

Approach Paper

Evaluation of the Development Effectiveness Framework

Copyright © 2023 Inter-American Development Bank. This work is licensed under a Creative Commons IGO 3.0 Attribution - NonCommercial - NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license

(<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) and may be reproduced with attribution to the IDB and for any non-commercial purpose. No derivative work is allowed.

Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IDB's name for any purpose other than for attribution, and the use of IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this CC-IGO license.

Note that link provided above includes additional terms and conditions of the license.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.



Inter-American Development Bank, 2023

Office of Evaluation and Oversight
1350 New York Avenue, N.W.
Washington, D.C. 20577
www.iadb.org/evaluation

RE-583-1
September 2023

CONTENTS

[ACRONYMS AND ABBREVIATIONS](#)

I.	INTRODUCTION	1
II.	BACKGROUND ON THE DEVELOPMENT EFFECTIVENESS FRAMEWORK	2
III.	THEORY OF CHANGE	7
IV.	EVALUATION OBJECTIVE, SCOPE, QUESTIONS, AND METHODOLOGY	10
	A. Evaluation objective and scope	10
	B. Evaluation questions and methodology.....	10
V.	EVALUATION TEAM AND TIMELINE	12

[REFERENCES](#)

Annex I	Evaluation design matrix
Annex II	Selected findings from previous OVE evaluations
Annex III	Similar evaluations from other International Financial Institutions

This document was prepared by Claudia Figueroa Garcia-Huidobro (team leader), Jorge Gallego, Luisa Riveros, Luis Fernando Corrales, Lucero Vargas, Clarisa Yerovi and Andreia Barcellos, with technical inputs provided by Carolina Romero and Cesar Bouillon, under the direction of Ivory Yong Protzel, OVE Director.

ACRONYMS AND ABBREVIATIONS

CRF	Corporate Results Framework
CS	Country Strategy
DEF	Development Effectiveness Framework
DEM	Development Effectiveness Matrix
DEO	Development Effectiveness Overview
IDB	Inter-American Development Bank
IDB-9	Agreement of the 9th General Capital Increase of the IDB
KIC	Knowledge, Innovation and Communications Sector
KLD	Knowledge and Learning Division
M&E	Monitoring and evaluation
NSG	Non-sovereign guaranteed
OPC	Operations Policy Committee
OVE	Office of Evaluation and Oversight
PCR	Project Completion Report
PMR	Progress Monitoring Report
SG	Sovereign guaranteed
SPD	Strategic Planning and Development Effectiveness
TOC	Theory of change

I. INTRODUCTION

- 1.1 This document defines the approach the Office of Evaluation and Oversight (OVE) will adopt to evaluate the Development Effectiveness Framework (DEF). It includes the objectives, scope, questions, and methodology that OVE will use for the evaluation. This corporate evaluation was included in OVE's 2022–2023 work program (document [RE-563](#)), as approved by the Inter-American Development Bank (IDB) and IDB Invest Boards.
- 1.2 **The IDB Group has tools and practices to support development effectiveness. The evaluation will focus on the IDB side.** At the Bank, the DEF was launched in 2008, and encompasses a series of instruments, governance arrangements, and reporting mechanisms to produce tangible and positive results for beneficiaries.^{1,2} The purpose of the evaluation is to assess whether the DEF has successfully enhanced the effectiveness of the Bank's projects by fostering an institutional culture of achieving and demonstrating results. This will be achieved through an analysis of the DEF's relevance, implementation, use, and results.
- 1.3 **This evaluation differs from previous assessments of the DEF in that it examines the DEF as a system, considering its three-pillar structure.** OVE evaluated the DEF as part of the mid-term and final evaluations of the Agreement of the Ninth General Capital Increase of the IDB (IDB-9)³ in 2013 and 2018, respectively. These evaluations focused on assessing compliance with the requirements stipulated in the IDB-9 Agreement for DEF instruments, Development Effectiveness Matrix (DEM), Progress Monitoring Report (PMR), and Project Completion Report (PCR)) and the Development Effectiveness Overview (DEO).⁴ Additionally, OVE conducts an annual validation exercise of PCRs and reports on its findings (please refer to Annex II for a summary of these evaluations). Unlike the previous ones, the current evaluation assesses all elements of the DEF for sovereign guaranteed (SG) projects, including governance and reporting mechanisms in addition to its instruments.
- 1.4 **Similar evaluations conducted by other institutions have identified relevant issues that have been considered in the design of this evaluation.** The self-evaluation systems of the African Development Bank, the European Bank for Reconstruction and Development, the World Bank Group, and the International Monetary Fund have been evaluated,⁵ and a common finding is that incentives influence the performance of such systems. Reports indicate that shortcomings tend to be more related to incentives and behaviors, rather than to the design of

¹ "Development Effectiveness," IDB website, <https://www.iadb.org/en/office-strategic-planning-and-development-effectiveness/development-effectiveness> . (Accessed April 24th, 2023).

² On the IDB Invest side, the Impact Management Framework is used to "build, measure, and manage a portfolio of financially sustainable investments that contribute to reaching the Sustainable Development Goals." Page 8, from: [IDB-Invest Managing a Portfolio for Impact: IDB Invest's Impact Management Framework](#) (Washington, D.C.: IDB-Invest, 2020). The assessment of IDB Invest's Impact Management Framework is included in the IDB Invest corporate evaluation. (Washington, D.C.: IDB-Invest, 2020).

³ IDB's Ninth General Capital Increase: Implementation and Results. Revised version (document [RE-515-6](#)) and Overview: Mid-term Evaluation of IDB-9 Commitments (document [RE-425](#)).

⁴ Annex I of the Technical Note presents the IDB-9 requirements for the DEF.

⁵ The EBRD's evaluation corresponds to an external review that also assessed the independent evaluation function.

self-evaluation instruments or processes. In most cases, evaluations found that self-evaluations lacked sufficient candor, which was primarily derived from a flawed incentive structure. In two instances, evaluations noted that even candid self-evaluations did not lead to action due to a lack of incentives to identify and address problems. The use of lessons learned from self-evaluations was often limited, partly due to issues with the quality and documentation of the lessons. The lessons were either too general or very specific to be applicable or were not supported by sufficient evidence (refer to Annex III for a summary of the findings from these evaluations). The evaluation design incorporates the assessment of incentives, learning mechanisms, and lessons learned.

- 1.5 **This evaluation complements the Bank's efforts to improve the development effectiveness of its projects.** The 2021 DEO claimed that “as a development institution, the most critical commitment of the IDB Group is driving toward development results.”⁶ Development effectiveness has been identified as a key priority in the past years. Most recently stemming from the Barranquilla Resolution of 2021 and the Washington Resolution of 2022, Management has developed and presented to the Board a series of analytical works on development effectiveness⁷ and has proposed several changes to DEF elements. Management’s proposals will be considered in the evaluation to the extent that sufficient information is available to make an informed assessment.

II. BACKGROUND ON THE DEVELOPMENT EFFECTIVENESS FRAMEWORK

- 2.1 **The term “development effectiveness” refers to achieving development objectives.** The OECD DAC Network on Development Evaluation (EvalNet) defines *effectiveness* as “the extent to which the intervention achieved, or is expected to achieve, its objectives and its results, including any differential results across groups.”⁸ The term “*development effectiveness*” refers to ensuring that development resources achieve their development objectives.⁹ The term gained relevance in the 2000s within the international community, especially after the Paris Declaration on Aid Effectiveness of 2005. This Declaration, endorsed by the Bank, defined five best practice principles for the effective management of development resources, including one on results: “developing countries and donors shift focus to development results and results get measured”.¹⁰ Subsequent high-level international meetings have reviewed these principles. The latest update

⁶ Document [GN-3059](#), page ix.

⁷ Documents [GN-3118-1](#), [GN-3118-2](#), and [PP-1262](#).

⁸ OECD DAC Evaluation Criteria. <https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm> (Accessed April 24th, 2023).

⁹ Adapted from: Center for Global Development, [Development Effectiveness in the “New Normal”: What Do the Changing Roles and Purposes of ODA Mean for the Effectiveness Agenda?](#) (Washington, D.C.: CGD, 2022).

¹⁰ Paris Declaration and Accra Agenda for Action. <https://www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm> (Accessed April 24th, 2023). The other four principles are the following: i) ownership: developing countries set their own strategies for poverty reduction, improve their institutions and tackle corruption; ii) alignment: donor countries align behind these objectives and use local systems; iii) harmonization: donor countries coordinate, simplify procedures and share information to avoid duplication; and iv) mutual accountability: donors and partners are accountable for development results.

to the principles was made with the Busan Partnership Agreement in 2011, which recognized the diversity of development actors and kept four principles: country ownership, transparency and mutual accountability, focus on results, and inclusive partnerships.¹¹

2.2 **The DEF aimed for “a greater focus on results, based on hard evidence, while improving the quality of the effort.”**¹² The DEF was launched in 2008 with a clear purpose: “Through the DEF, Management intends to increase the effectiveness of all of the Bank’s products through: (i) setting clear standards and metrics for the evaluation of all development interventions; (ii) providing clear guidance to staff about analytical requirements for meeting the standards; (iii) aligning governance structures to comply with those set out as good practice standards; and (iv) establishing a results framework incorporated in the Corporate Performance Framework to monitor progress in key development effectiveness indicators.”¹³

2.3 **OVE has reconstructed the intended general and specific objectives of the DEF.** The first document establishing the DEF in 2008 (document [GN-2489](#)) stipulated the DEF’s broad purpose and key elements for implementation (see paragraph 2.2). However, the broad purpose and implementation elements do not constitute a clear vertical logic, which is needed to evaluate the DEF. The purpose is at the level of impact—to increase effectiveness—and the key implementation elements are at the level of activities and outputs—setting clear standards and metrics [...], providing clear guidance to staff [...], aligning governance structures to comply with GPS [...]. Missing between these levels are the specific objectives that reflect expected changes at an outcome level and which help track progress toward the broad purpose or general objective. Following common practice in evaluation, when not all the objectives are explicitly formulated or are not at the appropriate level, OVE reconstructed the intended objectives of the DEF, filling in the gaps. This more comprehensive formulation of the objectives will be used to evaluate the DEF’s performance. Based on the initial document, OVE identified one general and four specific objectives of the DEF:

General objective: Enhance the effectiveness of the Bank’s products by fostering an institutional culture of achieving and demonstrating development results.¹⁴

- **Specific Objective 1: Governance arrangements enable the effective implementation of the DEF.** A reviewed governance aligning roles,

¹¹ Subsequent meetings have upheld these principles. The latest meeting was the [2022 Effective Development Co-operation Summit](#).

¹² Document [GN-2489](#) para. 2.5 and document [GN-2489-2](#) para. 1.2.

¹³ Document [GN-2489](#), para. p.1.1.

¹⁴ This intended general objective was reconstructed to reflect the profound institutional changes the DEF aimed to make. The document establishing the DEF stated that: “the proposal [establishing the DEF] will provide the Bank with the tools needed to continuously assess its performance. It will promote discipline in thinking about how to achieve and measure results. It will allow the Bank to learn from past experience, which ultimately will lead to increased effectiveness of all its interventions. The proposal differs from previous efforts in that it adopts the international standards jointly developed by the Multilateral Development Banks in the past five years and organizes the activity of the Bank under a single framework focused on achieving results. It not only focuses on metrics but also on the Bank’s incentive structure. By aligning performance incentives with the achievement of development results, the DEF establishes the means by which to recognize success and foster accountability” (Document [GN-2489](#), para 1.2).

responsibilities, incentives, and processes of DEF-related stakeholders will underpin the achievement of the DEF's general objective.

- **Specific Objective 2: Design, monitoring, and evaluation throughout the intervention's cycle are improved.** An improved design, monitoring, and evaluation system will enable the Bank to achieve and demonstrate the results of its interventions based on empirical and solid evidence.
- **Specific Objective 3: Accountability for development results is enhanced.** Improved mechanisms to track progress in the achievement of the Bank's objectives will increase results accountability.
- **Specific Objective 4: Learning from past experience is increased.** A stronger learning culture will allow the Bank to learn from experience, understanding what works and why to increase effectiveness.

2.4 **The DEF is currently implemented for both sovereign-guaranteed operations and Country Strategies, with the evaluation focusing only on the former.** Initially, the DEF intended to encompass both SG and non-sovereign guaranteed (NSG) operations, knowledge products, and country strategies. However, at present, it is only applied to SG projects and country strategy (CS). While the DEF for the NSG projects managed by the Bank was developed and put into effect,¹⁵ these operations were later transferred to IDB Invest after the private sector merge-out, and therefore the DEF for NSG was no longer applied.¹⁶ On the other hand, the DEF for knowledge products was not implemented; knowledge products follow decentralized approval and quality assurance processes that vary between originating units and funding mechanisms.¹⁷ The evaluation will solely focus on the DEF for SG projects (refer to paragraph 4.2 for the reasons for this scope).

2.5 **The DEF is a comprehensive system consisting of three interconnected pillars — instruments, governance arrangements, and reporting mechanisms.** The DEF was conceived as a system, focused not only on the *instruments* to measure the achievement of objectives but also on the Bank's incentive structure through the *governance* and *reporting* pillars.¹⁸ Within the first pillar, the DEF defined instruments to support the design, implementation, and evaluation of SG projects (figure 2.1). At *entry*, a development effectiveness matrix is filled out for SG projects. During *implementation*, a PMR is filled out. At *exit*, SG projects require a self-evaluation called the PCR. While not a part of Management's DEF system, OVE has DEF-related products including the validation of all Management's PCRs.

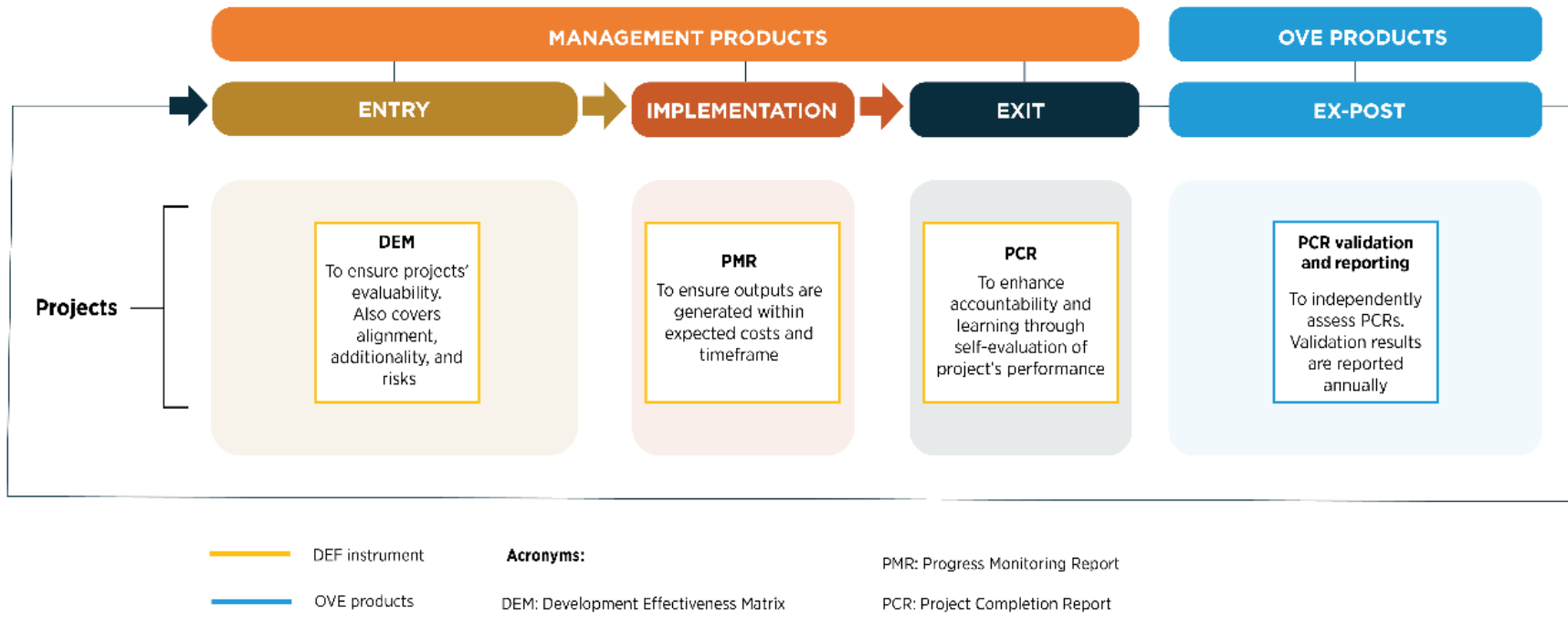
¹⁵ Documents [GN-2473-1](#) and [GN-2489-8](#).

¹⁶ IDB Invest operates under different development effectiveness arrangements.

¹⁷ OVE's 2019 Review of Knowledge Generation and Dissemination in the Inter-American Development Bank (document [RE-517-2](#)) found that knowledge production was still decentralized, and quality controls varied depending on the funding mechanism and originating unit. Management is working on improvements, as reported through the ReTS system. Additionally, there has been progress on the monitoring system for technical cooperation, which includes the use of a results matrix (document [OP-1385-4](#)).

¹⁸ *Ibid.*, para. 1.2.

Figure 2.1. DEF Instruments for SG projects



Source: OVE.

2.6 **The second pillar seeks to align all stakeholders' actions towards the effective implementation of the DEF and to avoid gaps and duplications.** The governance pillar focuses on defining the roles and responsibilities of all actors involved in the DEF implementation, supported by a set of regulations, processes, and incentives. Governance arrangements were outlined at a high level in the original DEF design and have since been incorporated into the Bank's system. OVE identified three main areas for which roles and responsibilities can be delineated. The first area refers to setting the standards, guidelines, regulations, and processes to implement the DEF. The second area involves the preparation of the DEF-related products (e.g., DEM, PMR, DEO, and CRF), and the third area involves the review and use of DEF-related products. The main stakeholders are the following:

- Strategic Planning and Development Effectiveness (SPD) is involved in all three areas. In the first area, SPD proposes quality standards for the design and execution of projects to ensure their evaluability, relevance, efficiency, and effectiveness. It develops instruments' guidelines and templates. In the second area, SPD provides technical support and guidance throughout the process of preparing the information for DEF instruments. SPD also updates the CRF indicators and prepares the DEO (together with the teams managing development effectiveness at IDB Invest and IDB Lab). In the third area, SPD reviews the information of DEF instruments for projects completely or partially.
- Project teams participate in the second area, primarily preparing and completing the information required by DEF instruments.
- Vice Presidency for Countries and Vice Presidency for Sectors: participate in both the first and third areas. In the first area, they prepare the regulations for SG operations, which includes the processes to fill out the DEF instruments. In the third area, they participate in reviewing DEF products.
- Knowledge, Innovation and Communications Sector (KIC), through the Knowledge and Learning Division (KLD), participates in the first and third areas. In the first area, KLD coordinates capacity building through operational training programs. In the third area, KLD has a role in capturing lessons learned from operational work and promoting their systematization, dissemination, and application.
- Operations Policy Committee (OPC) participates in the third area. One of the main functions of OPC is to review and approve operations before they are submitted to the Board.

2.7 **The third pillar —reporting mechanisms— seeks to report IDB's development effectiveness to internal and external audiences.** Through this pillar, the information is systematized, aggregated, and disseminated to keep track of the Bank's performance in relation to development effectiveness. The Bank has two corporate reporting tools related to development effectiveness,¹⁹ the CRF and the DEO, both publicly available. The CRF monitors the institutional strategy and includes two development effectiveness targets derived from the PMR and

¹⁹ DEF instruments feed other corporate reporting tools, like the Quarterly and Annual Business Reviews (QBR and ABR, respectively).

validated PCRs, as well as other DEF instrument data to report additional targets. The DEO is the IDB Group's annual report on development effectiveness.

- 2.8 **OVE participates in the DEF's implementation, although it is not subject to this framework.** OVE products respond to OVE's annual work program, which is approved by the Boards of IDB and IDB Invest. Nonetheless, OVE is involved in the implementation of the DEF as part of its responsibility to ensure the integrity of the IDB Group's self-evaluation system.²⁰ Specifically, OVE validates self-evaluation reports of SG projects and produces a synthesis report on an annual basis, following guidelines prepared by Management that are consulted with OVE and are based on ECG GPS. In the past, OVE has validated the evaluability of SG projects and has reviewed elements of the DEF as part of the mid-term and final evaluation of IDB-9 commitments.²¹

III. THEORY OF CHANGE

- 3.1 **OVE reconstructed the Theory of Change (TOC) of the DEF to present its objectives and the pathways to reaching them.** This TOC has been reconstructed using the first document establishing the DEF in 2008 (document [GN-2489](#)), along with corporate and operations regulations, and instruments' templates and guidelines. Organized around the DEF's three pillars, the TOC explains the causal logic that links inputs and actions by stakeholders to key outputs, which correspond to the organizational structure and tools used to assess the Bank's performance. These outputs then lead to outcomes that contribute to the overarching impact of enhancing the effectiveness of the Bank's products by fostering an institutional culture of achieving and demonstrating development results.²² The TOC also articulates a series of key underlying assumptions required for the planned interactions to occur from the bottom to the top levels of the causal chain (see Figure 3.1). The DEF inputs, activities, outputs, and outcomes (intermediate and final) identified in the TOC are described below. Assumptions are highlighted in italics between links:

- Inputs comprise of staff, information systems, instruments templates and guidelines, a results framework, and the description of coordination mechanisms and responsibilities.

Instruments and corporate reporting mechanisms contain the relevant areas and domains, while the roles and responsibilities are comprehensively and clearly defined with no duplicities.

- Activities involve the preparation of DEF instruments, systematization and aggregation of progress on CRF indicators and DEO, the formulation of regulations and processes, as well as the definition of incentives mechanisms to operationalize the DEF.

The execution of activities requires that processes and regulations cover all DEF-related implementation areas and that staff possess the capacity

²⁰ Document [RE-538-5](#).

²¹ IDB's Ninth General Capital Increase: Implementation and Results. Revised version (document [RE-515-6](#)) and Mid-term Evaluation of IDB-9 Commitments (document [RE-425](#)).

²² Although the DEF is currently applied to SG projects and CS, the evaluation will be centered on the former. Section V explains the details of the scope of the evaluation.

to prepare and report information at both the project and the corporate levels.

- Outputs are the concrete DEF instruments, the data for corporate-level indicators, the DEO, and the organizational structure.

Information provided by instruments is timely and complies with quality standards to raise signals for course correction. Reporting mechanisms information as well as independent validation also follow quality standards. All stakeholders understand the organizational structure and there are clear lines of communication coordination and decision-making mechanisms.

- Intermediate outcomes reflect expected changes at each pillar level. Within the instruments pillar, utilizing the information provided by DEF instruments leads to i) improved project's capacity to demonstrate results at completion, ii) supported decision-making and accountability by tracking activities and outputs, and iii) enhanced accountability and learning at the project level. Within the reporting pillar, the intermediate outcome materializes by informing internal and external audiences of the Bank's progress in development effectiveness. Within the governance pillar, the organizational structure leads to aligned DEF stakeholders.

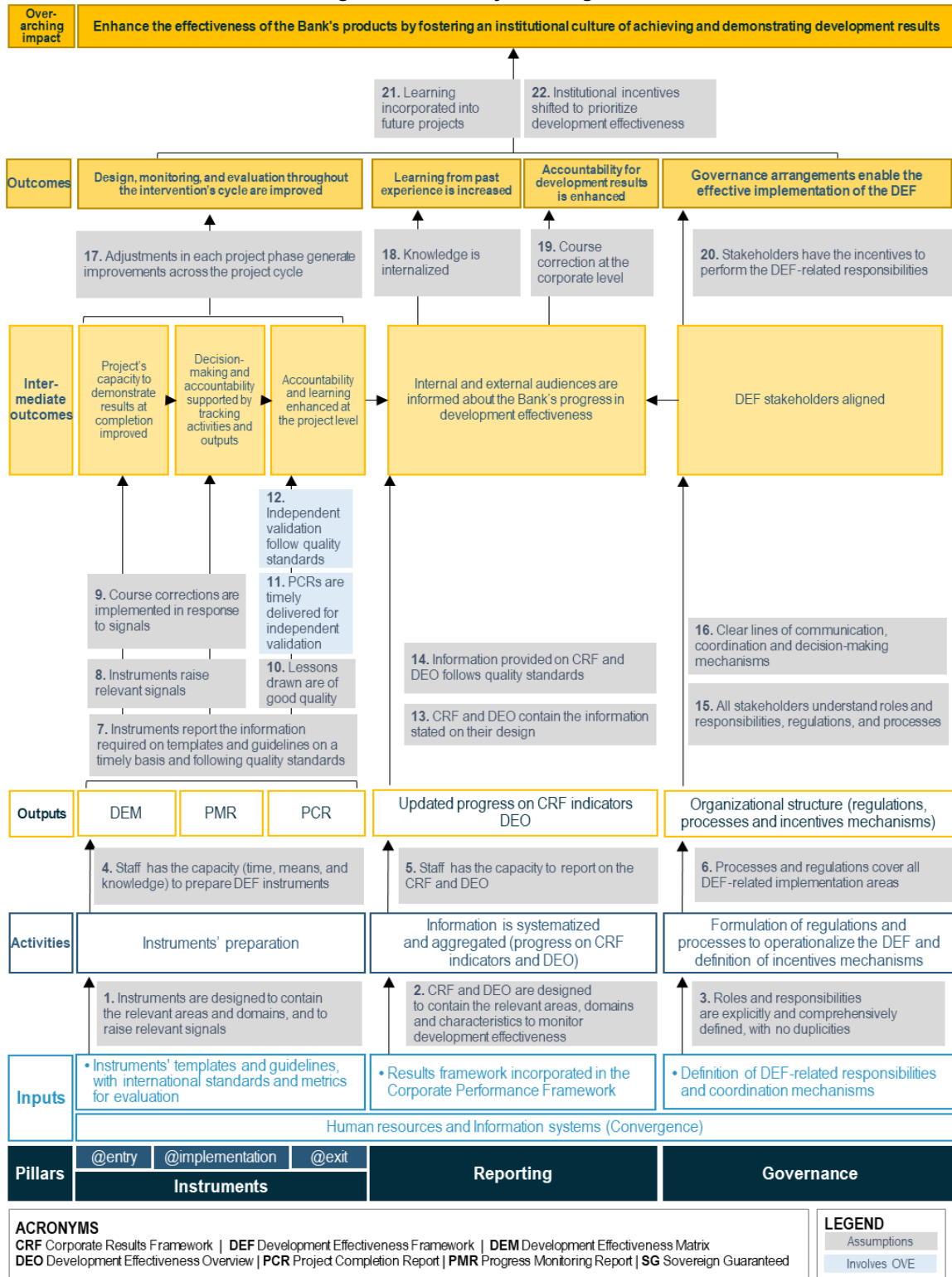
Adjustments in each project phase generate improvements across the project cycle, knowledge is internalized and course is corrected at the corporate level. Stakeholders have the incentives to perform the DEF-related responsibilities.

- Outcomes refer to the expected improvements at the corporate level, which are defined by the four specific objectives OVE reconstructed for the DEF — improved design, monitoring, and evaluation throughout the intervention's cycle, increased learning from past experience, enhanced accountability for development results, and effective implementation of the DEF enabled by governance arrangements for (see para. 2.3).

Learning is incorporated into future projects and Institutional incentives are shifted to prioritize development effectiveness.

- 3.2 **The TOC serves as guide and anchor for evaluating the DEF.** To evaluate the relevance of the DEF, OVE will assess the extent to which it is designed to fulfill its stated objectives, examining the inputs and assumptions necessary for carrying out its activities. Additionally, OVE will evaluate the extent to which the DEF has been implemented as expected, through the execution of activities and outputs. Furthermore, OVE will evaluate the fulfillment of critical assumptions required for realizing outcomes, as well as the degree to which intermediate outcomes have been realized. Finally, OVE will examine the DEF's contribution to its four objectives.

Figure 3.1. A theory of change for the DEF



Source: OVE.

IV. EVALUATION OBJECTIVE, SCOPE, QUESTIONS, AND METHODOLOGY

A. Evaluation objective and scope

- 4.1 **The evaluation will assess the degree to which the DEF is achieving its objectives and identify the factors influencing its performance.** The evaluation approach is objectives-based, using the general and specific reconstructed objectives (as detailed on para. 2.3) as the basis for assessing the DEF's performance. To evaluate the DEF's contribution to its objectives, OVE will consider both the design and implementation of the framework. This will allow determining whether observed results (and perceptions of results in some cases) are related to the DEF's design or implementation, or a combination of both.
- 4.2 **The evaluation will focus on the DEF for SG projects, despite also being applied to CS.** Two reasons explain this scope. Firstly, the DEF for SG projects serves as the basis to demonstrate results at the country level, so an assessment at the project level is a necessary starting point. Secondly, Board-level discussions and Management's proposals to enhance the DEF have been focused on SG projects, underscoring the importance of assessing it at this level.
- 4.3 **The temporal scope of the evaluation is the period 2008 to 2022.** The evaluation period, 2008-2022, covers the entire duration during which the DEF has been implemented. However, specific analyses focus on particular timeframes within this period, as explained in the Evaluation Design Matrix.

B. Evaluation questions and methodology

- 4.4 The overarching question the evaluation seeks to answer is whether the DEF has been successful in enhancing the effectiveness of the Bank's projects by fostering an institutional culture of achieving and demonstrating results. To answer this question, the evaluation will examine the sub-questions enumerated below, which cover four areas: relevance, implementation, use, and results of the DEF.

Relevance:

1. To what extent do the objectives of the DEF adequately respond to the needs identified in the diagnostic and the priorities of the IDB?
2. To what extent is the DEF adequately designed to fulfill its objectives?
3. To what extent has the DEF's design adapted to address the evolving needs and changing priorities of the IDB, as well as the challenges encountered during its implementation?

Implementation:

4. To what extent has the DEF been implemented as expected, including quality standards?

Use:

5. To what extent are DEF products (instruments and reporting mechanisms) being used as expected?

Results:

6. To what extent have the DEF's expected outcomes of (i) having governance arrangements enabling the effective implementation of the DEF, (ii) improving the design, monitoring, and evaluation throughout the

intervention's cycle, (iii) enhancing accountability for results, and (iv) increasing learning from past experience been achieved?

4.5 **OVE will employ mixed methods to answer the evaluation questions.** These methods will include (i) a document review; (ii) an analysis of trends, correlations, and drivers of project scores and ratings from the DEM, PMR, and PCR; (iii) interviews and surveys with Bank staff; and (iv) a benchmarking exercise across a group of similar organizations. In the case that more, fewer, or different types of analysis on certain topics are needed, OVE may adjust the evaluation methods. For more detail on the evaluation questions, sources of information, and methods, please refer to the Evaluation Design Matrix in Annex I.

- Document review: OVE will analyze corporate documents describing the DEF; instruments' guidelines and templates; corporate and operational regulations;²³ the 2020-2023 CRF; DEOs from 2021 to 2023; loan proposals on a sample basis; sector framework documents; and other documents as necessary.
- Quantitative analysis: OVE will compile and analyze a database of project-level data, context-specific data, and scores, ratings, and sub-ratings stemming from DEF instruments.
- Interviews and surveys: OVE will conduct interviews and surveys to gather the perspectives of all stakeholders involved in the DEF (see p. 2.6 for the identification of stakeholders). Interviews will be conducted with SPD, Division Chiefs, Sector Managers, KIC, and selected project team leaders. OVE will approach Board Members, Chiefs of Operations, other project team leaders, and operations analysts via surveys.
- Benchmarking exercise: OVE will assess how the DEF compares to the development effectiveness performance management systems of other institutions for public sector operations. This exercise will involve characterizing the instruments and reporting mechanisms utilized by different institutions and how they are used for course correction. The analysis will also encompass the governance structures and institutional set-up of these systems.

²³

DEF-related responsibilities and processes are reflected in these regulations.

Table 4.1. Evaluation questions and methods

Main evaluation questions	Document review	Trends and Correlations Analysis	Benchmark Review	Interviews and surveys
1. To what extent do the objectives of the DEF adequately respond to the needs identified in the diagnostic and the priorities of the IDB?	X			
2. To what extent is the DEF adequately designed to fulfill its objectives?	X		X	X
3. To what extent has the DEF's design adapted to address new needs and changing priorities of the IDB, as well as the challenges encountered during its implementation?	X			X
4. To what extent has the DEF been implemented as expected, including quality standards?	X			X
5. To what extent are DEF products (instruments and reporting mechanisms) being used as expected?	X	X	X	X
6. To what extent have the DEF's expected outcomes of (i) having governance arrangements enabling the effective implementation of the DEF, (ii) improving the design, monitoring, and evaluation throughout the intervention's cycle, (iii) enhancing accountability for results, and (iv) increasing learning from past experience been achieved?	X	X	X	X

Source: OVE.

V. EVALUATION TEAM AND TIMELINE

- 5.1 Evaluation team: The evaluation team includes Claudia Figueroa (team leader), Jorge Gallego, Luisa Riveros, Luis Fernando Corrales, Mariana Gamarra, Lucero Vargas, Clarisa Yerovi, and Andreia Barcellos. An external consultant will also support the team. The work will be carried out with technical inputs and guidance from Cesar Bouillon, principal economist, and under the supervision of the OVE Director, Ivory Yong-Protzel.
- 5.2 **Timeline:** OVE will conduct its evaluation activities during 2023 and early 2024 and will deliver its final report to the Board of Directors by the end of Q1 2024.

Table 5.1. Indicative timeline of activities

Activity	Date
Approach paper to Board of Executive Directors	September 2023
Draft for Management review	February 2024
Submission to Office of the Secretary for translation and subsequent distribution to the Board of Executive Directors	End of Q1 2024

REFERENCES

- African Development Bank (AfDB). 2020. [*Evaluation of the AfDB's Self-Evaluation Systems and Processes Summary Report, an IDEV Corporate Evaluation*](#). Abidjan: AfDB.
- Center for Global Development (CGD). 2022. *Development Effectiveness in the "New Normal": What Do the Changing Roles and Purposes of ODA Mean for the Effectiveness Agenda?*, CDG Policy Paper 255. Washington, D.C.: CGD.
- Independent Evaluation Group. 2016. [*Behind the Mirror: A Report on the Self-Evaluation Systems of the World Bank Group*](#). Washington, DC: WB
- International Monetary Fund. 2015. *Self-Evaluation at the IMF: An IEO Assessment*. <https://ieo.imf.org/en/our-work/Evaluations/Completed/2015-1001-self-evaluation-at-the-imf-an-ieo-assessment>
- Inter-American Development Bank. 2008. *Development Effectiveness Framework (DEF)*, document [GN-2489](#). Washington, D.C.: IDB.
- .2008. *Development Effectiveness Framework for SCF projects*, document [GN-2473-1](#), Washington, D.C.: IDB.
- .2009. *Development Effectiveness Framework: Implementation Progress Report*, document [GN-2489-2](#). Washington, D.C.: IDB.
- .2012. *Overview: Mid-term Evaluation of IDB-9 Commitments*, document [RE-425](#). Washington, D.C.: IDB.
- .2013. [*Mid-term Evaluation of IDB-9 Commitments: The Development Effectiveness Framework and the Development Effectiveness Overview*](#), Background Paper Washington, D.C.: IDB.
- .2013. *Revised Development Effectiveness Framework for Non-Sovereign Guarantee Projects. Revised version*, document [GN-2489-8](#), Washington, D.C.: IDB.
- .2018. *IDB's Ninth General Capital Increase: Implementation and Results. Revised version*, document [RE-515-6](#). Washington, D.C.: IDB.
- .2018. *Review of Knowledge Generation and Dissemination in the Inter-American Development Bank*, document [RE-517-2](#), Washington, D.C.: IDB.
- .2019. *The Technical Cooperation Monitoring and Reporting System. Approved version*, document [OP-1385-4](#), Washington, D.C.: IDB.
- .2021b. *Development Effectiveness Overview 2021*. Washington, D.C.: IDB. [GN-3059](#). Washington, D.C.: IDB.
- .2022. *Analytical Work: Development Effectiveness. Revised version*, document [GN-3118-1](#). Washington, D.C.: IDB.
- .2022. *Analytical Work: Development Effectiveness. Second revised version*, document [GN-3118-2](#), Washington, D.C.: IDB.
- .2022. *Project Completion Report. Understanding the Causes Behind the Effectiveness Ratings of Operations. Audiovisual presentation*, document [PP-1262](#), Washington, D.C.: IDB.
- .2022. *OVE's Review of Project Completion Reports and Expanded Supervision Reports: The 2022 Validation Cycle*, document [RE-575](#). Washington, D.C.: IDB.

- Inter-American Investment Corporation. 2020. [IDB Invest's Impact Management Framework: Managing a Portfolio for Impact](#): Washington, D.C.: IIC
- Kirk, C. 2019. [Independent external evaluation of EBRD's evaluation system - Main Report](#).
- OECD. 2002. *Glossary of Key Terms in Evaluation and Results Based Management*. Paris: OECD.

ANNEX I – EVALUATION DESIGN MATRIX

Area, evaluation questions, and sub-questions	How judgment will be formed to answer the evaluation questions	Sources	Methods and scope	Limitations	
Relevance:					
1. To what extent do the objectives of the DEF adequately respond to the needs identified in the diagnostic and the priorities of the IDB?					
1.a.	To what extent was there a clear diagnosis of the Bank's needs and a definition of the Bank's priorities?	<ul style="list-style-type: none"> • Identification and analysis of the needs and priorities. 	<ul style="list-style-type: none"> • OVE evaluations, Realignment documents, DEF-related documents, and other relevant reports. 	<ul style="list-style-type: none"> • Analysis of documents • Scope: relevant documentation before and after the DEF launch. 	<ul style="list-style-type: none"> • Information from documents may not be complete.
1.b.	To what extent were the objectives of DEF aligned with the needs and priorities of the Bank?	<ul style="list-style-type: none"> • Analysis of the alignment between the objectives of the DEF and the needs and priorities of the Bank. 			
2 To what extent is the DEF adequately designed to fulfill its objectives?					
2.a.	To what extent are the objectives and design of DEF instruments adequate?	<ul style="list-style-type: none"> • Assessment of the strength of the vertical logic of DEF instruments. 	<ul style="list-style-type: none"> • DEF-related documents, DEF instruments guidelines and templates, OVE evaluations. • ECG Good Practice Standards. • Information from benchmarked organizations. • Interviews with SPD. 	<ul style="list-style-type: none"> • Analysis of documents and templates. • Qualitative analysis of interviews. • Analysis and incorporation of findings from benchmarked organizations. • Scope: latest version of DEF instruments. 	<ul style="list-style-type: none"> • Standards may not be specific or applicable enough to each instrument. • Information from the benchmarked organizations may not be homogeneous. • Benchmarked organizations may have particularities that make the comparison with the IDB difficult. • Difficulties in getting information from interviewees.
2.b.	To what extent are the objectives and design of reporting mechanisms adequate?	<ul style="list-style-type: none"> • Assessment of the strength of the vertical logic of reporting mechanisms. 	<ul style="list-style-type: none"> • DEF-related documents, CRF and DEO documentation, AUG report on the DEO, and other relevant documents. 	<ul style="list-style-type: none"> • Analysis of documents • Qualitative analysis of interviews. • Analysis and incorporation of 	<ul style="list-style-type: none"> • Information from the reviewed organizations may not be homogeneous.

Area, evaluation questions, and sub-questions		How judgment will be formed to answer the evaluation questions	Sources	Methods and scope	Limitations
			<ul style="list-style-type: none"> Information from other organizations. Interviews with SPD. 	findings from reviewed organizations. <ul style="list-style-type: none"> Scope: latest version of the CRF, last three DEOs. 	<ul style="list-style-type: none"> Difficulties in getting information from interviewees.
2.c.	To what extent are DEF-related roles and responsibilities clearly and comprehensively defined, without gaps or duplications?	<ul style="list-style-type: none"> Mapping and analysis of roles and responsibilities described in corporate and/or operational regulations, guidelines, and procedures. 	<ul style="list-style-type: none"> DEF-related documents, corporate and operations regulations, and instrument guidelines. Interviews with DEF stakeholders. 	<ul style="list-style-type: none"> Analysis of documents. Qualitative analysis of interviews. Scope: latest version of corporate and operations regulations, and of instruments guidelines. 	<ul style="list-style-type: none"> Information from documents may not be complete. Difficulties in getting information from interviewees.
3.	To what extent has the DEF's design adapted to address the evolving needs and changing priorities of the IDB, as well as the challenges encountered during its implementation?				
3.a.	To what extent were evolving needs, changing priorities, and challenges identified?	<ul style="list-style-type: none"> Identification and analysis of evolving needs, priorities, and challenges. 	<ul style="list-style-type: none"> DEF-related documents, OVE evaluations, IDB institutional strategy, SPD reports, and other relevant documents. Interviews with key informants involved in the DEF design and implementation. 	<ul style="list-style-type: none"> Analysis of documents Qualitative analysis of interviews. Scope: relevant documentation after the DEF launch. 	<ul style="list-style-type: none"> Information from documents may not be complete. Difficulties getting information from interviewees.
3.b.	To what extent do the changes to the DEF align with the evolving needs and priorities?	<ul style="list-style-type: none"> Mapping of DEF changes Analysis of the alignment between DEF changes and evolving needs, priorities, and challenges. 			
Implementation:					
4.	To what extent has the DEF been implemented as expected, including quality standards?				
4.a.	To what extent have DEF instruments been implemented as expected?	<ul style="list-style-type: none"> Analysis of project teams' understanding of instrument objectives, guidelines, and templates from both their own perceptions and the viewpoint of SPD. Mapping of DEF instrument changes. 	<ul style="list-style-type: none"> Interviews and surveys with project team leaders, operations analysts, and PCR team leaders. Interviews with SPD. 	<ul style="list-style-type: none"> Qualitative analysis of interviews. Quantitative analysis of survey responses. Document review. Scope: latest version of corporate and operations regulations, 	<ul style="list-style-type: none"> Difficulties getting information from interviewees. Low survey response rate Application of quality control mechanisms may not be homogeneous.

Area, evaluation questions, and sub-questions		How judgment will be formed to answer the evaluation questions	Sources	Methods and scope	Limitations
		<ul style="list-style-type: none"> • Mapping of quality control mechanisms applied during the preparation of DEF instruments. • Analysis of project teams' and SPD's perceptions regarding the incentives that affect the implementation of DEF instruments and the effectiveness and adequacy of quality control mechanisms. • Descriptive statistics of DEF implementation variables 	<ul style="list-style-type: none"> • Corporate and operational regulations, and instrument guidelines. • Data from projects and DEF instruments. 	and of instruments guidelines.	
4.b.	To what extent have reporting mechanisms been implemented as expected?	<ul style="list-style-type: none"> • Analysis and follow-up to previous assessments of the DEO and CRF (from AUG and OVE) 	<ul style="list-style-type: none"> • Audit Report: Process to Report Information in the Development and follow-up documents. • OVE evaluation of IDB-9 commitments (2013 and 2018) • CRF and DEO documents. • Interviews with SPD and AUG. 	<ul style="list-style-type: none"> • Document review. • Qualitative analysis of interviews. • Scope: CRF 2020-2023, last three DEOs. 	<ul style="list-style-type: none"> • Follow-up documents may not be up to date. • Difficulties getting information from interviewees.
4.c.	To what extent are the roles and responsibilities, regulations, and processes understood and followed by all stakeholders?	<ul style="list-style-type: none"> • Analysis of stakeholders' perceptions regarding their comprehension, ability, and challenges encountered to fulfill DEF-related responsibilities. • Analysis of the degree to which DEF-related responsibilities are delegated, as reported by stakeholders. 	<ul style="list-style-type: none"> • Interviews and surveys with DEF stakeholders 	<ul style="list-style-type: none"> • Qualitative analysis of interviews. • Quantitative analysis of survey responses. 	<ul style="list-style-type: none"> • Difficulties getting information from interviewees. • Limited availability of stakeholders. • Low survey response rate.

Area, evaluation questions, and sub-questions	How judgment will be formed to answer the evaluation questions	Sources	Methods and scope	Limitations	
Use:					
5. To what extent are DEF products (instruments and reporting mechanisms) being used as expected?					
5.a.	To what extent have DEF instruments signaled problems?	<ul style="list-style-type: none"> • Analysis of project teams' ability and incentives to signal problems when applying DEF instruments from both their own perceptions and the viewpoint of SPD. • Analysis of the trends and correlations between the scores and ratings from DEF. 	<ul style="list-style-type: none"> • Interviews and surveys with project team leaders and operations analysts. • Interviews with SPD • Data from DEF instruments. 	<ul style="list-style-type: none"> • Qualitative analysis of interviews. • Quantitative analysis of survey responses. • Quantitative analysis of DEF instruments. • Scope: closed projects with a validated PCR. 	<ul style="list-style-type: none"> • Changes in instrument templates and guidelines may affect comparability across time. • Difficulties getting information from interviewees. • Low survey response rate.
5.b.	To what extent has DEF instruments been used for course correction?	<ul style="list-style-type: none"> • Analysis of the perceptions from project teams', Chiefs of Operations, and SPD staff on the usefulness of DEF instruments for course correction and the incentives for using them. 	<ul style="list-style-type: none"> • Interviews and surveys with project team leaders, operations analysts, Chiefs of Operations. • Interviews with SPD. 	<ul style="list-style-type: none"> • Qualitative analysis of interviews. • Quantitative analysis of survey responses. 	<ul style="list-style-type: none"> • Difficulties getting information from interviewees. • Low survey response rate.
5.c.	To what extent have the corporate reporting mechanisms been used for course correction at the corporate level?	<ul style="list-style-type: none"> • Analysis of the perceptions from SPD staff, Senior Management, and Board on the usefulness of the CRF and DEO for course correction at the corporate level and the incentives for using them. 	<ul style="list-style-type: none"> • Interviews and surveys to SPD, Senior Management, and IDB Board members. 	<ul style="list-style-type: none"> • Qualitative analysis of interviews. • Quantitative analysis of survey responses. 	<ul style="list-style-type: none"> • Difficulties getting information from interviewees. • Low survey response.
5.d.	How does the use of DEF products compare to other MDBs?	<ul style="list-style-type: none"> • Identification and analysis of the use of instruments and reporting mechanisms from other international financial institutions. • Comparison of the use of instruments and reporting mechanisms between the IDB and other international financial institutions. 	<ul style="list-style-type: none"> • Information from benchmarked organizations. • Same as questions 5.a. to 5.c. 	<ul style="list-style-type: none"> • Analysis and incorporation of findings from benchmarked organizations. • Same as questions 5.a. to 5.c. 	<ul style="list-style-type: none"> • Information from the benchmarked organizations may not be homogeneous. • Benchmarking organizations may have particularities that make the comparison with the IDB difficult. • Same as questions 5.a to 5.c.

Area, evaluation questions, and sub-questions	How judgment will be formed to answer the evaluation questions	Sources	Methods and scope	Limitations	
Results:					
6. To what extent have the DEF's expected outcomes of (i) having governance arrangements enabling the effective implementation of the DEF, (ii) improving the design, monitoring, and evaluation throughout the intervention's cycle, (iii) enhancing accountability for results, and (iv) increasing learning from past experience been achieved?					
6.a.	To what extent have governance arrangements enabled the effective implementation of the DEF?	<ul style="list-style-type: none"> Analysis of the perceptions of DEF stakeholders. 	<ul style="list-style-type: none"> Interviews and surveys to SPD, KIC, Division Chiefs, Sector Managers, Chiefs of Operations, project teams and IDB Board members. 	<ul style="list-style-type: none"> Qualitative analysis of interviews. Quantitative analysis of survey responses. 	<ul style="list-style-type: none"> Difficulties getting information from interviewees. Low survey response.
6.b.	To what extent has there been improvement in the design, monitoring, and evaluation throughout the intervention's cycle?	<ul style="list-style-type: none"> Analysis of aggregated indicators comparing the period before and after the DEF launch along with a qualitative description of the observed changes. Analysis of the data from previous evaluation questions to build a logical argument establishing that when DEF instruments are well-designed and properly implemented, improvements are inherently realized. Descriptive analysis and correlation of variables in project design, monitoring, and evaluation derived from DEF instruments. Quantitative analysis of the determinants of project performance (e.g., project cancellations). Analysis of the perceptions of project teams, Division Chiefs, Sector Managers, Chiefs of Operations, and 	<ul style="list-style-type: none"> OVE evaluations conducted prior to the DEF implementation (e.g., evaluability reviews and PCRs assessment) Interviews and surveys to SPD, Division Chiefs, Sector Managers, Chiefs of Operations, project teams. Project-level data. Same as questions 2.a., 3.b., 4.a., 5.a, 5.b. 	<ul style="list-style-type: none"> Qualitative analysis of interviews. Quantitative analysis of survey responses. Quantitative analysis of PCR ratings and their predictors. Scope: closed projects with a validated PCR. 	<ul style="list-style-type: none"> Difficulties getting information from interviewees. Low survey response. Changes in instrument templates and guidelines may affect comparability across time. Small sample sizes may affect the ability to draw definitive conclusions.

Area, evaluation questions, and sub-questions		How judgment will be formed to answer the evaluation questions	Sources	Methods and scope	Limitations
		SPD staff regarding improvements across the project cycle due to the use of DEF.			
6.c.	To what extent has the accountability for development results been enhanced?	<ul style="list-style-type: none"> • Descriptive analysis of the proportion of projects with validated development results and the gap in project performance ratings reported by Management and OVE. • Analysis of the perceptions on whether PCR scores and project performance is being used to hold project team leaders, Chiefs of Operations, Division Chiefs, Sector Managers accountable. • Analysis of the perceptions of Division Chiefs, Sector Managers, and SPD on whether the CRF and DEO have generated a culture of accountability. 	<ul style="list-style-type: none"> • Interviews and surveys to SPD, Division Chiefs, Sector Managers, Board members, and project teams. 	<ul style="list-style-type: none"> • Qualitative analysis of interviews. • Quantitative analysis of survey responses. 	<ul style="list-style-type: none"> • Difficulties getting information from interviewees. • Low survey response.
6.d.	To what extent has learning from past experience been increased?	<ul style="list-style-type: none"> • Analysis of the perceptions of project team leaders, Division Chiefs, Sector Managers, SPD, KIC, and other Bank staff on whether DEF products have informed the design of future projects. • Analysis on whether Loan Proposals and Sector Framework Documents, include lessons learned from DEF products. 	<ul style="list-style-type: none"> • Interviews and surveys to project team leaders, Division Chiefs, Sector Managers, SPD, KIC. • Loan proposals and Sector Framework Documents. 	<ul style="list-style-type: none"> • Qualitative analysis of interviews. • Quantitative analysis of survey responses. • Document review supported by AI and computational methods. • Scope: representative sample of projects approved 2020 onwards, most updated version of sector framework documents. 	<ul style="list-style-type: none"> • Difficulties getting information from interviewees. • Low survey response. • Loan proposals and Sector Framework Documents may not explicitly identify whether lessons learned come from DEF products.

ANNEX II - SELECTED FINDINGS FROM PREVIOUS OVE EVALUATIONS

In previous evaluations, OVE has assessed the DEF's instruments. The following paragraphs present selected findings that have been grouped by instrument. These findings come from OVE's IDB-9 evaluation (document [RE-515-6](#)), and OVE's Review of Project Completion Reports and Expanded Supervision Reports: The 2022 Validation Cycle (document [RE-575](#)).

1. **Findings related to the DEM:** While OVE's IDB-9 evaluation from 2018 found that most IDB-9 requirements related to the DEM had been implemented, it was unclear whether DEMs remained an accurate measure of project evaluability. The evaluation found that project teams had become adept at designing projects with high DEM scores, and, therefore, the score threshold of 5 was easily met. OVE's latest validation report still found that poor monitoring and evaluation (M&E) affected projects' effectiveness, in part due to problems with results matrix quality (e.g., insufficient indicators to measure all relevant dimensions of objectives, and outputs incorrectly identified as outcome indicators).
2. **Findings related to the PMR:** OVE's IDB-9 evaluation found that most IDB-9 requirements were met, though further refinements were needed. Most importantly, OVE found that the PMR did not monitor outcomes, as they did not require a judgment on whether a project was on track to achieve its development objectives. Rather, attention was placed on outputs and expenditures, which drove the classification of projects among the three categories of "satisfactory," "alert," or "problem." Many IDB staff interviewed by OVE at that time did not believe this classification system accurately reflected project performance. The guidelines allowed the reclassification of projects, with justification by the project team leader and approval by the country representative, and the number of such reclassifications had been growing.
3. **Findings related to the PCR:** OVE's IDB-9 evaluation found that significant progress had been made in establishing a credible and consistent objectives-based self-evaluation system but there were still challenges to address. Stronger efforts were needed at that time to ensure that PCRs were in line with the objectives-based methodology and were consistently delivered on a timely basis. Further clarity was needed on what type of economic analysis to carry out to assess efficiency for various types of operations. OVE's 2022 validation report (document [RE-575](#)) found that differences between OVE's and Management's ratings for SG operations had continued to widen. During the 2022 validation cycle, while OVE rated the overall outcome positively in 53% of validated SG operations, Management did so in 81% of them. In the 2022 and preceding three validation cycles discrepancies in ratings have been observed. In terms of quality, fewer than half of the PCRs were of "satisfactory" quality. Over half of the PCRs were missing key information necessary to assess performance, which required Management to submit additional information and OVE to revisit evidence and reassess the relevant ratings.

ANNEX III - SIMILAR EVALUATIONS FROM OTHER INTERNATIONAL FINANCIAL INSTITUTIONS

In the last decade, four ECG members (AfDB, EBRD, WBG, and IMF)²⁴ have evaluated their self-evaluation systems. These evaluations frame the objectives of the self-evaluation system (or evaluation system in the case of EBRD) in very similar terms: performance management, accountability, and learning (AfDB); institutional performance, learning, and accountability (EBRD); systematic learning, accountability, institutional effectiveness, and transparency (IMF); and operational performance management, accountability for results, and learning (WBG). Performance management, accountability and learning are the three main aspects every evaluation covers.

The following are the main findings for each of these reports. At the AfDB, the self-evaluation system and process were found to have many positive features, laying out strong standards and procedures and a cogent articulation with the independent evaluation function. The main weaknesses were in the application of the established procedures, standards and norms. The self-evaluation system at EBRD was found to be extensive and that required significant time and attention from operational staff and senior management. Among senior managers, the system was widely perceived as an overhead rather than an essential feature of organizational learning and accountability. At the World Bank Group, self-evaluation systems compliance with requirements was found to be mostly strong. The self-evaluation systems primarily focused on results reporting and accountability needs and did not provide the information necessary to help the World Bank Group transform into a “Solutions Bank” or develop learning to enhance performance. Considerable self-evaluation takes place at the IMF; many self-evaluation activities and reports were found to be of high technical quality; and that self-evaluation informed reforms in policies and operations. Yet, there were gaps in coverage, weaknesses in quality, and shortcomings in the dissemination of lessons.

The following table summarizes more specific findings, grouped by main topic:

Main topic	Findings
Candor of self-evaluation	<ul style="list-style-type: none"> Insufficient candor was highlighted by the AfDB, WBG and IMF reports. The first two reports explain that the lack of candor derives from a defective incentive structure. At the WBG, candor was limited by the fear of reputational damage: “acknowledging that a project was not performing well was described as “exposing one’s dirty laundry” and best avoided”. Similarly, at the AfDB, “the perception that project performance is equated to staff performance undermines the motivation to rate poorly-performing projects candidly. The IMF report found that although self-evaluation products (EPAs and EPEs) were generally effective tools for reflecting on experience, there was room for improvement in their candor.
Ratings	<ul style="list-style-type: none"> The WBG and AfDB reports emphasize negative effects that may arise from the rating system. The WBG report explains that an excessive focus on ratings also affects the candor of self-evaluation due to the fear of repercussions from a bad rating. For the AfDB, the report states that “the rating itself could become an obstacle to learning because it potentially makes the discussion unnecessarily contentious and personalized”. At the WBG, measuring and rating project outcomes at closing against objectives stated at design years earlier has become a source of tension and perceived rigidity, given that the quality assurance of results frameworks at the time of project design is insufficient and that the options of restructuring and adaptive project management have not taken root.

²⁴ The ERBD evaluation, out of the four, is the only one covering independent evaluation.

Main topic	Findings
Lessons learned	<ul style="list-style-type: none"> Documentation of lessons: for the AfDB, lessons were not institutionalized and documented. For the WBG the “right” lessons were not being captured in self-evaluation products. Quality of lessons: the reports from the IMF, AfDB, and WBG highlighted issues with the quality of lessons learned. Lessons were found to be either too generic or too specific to be applicable. In some cases, the evidence behind the lesson was weak. Use of lessons: at the WBG there was little effort to extract and synthesize evidence and lessons or to inform operations. Staff were more likely to rely on tacit knowledge.
Who conducts self-evaluation	<ul style="list-style-type: none"> At the WBG, many self-evaluation products are written by consultants rather than staff. The report explains that while the reasons for outsourcing are varied and legitimate, by using consultants, the World Bank’s forgoes an opportunity for contextual learning by staff and also signals the low priority placed on self-evaluation.
Incentives and tools	<ul style="list-style-type: none"> Shortcomings in self-evaluation go beyond the specific tools and instruments. The WBG report states that the main reasons for the observed shortcomings lie in incentives and behaviors rather than templates and processes. At the AfDB, the main weaknesses were not in the established procedures, standards, and norms but in their applications, particularly, the low level of compliance with established procedures, deficient candor and over optimism in assessing performance, and limited resources for M&E during supervision. The EBRD report states that revising templates and processes may be necessary but not sufficient. It explains that to reform the self-evaluation system, significant organizational change, fresh incentives, committed leadership, and organization-wide pursuit of a vision which recognizes the transformative potential of high-quality performance information when applied in a receptive organizational environment, will be required.
Incentives and use of self-evaluation	<ul style="list-style-type: none"> Reports from the WBG and the AfDB explain that even when evaluations are candid, there are no incentives to act on them. At the WBG incentives do not reward good M&E and the identification and fixing of problems and at the AfDB there is no recognition for fixing problems. At the WBG, knowledge from systems was rarely valued or used. At the EBRD, the self-evaluation system was widely perceived as an overhead rather than an essential feature of organizational learning and accountability. Some senior managers confused monitoring and evaluation and evince skepticism about the benefits of the self-evaluation process (‘too much, too late and too little learning’).
Avoidance of project restructuring	<ul style="list-style-type: none"> At the AfDB, the report found a strong tendency to avoid addressing issues through formal project restructuring because the transaction costs were considered too high. This resulted in a failure to introduce corrective measures and led to the retention of appraisal targets that were no longer in line with the project reality.

Sources: ^{a)} AfDB (2020). Evaluation of the AfDB’s Self-Evaluation Systems and Processes Summary Report, An IDEV Corporate Evaluation, April 2020.

^{b)} EBRD (2019). Independent external evaluation of EBRD’s evaluation system - Main Report, author Colin Kirk.

^{c)} IMF (2015). Self-Evaluation at the IMF: An IEO Assessment Independent Evaluation Group (2016).

^{d)} Behind the Mirror: A Report on the Self-Evaluation Systems of the World Bank Group. World Bank, Washington, DC. World Bank.


Office of Evaluation and Oversight - OVE

Established in 1999 as an independent evaluation office, OVE evaluates the performance and development effectiveness of the activities of the Inter-American Development Bank Group (IDB Group). These evaluations seek to strengthen the IDB Group through learning, accountability and transparency.

OVE evaluations are disclosed to the public in accordance with IDB Group policies to share lessons learned with the region and the development community at large.

 iadb.org/evaluation

 [linkedin.com/showcase/idb-ove](https://www.linkedin.com/showcase/idb-ove)

 [@BID_evaluacion](https://twitter.com/BID_evaluacion)