



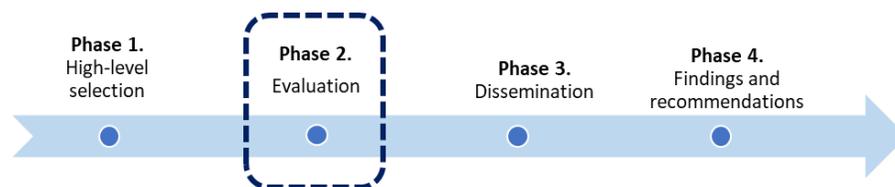
11. Process. All departments are free to propose evaluation topics. To be included in the strategic work plan, evaluation topics should be proposed to the CCB for consideration at its summer meeting. These topics could cover a CD workstream provided in a region/country grouping by a single department (e.g., banking supervision CD by Monetary and Capital Markets Department (MCM) in Asia) or by several departments (e.g., governance CD by Fiscal Affairs (FAD) and Legal (LEG) departments). In the proposal, the proposing department(s) will describe the high-level purpose/goals, scope, coverage, and the start and completion dates of the evaluations. The goals will depend on the context and how the findings will be used (e.g., improve the performance of future CD interventions, provide input to the next phase of a project, mitigate risks, or ensure accountability to different stakeholders on results). In its summer meeting, the CCB will discuss departmental proposals, and accounting for key considerations, endorse the evaluation topics that will start in the current fiscal year and two years after. Finally, should there be an urgent need, departments could propose new evaluations as part of the strategic workplan, and the CCB could endorse outside the summer meeting. Once a proposal has been accepted, the department(s) who proposed the topic and who will also typically manage the evaluation, will include it in their accountability frameworks and set aside budget resources.³

³ Donor-mandated evaluations have specific earmarked budget allocations for an independent evaluation from the trust fund resources provided by donors (i.e., IMF02 resources). Internal evaluations are financed from the Fund’s own resources (i.e., IMF01 resources) and have no earmarked budgetary allocation. Rather departments sponsoring/conducting the evaluation allocate financing from their existing budget resources.

12. Key considerations. The CCB will consider evaluation topics that are relevant to learning and operational/strategic CD decision making. These topics could emerge from other evaluations (e.g., previous project-based evaluations, Independent Evaluation Office evaluations, policy papers on surveillance/program topics) and from the analysis of monitoring data, such as the Fund’s RBM framework and project self-assessments. RBM data could specifically point to topics with high potential for learning, such as evaluations of highly successful or unsuccessful initiatives or a quick mid-term or formative evaluation of a project/program to provide feedback for course correction. Choosing such projects with high potential for learning could introduce a deliberate selection bias, for instance, by focusing only on highly unsuccessful (or highly successful) projects. In such cases, the proposal should highlight upfront the key goal of the evaluation as learning, along with the direction of deliberate selection bias, and deemphasize the accountability purpose accordingly. Furthermore, risk-based considerations may lead to evaluating projects that have a critical impact on the Fund’s broader work program (e.g., on surveillance and lending) or reputation.

PHASE II. EVALUATION

This section covers the process of conducting an evaluation. This includes three stages of designing, implementing, and reporting on evaluations.



A. Design

This stage covers the design of an evaluation, including key processes and elements.

Specifying the Broad Elements of an Evaluation

13. The department(s) managing the evaluation, or the managing department,⁴ will assign an evaluation manager, who will oversee the evaluation process but will not take part in the evaluation. For donor-mandated evaluations, the Evaluation Subcommittee (ESC),⁵ who will agree by consensus on the evaluation design, is the evaluation manager. To ensure impartiality, the evaluation manager should not be involved in the delivery, backstopping, and/or oversight of any of the CD projects covered by the evaluation. The evaluation manager should decide on the elements listed below.

⁴ The managing department(s), i.e., the departments managing the evaluation, will typically be the CD and/or area department who proposed the evaluation to the CCB and whose CD provision is the focus of the evaluation.

⁵ The ESC is typically comprised of development partners, member countries, and Fund staff. The AD/thematic fund coordinator acts as the lead representative of the ESC and ICD acts as the secretariat of the ESC.

- **Purpose, scope, coverage, and aggregation.** The evaluation manager should identify the purpose/goals of the evaluation, complete set of projects that will be evaluated, and if the set is large, whether a sample should be taken, to make the evaluation possible. In addition, the evaluation manager should decide on the level of aggregation (e.g., sub-topics, country/country groups) of the evaluation ratings and assessment. Finally, in the case of evaluations of projects under funding vehicles, the evaluation manager should indicate if the evaluation is also expected to cover some broad operational issues (e.g., efficiency of a regional CD center or RCDC), and specify these issues separately.
- **Timeframe of the evaluation.** The evaluation manager will set the timeline for the main deliverables of the evaluation, which includes the inception note, the draft evaluation report, the revised evaluation report, a presentation of findings and recommendations, and the final evaluation report.
- **Attributes of the evaluation team.** The evaluation team should have the following attributes: (i) no involvement in the delivery, backstopping, and/or oversight of any of the CD projects covered by the evaluation; (ii) thematic and regional expertise; (iii) experience in macroeconomic policy analysis and institutional knowledge on the IMF operations; and (iv) evaluation experience and understanding of the CEF and the RBM Governance Framework. The team leader should have an understanding of the CEF and its application, as reflected in this guidance note, and the OECD-DAC criteria,⁶ though prior experience in evaluation is not required. The evaluation manager could specify additional skills required. For internal evaluations, the ToR will identify the evaluation team, based on the required attributes. To avoid conflicts of interest, internal evaluations should be undertaken by an interdepartmental team led by a mission chief that is not from the CDD or AD whose projects account for a significant part of the evaluation coverage.⁷ The team members should include representatives from the CD and area departments whose CD is the focus of the evaluation. The managing department will nominate the team leader (i.e., the mission chief), as well as the departments from which the team members will come from. These departments will in turn identify the choice of representative. The evaluation manager will not be part of the team.
- **Evaluation inputs.** Evaluations typically use the following inputs: (i) output from monitoring and self-assessment tools, such as RBM information, interim and end-of-project assessments; (ii) output of CD activities, such as CD mission briefs, back-to-office reports, TA reports, and presentations;⁸ (iii) country reports, such as policy notes, Article IV reports, program documents, and Financial Sector Assessment Program (FSAP) reports; (iv) strategy and planning CD documents, such as CD country strategy notes, regional strategy notes, and reports and steering

⁶ OECD-DAC Network on Development Evaluation, [Better Criteria for Better Evaluation Revised Evaluation Criteria Definitions and Principles for Use](#), December 2019.

⁷ The guidance provided here on the attributes of the team leader and the composition of the team will be piloted on two to three internal evaluations and revisited based on lessons learned from these pilots.

⁸ Back-to-office reports are not shared with external evaluators, as they are considered internal documents.

committee meeting minutes of the funding vehicle; and (v) CD budget and spending data. For training projects, inputs also include results from end-of-course surveys, the pre-and post-course tests, and follow-up surveys that are customized to the learning objectives of the course. The team is also encouraged to collect and analyze external information for a comprehensive view of factors that led to the success or failure of reforms.

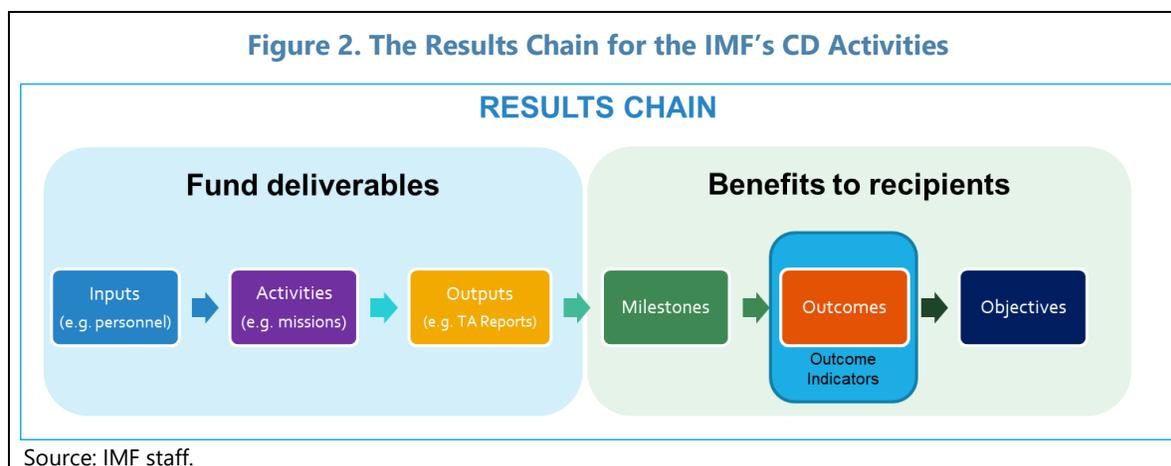
14. The evaluation manager will draft a ToR to convey the design elements to the evaluation team. The ToR will be reviewed by relevant departments, including CD and area departments that are not managing departments, but are involved in the projects being evaluated, as well as ICD, who will review from the angle of ensuring the evaluation is consistent with the CEF guidelines. For donor-mandated evaluations, this may include members of the ESC. In addition, the ToR will specify that the evaluation will (a) use the logical framework from RBM, that lays out the sequence of steps on how CD actions are expected to lead to outcomes and objectives (¶15), and (b) assess the achievement of objectives using the relevant OECD-DAC criteria (¶16).⁹ The ToR will be cleared by a deputy director of the managing department for internal evaluations while the ESC will clear for donor-mandated evaluations.

Using the RBM Logical Framework

15. All evaluations should base their assessment on the Fund's RBM log frames (Figure 2). Most Fund CD projects have a logical framework (in short, log frames),¹⁰ established at the beginning of the project. The starting point for a log frame is the *objective*, i.e., the outcomes and results the planned CD aims to achieve/contribute to. The causal chain for meeting the objectives describes how *inputs* (e.g., financial and human resources) are translated into *activities* (e.g., missions, backstopping, delivering a training course) to produce the *outputs* of the Fund's work (e.g., a technical assistance report). It then describes what *outcomes* (i.e., the actual capacity improvements) are expected along with any *milestones* (interim steps) that will be completed en route to meet the *outcomes and objectives* of a specific CD activity. The elements of a log frame, which include objectives, outcomes, indicators (baseline and target), milestones, and risks assessments, are key to understanding the project. The evaluation should assess projects according to the realization of outcomes and objectives, specified at the beginning of a project, as captured by the indicators, milestone achievement (including ratings), and risk assessments. In particular, the assessment under the effectiveness criteria (¶16) should take into account the outcome ratings, that are based on the status of milestones as well as indicators relative to target and baseline.

⁹ Compared to that of an externally delivered evaluation, the ToR for an evaluation team led by Fund staff is expected to be more concise.

¹⁰ While the Fund's CD projects are increasingly linked to logical frameworks, full coverage has not been accomplished. The share of IMF01 CD activities covered by a log frame is lower than that of IMF02. The difference is explained by the typically short-term, one-off, and urgent nature of IMF01 projects.



Applying the Relevant OECD-DAC Criteria

16. All evaluations should be assessed using the OECD-DAC criteria. The OECD-DAC criteria cover a set of six dimensions against which to assess development interventions: relevance, coherence, effectiveness, impact, efficiency, and sustainability. It is not required that an evaluation covers *all* criteria. The evaluation manager should discuss with the evaluation team which DAC criteria will be relevant to achieving the objectives of the evaluation and the specific evaluation questions that the team will aim to answer. For an evaluation focused on learning if CD has supported a concurrent Fund-supported arrangement, the efficiency and sustainability criteria may not be relevant. For a set of ongoing projects whose impact may be too early to assess, the impact criteria could be dropped. The annex presents definitions of the criteria, guidelines on applying to the Fund context, and indicative evaluation questions that may be used for each OECD-DAC criterion.

Specifying the Details of an Evaluation

17. Once ToR have been cleared, the evaluation team will work on the details of the evaluation design. In particular, the team will perform a desk review of the evaluation inputs available to them and, guided by the ToR, proceed to:

- **Analyze the RBM log frames and reconstruct/adjust them, if needed.** The evaluation team will use the RBM log frames to understand project goals and design. If the project is either missing a log frame or has a weak or incomplete log frame, the evaluation team should reconstruct what the project was aiming to achieve and how. Such cases are expected to become fewer, after the introduction of CDMAP and the RBM Governance Framework. In such cases, the team will need to discuss with the evaluation manager to determine the objective, outcomes, and indicators (baseline and target) of the project, using information in the relevant CD reports at the design stage (e.g., mission brief or TA report) and/or interviews with staff who supervised the CD activity. The evaluation team is expected to use the RBM catalog when developing ex-post the objectives, outcomes, and indicators of the evaluation. In some cases, the objectives, outcomes, and indicator targets of the CD may have been modified during

project implementation.¹¹ In such cases, the team should use the new elements for the evaluation and make a note in the evaluation report. The team will not update the log frame in CDMAP, which is the responsibility of the project manager.

- **Propose an evaluation subsample.** If the evaluation manager decides to sample from the full set of projects, the team will propose the selection criteria for the sample. The selection should aim to be representative of the full set, unless justified by other considerations such as the objectives or cost of the evaluation.
- **Propose a specific set of evaluation questions,** for each of the relevant criteria, guided by the purpose of the evaluation. The questions should balance comprehensiveness to attain the purpose of the evaluation and focus, to ensure that the evaluation remains useful. The annex provides a list of indicative questions for each of the criteria.
- **Propose a methodology** (quantitative or qualitative), and *develop the instruments*, such as questionnaires, interviews, data request templates, and case studies, to respond to the evaluation questions (Box 2). An online survey could also be developed to collect information on the full sample of projects to triangulate the information collected from interviews and case studies.¹² The team should specify the respondents to relevant instruments. Typically, based on the methodology and instruments, the team will propose to visit recipient countries.

Box 2. Questionnaire and Interview Questions Best Practice

- The questions should be formulated in a simple and clear manner and avoid the technical evaluation terminology. The evaluation team should not presume familiarity with the DAC criteria and provide short definitions of each criterion before asking a set of questions specific to that criterion.
- The questions should be customized to different stakeholders, which typically include CD recipients, CD providers (including back stoppers), CD managers, area department teams, and, if relevant, RCDC staff, development partners, and steering committee (SC) members. For the first two, the questions should be project-specific and refer to project outcomes/objective. High-level or non-project related questions should be targeted separately to relevant stakeholders.
- The online survey should limit the number of open-ended questions and use close-ended questions with multiple choices while giving the “other” option to collect any open-ended comments. Too many open-ended questions in an online survey tend to reduce the response rate. Interview questions could be open-ended.

18. The evaluation team will draft an inception note to describe the evaluation design, building on the ToR, review of available evaluation inputs, and discussions with the evaluation manager. Like the ToR, inception notes will be reviewed by relevant departments, including ICD, and cleared by the evaluation manager. The final note should be made immediately available to Fund

¹¹ Under the RBM Governance Framework, objectives and outcomes could be modified only with the approval of the portfolio manager. Indicators could be modified by the project manager, with documented justification. These rules would not have applied to projects initiated before the new framework.

¹² An online survey is especially encouraged in evaluations covering a relatively large number of countries and when only a small portion of them could be included in case studies and interviews.

staff, especially to interview/survey respondents. The inception note should include the following elements:

- **Purpose, scope, and coverage**, as laid out in the ToR.
- **Evaluation subsample.** The team will discuss the selection criteria and present summary statistics relative to the full set of projects.
- **Aggregation level.** The main report will typically refrain from presenting project- or country-level assessments. Such granular assessments will be aggregated/summarized to the level that is most useful for the evaluation purpose. The inception note will describe the level of aggregation of project assessments (such as by topics, county groups, and modality). For example, if extracting lessons on modes of delivery is an evaluation objective, aggregating by delivery modality would be specified in the inception note.
- **Relevant DAC criteria and key evaluation questions.** The note will list the relevant DAC criteria that will be used, as defined by the ToR, and the key evaluation questions under each criterion. In the case of evaluations of projects under funding vehicles, if the evaluation is also expected to cover some broad operational issues, the note should specify these non-project-specific evaluation questions separately.
- **Methodology.** The inception note will present the approach (quantitative or qualitative), related instruments that the evaluation team will use to respond to the key evaluation questions, and the respondents for the questionnaires and interviews. If relevant, the team should also discuss the proposed case studies and any field work they intend to do to support the approach.
- **Annexes.** The note will include the following in annexes: (a) draft instruments such as questionnaires, interview questions, and data request templates; and (b) list of the complete set of projects covered by the evaluation. Other annexes can be added as appropriate.

B. Implementation

This stage covers the field phase of the evaluation and analysis of the key evaluation inputs.

19. Field phase. During the implementation phase, the evaluation team will execute the methodology laid out in the inception note, typically involving one or more in-person or virtual/remote visits to recipient countries that are identified from the evaluation sample. Here are some best practices during these visits:

- **Hold an opening and closing meeting with the recipient agency’s management.** Opening meeting will be used to explain the objectives of the evaluation, a closing meeting to discuss preliminary findings.
- **Hold a briefing meeting with the IMF resident representative/RCDC head** at the beginning and end of the visit.

- **Conduct interviews with stakeholders**, which typically include CD recipients, CD providers (including backstoppers), CD managers, area department teams, and, if relevant, RCDC staff, development partners, and SC members. Use a predetermined set of interview questions, differentiated by stakeholders.
- **Keep the anonymity of interviewees** to encourage more candid views and eliminate any risks of exposure to retribution.
- **Be open** to what data or new information collected in the field reveals, to ensure the credibility of the evaluation.
- **Form a preliminary assessment of key conclusions**, which will be the basis for discussions in the closing meetings, and for internally-led evaluations, preparing a back-to-office report.

20. Analysis. The evaluation team will analyze the available information, including those collected during the field phase, to respond to the evaluation questions. In particular, the team will:

- **Rate each project and relevant DAC criteria.** For *each* project, the team will summarize their response to the evaluation questions by rating *each* relevant DAC criteria on a 1–4 scale.¹³ For the effectiveness criteria, the team will use the RBM ratings as one input but will take into account other sources to cross-validate and assess the reliability of the RBM ratings. Whenever the requisite evidence needed to rate a criterion is not available, the evaluation report should convey this finding.
- **Explain the rating.** For *each* project, the team will provide a narrative of the extent to which (and how) the objectives and related outcomes were achieved and identify factors behind the rating. The team will also raise any exogenous factors that may have contributed to the outcome, as well as highlight alternative interventions, if any, that could have provided better results.
- **Aggregate the ratings.** Finally, the team will decide on the assignment of weights in aggregating the ratings. The project-level assessments in the previous two bullets will be reflected in the annex of the main evaluation report. The main report will reflect a weighted average/summary of the project-level ratings and explanations.

C. Reporting and Follow-up

This stage describes the final report as well as the review process and follow-up in response to the report.

21. Main report. The report should not exceed 25 pages in length (excluding annexes), including the executive summary. It is expected to include the following sections:

¹³ The rating will correspond to the following performance assessments against each criterion: 4—Fully (or substantially fully) met; 3—Mostly met; 2—Partially met; and 1—Not met.

- **Executive summary.** After a short paragraph on the context for the evaluation, the executive summary will concisely focus on the main evaluation findings and evaluation recommendations.
- **Introduction.** The introduction will briefly present the purpose and the scope of the evaluation.
- **Project-based evaluation.** This section should focus on presenting an aggregated assessment of projects covered in the evaluation, based on the bottom-up project-by-project assessment, which will be presented separately in an annex. The project evaluation section will cover the following:
 - *Scope.* The total number and scope of projects covered in the evaluation and descriptive statistics on the evaluation sample will be presented.
 - *Assessment and analysis at an aggregated level (i.e., topic, country groups, modality) using the OECD-DAC criteria.* The evaluation team will present the assessment according to the aggregation approach that is most useful for the purpose of the evaluation and laid out in the inception note. The team could also present its assessment by the relevant DAC criteria, if deemed useful.¹⁴

In addition to the assessment, this section will present the factors that affected the assessment, including obstacles to the achievement of objectives, such as constraints facing recipient countries, and alternative interventions that would have provided better results. The evaluation team is expected to highlight common factors affecting all projects as well as any significant differences across aggregated groups (e.g., specific patterns observed by CD topics, country groups, modality).

To facilitate dissemination and publication of the main report, the evaluation team will avoid presenting project- or country-specific information and/or final TA advice, where feasible (see Phase III on dissemination).

- *Assessment of RBM log frames.* The evaluation team can also present an overall assessment of the quality of log frames of the projects (e.g., whether the projects have clearly defined objectives and log frames marking the results chain from input, activities, output, and milestones to outcomes and objectives with well-defined verifiable indicators).
- **Non-project-specific evaluation.** For RCDC evaluations, the evaluation manager and evaluation team may agree to assess the entity's operations with some entity-level questions, in addition to the project-based evaluation. In this section, the evaluation team will present their

¹⁴ For example, if the evaluation topic is Public Financial Management (PFM) in the African region and it was deemed useful during the inception phase to gather lessons on the relevance and effectiveness of low-income and frontier/emerging economies in the region, the section could have sections on Low-income Economies and Frontier/Emerging Economies and subsections on Relevance and Effectiveness for each section. Alternatively, the team could drop the subsections if they are considered overly granular.

assessment for such non-project related questions (as opposed to project-based questions). The OECD-DAC criteria do not have to apply to this part of the evaluation.

- **Conclusions and evaluation recommendations.** This section lays out the conclusions, particularly lessons learned from the evaluation, and recommendations. The report should contain no more than 10 recommendations and they should be:
 - *Prioritized*, in terms of urgency and timing, and sequenced;
 - *Actionable* (under the control of the IMF), feasible, and reflecting an understanding of potential constraints to implementation; and
 - *Cost effective* (i.e., focused on affordable alternatives to achieve the objectives).

22. Annexes. The evaluation report will have the following annexes:

- **Evaluation of individual projects.** This annex will describe each CD project succinctly, listing major activities and the channels through which they were expected to achieve outcomes and objectives. In a table, the evaluation team should present the DAC criteria rating by project and in aggregate, as well as the RBM rating, where available.¹⁵ For multi-country evaluations, the annex is also expected to present a brief assessment of performance by country, highlighting differences in project performance across countries.
- **Methodology.** The evaluation methodology as conveyed in the inception note will be presented. This annex will also include the final interview and survey questions along with the results of the survey.
- **Other annexes.** The list of missions and interviews will be presented. If deemed necessary, the evaluation team can add other information in annexes.

23. Review. All evaluation reports will go through an interdepartmental review process, that includes all CDDs and area departments that either delivered or received CD projects included in the scope of the evaluation, ICD who will ensure consistency with these guidelines, as well as the ESC for donor-mandated evaluations. Once a report has been delivered, the evaluation team will circulate it for review to relevant departments. In the case of donor-mandated evaluations, ICD will circulate the draft report to departments and the ESC. The report will also be reviewed by the country authorities of projects that were selected in the evaluation sample for factual checks. Its analysis and conclusions should not be subject to negotiation with the authorities. Final decisions on the content of the evaluation report are the responsibility of the evaluation team leader. Internal evaluation reports will be cleared by the managing department, typically the deputy director.

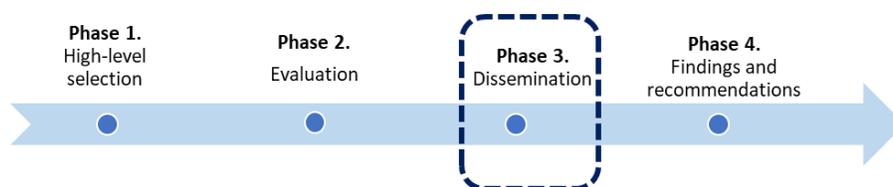
24. Response to recommendations. The response to each recommendation will be the responsibility of at least one department. This department will draft a response in coordination with

¹⁵ Conceptually, the RBM ratings for the achievement of objectives and outcomes feed into the effectiveness ratings.

relevant departments, within six weeks after the release of the evaluation report. It should be clear whether staff accepts, accepts conditionally, or rejects recommendations. Staff may reject a recommendation if (a) it would require more resources than are currently available to implement; (b) it is out of scope of activities the IMF is responsible for or mandated to deliver; or (c) staff disagrees that the recommendation will improve CD outcomes. For all evaluations, ICD will be responsible for putting together the draft staff responses. The draft staff response will be sent to the evaluation team for comments before finalizing. The response could be in a matrix format that presents feedback on each recommendation, a list of agreed actions by the responsible unit(s) for accepted or conditionally accepted recommendations, and a specific time frame for their implementation. When more than one IMF department or stakeholder (i.e., CD users, other development partners) is mentioned, it should be clear which department or stakeholder is responsible for which action(s). Based on this staff response, the implementation of accepted or conditionally accepted recommendations will be tracked by the CCB (see Phase IV).¹⁶

PHASE III. DISSEMINATION AND PUBLICATION

This section outlines the process of disseminating and publishing evaluations included in the strategic work plan. The guidelines, which draws from the [Staff Operational Guidelines on Dissemination of Technical Assistance Information](#) and RBM Governance Framework, aims at increasing transparency and promoting learning, without compromising the Fund’s role as a trusted advisor.



25. Dissemination to staff. All documents related to evaluations in the strategic work plan will be accessible to Fund staff. This includes the final evaluation report (main and annex), inception report, ToRs, as well as the response to recommendations, unless classified as confidential or strictly confidential. Nevertheless, any classified evaluation document should be the exception, which requires careful consideration at the selection stage (Phase I). If the documents contain information provided by the recipient on the understanding that it will remain confidential, then the recipient’s consent will be required.¹⁷

26. Dissemination to the Executive Board and partners. All final main reports (excluding annexes) and staff’s response to recommendations will be made available to the Executive Board and shared with partners who contributed to financing a project included in the evaluation, unless classified as confidential or strictly confidential. Main reports are not expected to contain final TA

¹⁶ The RCDCs also follow up annually on the implementation of relevant recommendations at their Steering Committee meetings. Annual reports on RCDC activities and program documents for their subsequent phases explicitly present how recommendations have been or will be implemented.

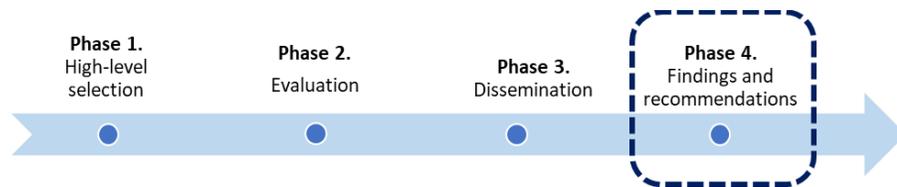
¹⁷ According to the Transparency Guidelines for Non-Board Documents (Appendix I of [Guidance Note on the Fund’s Transparency Policy](#)).

advice, which the dissemination policy considers confidential. Main reports are also expected to avoid reflecting country-specific assessments or ratings, if feasible. If the report contains final TA advice or country-specific assessments, then the consent of the CD recipients in the countries referred to in the report will be required on a *non-objection basis*, prior to dissemination.¹⁸ Other documents such as inception reports and ToRs, as well as annexes, will not be disseminated to the Board or partners, unless the partner is an ESC member and will therefore have access in this role.

27. Publication. All final main reports (excluding annexes) and responses to recommendations will be published, unless classified as confidential or strictly confidential. If the main report contains final TA advice or country-specific assessments, then the *explicit consent* of the CD recipients in the countries referred to in the report and the approval of the head of the managing department will be required, prior to publication. Internally led evaluations will be issued to the Board for information, before publication. Inception reports and ToRs will not be published.

PHASE IV. ADDRESSING FINDINGS AND RECOMMENDATIONS

This section covers distilling findings and lessons across evaluations and follow up on recommendations.



28. Periodic report to the Board. The CD Strategy Review, conducted every five years, will include a report that will consolidate key findings from evaluations in the strategic work plan and report on efforts to address the recommendations. In the interim period, a similar report will be shared with the Board as a stand-alone document. The interim report is expected to be sent to the Board no later than three years after the publication of the last CD Strategy Review.

29. Annual report to the CCB. ICD, in coordination with the managing departments, will present an annual summary of findings and recommendations from evaluations completed in the previous fiscal year. The implementation status of accepted recommendations from previous evaluations will also be reported and discussed in the CCB each year and posted on the ICD evaluations landing page.

30. Retiring/revising recommendations. When needed, revising responses to recommendations is possible to ensure their relevance to changing conditions. Moreover, if some

¹⁸ Similar to a TA report, consent will be deemed obtained unless the CD recipient objects to such dissemination within 60 days of the transmittal of the final evaluation report. After the expiration of the 60-day period, the recipient may still object to the dissemination provided the report has not yet been disseminated.

recommendations are assessed as superseded by events, no longer valuable, and/or redundant, they may be proposed to be retired. The CCB will approve the basis for retiring or revising recommendations in the summer CCB meetings. The tracking of follow-up to recommendations is also expected to expire after four years and the CCB should decide on the basis for retiring or revising a recommendation by then.

Annex I. The OECD-DAC Criteria and Example Questions

1. **This Annex presents the definitions of each criterion, based on the OECD-DAC,¹** and guidelines to their application in the context of Fund CD projects. Table A1 also presents the six OECD-DAC criteria and example questions. The questions are indicative—the evaluation manager and evaluation team may refine the questions to best serve the purpose of the evaluation.
2. **The criteria should be applied thoughtfully, not mechanistically.** Initially, all six criteria should be considered alongside the purpose, scope, and context of the evaluation. They should then be narrowed down to address the needs of the users of the evaluation. Data availability, resource constraints, timing of the evaluation, and methodological considerations may also influence how (and whether) a particular criterion is covered. It is however important to strike a balance between flexibility and cherry picking. The criteria are interlinked, not mutually exclusive, and should be considered together to reach a higher level of understanding of how CD projects have helped recipients.
3. **The DAC criteria are broadly assessed with respect to the CD’s net benefits to the recipient country’s institutions.** This is straightforward with regards to TA as its objectives and outcomes are defined at the country level. For multi-country training (face-to-face or online) and workshops, a CD project potentially benefits several countries’ institutions. In this case, the criteria would still be assessed with respect to the net benefits to the participant countries’ institutions, and not just the training/workshop participants.

Relevance: Is the CD Project Doing the Right Thing?

4. **Relevance is the extent to which the project objectives and design respond to** beneficiaries’ global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change.

IMF context. The relevance criterion will focus on the needs and priorities of the country and the specific agency that received the CD and assess three complementary dimensions below:²

- **Addressing critical capacity gaps.** The evaluation team will assess whether the project objectives and design, as defined in the RBM log frame, tackle the critical capacity gaps of the country and the specific agency that received the CD. Critical capacity gaps and the need to tackle them could be identified by (a) the recipient agency and national government; (b) IMF surveillance/program and technical diagnostics; and (c) international norms and standards. Note

¹ OECD-DAC Network on Development Evaluation, [Better Criteria for Better Evaluation Revised Evaluation Criteria Definitions and Principles for Use](#), December 2019.

² If a development partner requests an assessment of the alignment of the objectives and design of an externally financed project with its own needs/policies, such an assessment will not be covered under relevance or any other DAC criteria. Nevertheless, if the ESC agrees to it, the evaluation can respond to this request as a separate question under the non-project-specific evaluation section, i.e., outside of the core evaluation covered by the CEF.

that for multi-country trainings and workshops, relevance will be assessed with respect to the critical capacity gaps of the participants' agencies and government institutions.

- **Tailoring.** The evaluation team will look at whether the project design was sufficiently tailored to the economic, political economy, and capacity conditions of the country/agency. This aspect is particularly important in the context of fragile states and low-income countries, where absorptive capacity is low. In these countries, milestones, target indicators, and the general pace of reform should on average be less ambitious compared to emerging market and middle-income economies. The evaluation team will also assess whether the choice of modality and training/TA mix matches the technical capacity of the individual and institutional CD recipients.
- **Adapting to change.** The evaluation team will look at how project objectives and design were adapted to changes in the economic, political economy, and capacity conditions during implementation. The evaluation will assess whether the potential risks were identified adequately and to what extent the project management mitigated risks that would undermine the project objectives, impact, and sustainability. The objectives and design of CD are also expected to duly consider and respond to any intended and unintended effects of CD on environmental, equity, and social conditions.

Coherence: How Well does the CD Project Fit with Other Engagements?

5. Coherence is the compatibility of the intervention with other interventions in a country, sector, or institution.

IMF context. The coherence criterion will focus on three aspects.

- **Internal coherence on the recipient side** will assess the compatibility of CD objectives with other actions of the government and the specific agency that receive the CD, as well as with government commitments to international norms and standards. Internal coherence captures related policy actions and other government initiatives (e.g., legal, human resource/staffing, IT) that support or undermine the CD project. Internal coherence will also assess country ownership, i.e., whether both the government and the agency agree that the CD tackles a critical capacity gap and prioritize its implementation. If the CD is supported only by the agency but not by the government, the evaluation team will flag the lack of ownership at the national government level.
- **Internal coherence on the Fund side** will assess if CD objectives are consistent with IMF surveillance/program priorities, as highlighted in surveillance/program country reports or by country teams, given the leading role of area departments in establishing Fund CD strategies and priorities. In addition, the evaluation team will assess the consistency of the CD project with policy recommendations from surveillance and Fund-supported arrangements, and technical recommendations from other Fund CD.

- **External coherence** will assess Fund collaboration with other CD providers in the same area to achieve higher impact, using opportunities for complementary CD and avoiding duplication of effort. The evaluation team will identify other providers on the ground, presence of duplicate CD delivered by other providers, and evidence of collaboration or missed opportunities for complementary CD.

Effectiveness: Is the CD Project Achieving its Objectives?

6. Effectiveness is the extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.

IMF context. The effectiveness criterion will assess the extent to which the CD project achieved, or is expected to achieve, the *intended* results envisaged in the RBM log frame. Here are some guidelines:

- **Ascertain the intended results of the CD project.** In the Fund’s RBM framework, the intended results are reflected by the CD project’s objectives and related outcomes. However, as these are chosen from the RBM catalog, they do not always follow SMART (specific, monitorable, attainable, relevant, and time-bound) principles, making their achievement difficult to assess, and they are not always tailored to country conditions. The evaluation team will therefore need to use other elements of the RBM log frame, such as target indicators and milestones, and check project assessments, to determine the specific intended results of the project. If these elements are not available or insufficient (e.g., missing or vague entries), the evaluation team should consult the project manager and CD expert. If the project has more than one objective, the evaluation team should weigh their relative importance.
- **Use the RBM outcome ratings as one input to the assessment.** As RBM ratings are largely self-assessments, the evaluation team will need to corroborate the rating through other instruments included in the evaluation methodology, such as interviews with different stakeholders, surveys, and case studies.
- **Describe the project’s contribution to the intended results.** The evaluation team will assess to what extent the project *contributed to* the achievement of the intended results, as other factors could have affected the outcome, including but not limited to projects by other IFIs as well as exogenous factors. This could be done qualitatively, in the absence of a good counterfactual.³
- **Assess differential results across groups, if deemed relevant for the purpose of the evaluation.** After discussing with the evaluation manager its usefulness to the purpose of the evaluation, the team could examine how the intended results affected different groups.

³ Alternative approaches to assess causality, i.e., the attribution/contribution of the project to its intended outcomes and impact, include experimental, theory-based, case-based, and participatory approaches. The most suitable approach or the combination of approaches depends on the context. For further guidance see Stern, Stame, Mayne et al. (2012).

Efficiency: How Well are Resources Being Used?

7. Efficiency is the extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.

IMF context. The efficiency criterion will assess the net value of the project relative to similar interventions by the Fund and/or other development partners. The evaluation team will strive to quantify both benefits and costs to perform a cost-benefit analysis. Data however may not always be available. Here are some guidelines:

- **Be familiar with available cost data.** Project-level costs would be available for projects initiated after the CDMAP launch. For older projects, the evaluation team would have to assess at a higher level of aggregation (e.g., topic-country in an RCDC), or estimate costs from Travel Information Management System (TIMS) mission counts and/or Time Reporting for Analytic and Costing Estimation System (TRACES) full-time equivalent. Costs are typically measured in terms of full-time equivalent or U.S. dollar.
- **Use the RBM indicators, if available, to capture benefits.** Value-based verifiable indicators can be used to normalize costs (e.g., cost per percentage reduction of the energy subsidy). If the indicators are binary or qualitative, the evaluation team could choose a comparison group of projects that aimed at and achieved the same objectives and outcomes (based on the indicator), and compare their costs. Note, however, that the evaluation teams should also take into consideration the contribution of Fund CD to the benefits, as other factors could have led to the outcome.
- **Assess operational efficiency.** The evaluation team could also examine the process and implementation of the project and include measures of excessive staff turnover or unnecessary delays, and analysis of quality of experts, outputs, and backstopping, as well as mix of delivery modalities.

Impact: What Differences does the CD Project Make?

8. Impact is the extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.

IMF context. The impact criterion will assess the CD project's contribution to the Fund CD's strategic goal of helping member countries achieve macro and financial stability and inclusive growth by strengthening the institutional capability of their agencies/governments to implement sound policies. A narrower definition of the criterion is the CD project's intended or unintended effects on the Fund's strategic priorities (e.g., inequality, climate change, gender) in the country. Here are some guidelines:

- **Distinguish impact from effectiveness.** Effectiveness is assessed relative to the objectives and related outcomes in the RBM log frame. Impact is assessed on goals that are (a) higher-level than RBM objectives, such as the Fund's mandate of macro and financial stability in the member

country and/or (b) based on Fund priorities of inclusive growth (inequality, gender, financial inclusion) and climate change in the member country.

- **Discuss with the evaluation manager how to assess the impact criterion.** Unlike objectives and outcomes used in assessing effectiveness, higher-level goals are not defined in the RBM log frame. Moreover, the impact of CD on many higher-level goals such as macro and financial stability may take years to realize. For such goals, the assessment of impact of ongoing or recently completed CD may not be appropriate. Some CD however could have an immediate impact on vulnerable groups. In this case, the evaluation manager may decide to assess the impact criterion. As such, the evaluation team should clarify how impact will be assessed in the evaluation.
- **Describe the project's contribution to higher-level goals.** As in assessing effectiveness, the evaluation team's assessment will describe not only if the project has an impact but also the extent of the impact, taking into account other exogenous developments/events that affect the achievement of higher-level goals. In the absence of a good counterfactual, this could be done qualitatively.

Sustainability: Will the Benefits Last?

9. Sustainability is the extent to which the net benefits of the intervention continue, or are likely to continue.

IMF context. The sustainability criterion will assess the continuation of the *actual* or *projected* net benefits of CD (i.e., *positive* effects attributed to CD under effectiveness and impact) over the medium and long term, after the CD is completed. Sustainability will assess to what extent the financial, economic, social, environmental, and institutional capacities of the systems are (or are likely to be) in place to sustain these net benefits. The evaluation team will examine the political support from the government, the adequacy of financial and IT resources, socio-cultural environment, and institutional capability (including sufficiently trained staff with required skills, systems and incentives in place to sustain behavior change, and higher-level support within the agency).

Table A1. Common Definitions for the OECD-DAC Criteria and Example Questions	
DAC Criteria	Key Evaluation Questions
Relevance	<ul style="list-style-type: none"> • To what extent were the CD objectives derived from capacity gaps identified by the authorities, IMF surveillance/program, and other partners/institutions? Do the national authorities consider the CD objectives among the priorities of the country and/or agency? • Was the design sensitive to the context (e.g., economic, political economy, technical capacity) in which it took place? • Were the CD objectives and design successfully adapted to changing circumstances?
Coherence	<ul style="list-style-type: none"> • To what extent does the CD project support or undermine other interventions (particularly policies), and vice versa? <ul style="list-style-type: none"> ○ <i>Internal coherence:</i> Assess synergies and interlinkages between the CD project and other interventions carried out by the agency/ government. Also, assess the consistency of the CD project with other IMF recommendations from surveillance, program, and other CD. ○ <i>External coherence:</i> Assess the consistency of the CD project with interventions by development partners.
Effectiveness	<ul style="list-style-type: none"> • To what extent were the CD outcomes and objective, as defined by the RBM framework, achieved or are likely to be achieved? • In the process, validate the RBM ratings for outcomes and objective. • Assess to what extent the observed direct results were attributed to/ happened as a result of the CD.
Efficiency	<ul style="list-style-type: none"> • Were the CD inputs converted to outputs, outcomes, and impact in the most cost-effective way possible, as compared to feasible alternatives?
Impact	<ul style="list-style-type: none"> • To what extent has the CD project enhanced the country's macroeconomic and financial stability and supported inclusive growth? • Assess any consequences of the CD project on relevant Fund strategic priorities, such as climate change and inequality (economic, gender, and financial inclusion).
Sustainability	<ul style="list-style-type: none"> • To what extent does preserving the net benefits of the CD project hinge on its continuation? Discuss the likelihood of the net benefits continuing over the medium and long term. • Assess the capacity of the recipient country's system to sustain the net benefits of the CD project over time.