

60th GEF Council Meeting
June 14–18, 2021
Virtual Meeting

Agenda Item 12

**RESULTS BASED MANAGEMENT – EVALUATIONS OF THE AGENCY SELF-
EVALUATION SYSTEMS AND THE GEF PORTAL**

(Prepared by the Independent Evaluation Office of the GEF)

Recommended Council Decision

The Council, having considered document GEF/E/C.60/07, Results Based Management – Evaluations of the Agency Self-Evaluation Systems and the GEF Portal, and the Management Response, takes note of the related evaluation recommendations and endorses the management response to address them.

TABLE OF CONTENTS

Executive Summary	v
Findings	v
Conclusions	viii
Recommendations	ix
Introduction	1
PART A: Evaluation of Agency Self-Evaluation Systems.....	2
Background	2
Literature and Coverage by GEF Agencies	4
Key Questions	5
Methodology.....	6
Evaluation Design.....	6
Sources of Information	6
Conduct of the Evaluation	8
Findings	8
Structure of Self Evaluation Systems	8
Framework for Self-Evaluation in Agencies	9
Quality assurance.....	11
Self-Evaluation Products: Submission, Compliance, and Quality	17
PART B: Evaluation of the GEF Portal.....	26
Background	27
Key Questions and Methodology.....	28
Key questions.....	28
Methodological approach.....	28
Findings	31
Conclusion.....	41
Recommendations	42
References	43
Part A: References for Evaluation of the Agency Self Evaluation Systems	43
Part B. References for Evaluation of the GEF Portal.....	45
Annexes.....	47
Annexes - Part A. Evaluation of Agency Self Evaluation Systems	47
Annexes - Part B. Evaluation of the GEF Portal.....	79

LIST OF TABLES AND FIGURES

TABLES

Table 1: Availability of project implementation reports and mid-term reviews for completed projects	17
Table 2: Level of compliance on reporting requirements for terminal evaluations of full-sized projects	18
Table 3: Terminal evaluations rated in the satisfactory range for quality – by Agency and project type	19
Table 4: Coverage through interviews	29
Table 5: Participation figures for survey	30
Table 6: Comparison of the GEF Portal with good practices in portal design	35

FIGURES

Figure 1: Timeline of GEF Agencies joining the GEF partnership	2
Figure 2: Purposes of Self Evaluation	9
Figure 3: Coverage of Issues in Guidance Documents of GEF Agencies	10
Figure 4: Three-year rolling average of terminal evaluation quality ratings presented—based on year of terminal evaluation completion (N=1,774)	19
Figure 5: User perceptions on Portal performance, on a scale of 6 to 1	33

EXECUTIVE SUMMARY

1. As part of its ongoing work for the Seventh Comprehensive Evaluation (OPS-7) of the Global Environment Facility (GEF), the GEF Independent Evaluation Office (IEO) has undertaken several evaluations and reviews that address different aspects of the GEF results-based management (RBM) system. First of these was the Review of the GEF Terminal Evaluation Validation Process, which was presented to the GEF Council as an information document at the Council's December 2020 meeting. This report brings together evidence from two evaluations: i) The Evaluation of the Agency Self-Evaluation Systems; and, ii) The Evaluation of the GEF Portal. Another evaluation that addresses the RBM-related arrangements managed at the corporate level by the GEF Secretariat is presently underway; it will be prepared in time for inclusion in the OPS-7 and will be presented in December 2021 Council meeting.

2. This report is divided into two parts. Part A presents the Evaluation of the Agency Self-evaluation Systems and Part B presents the Evaluation of the GEF Portal.

A. Evaluation of the Agency Self Evaluation Systems

3. The Evaluation of the GEF Agency Self-Evaluation Systems aimed at assessing the extent to which the Agency self-evaluation systems meet the GEF requirements and provide credible, quality, and timely information to support accountability and learning.

4. The evaluation is based on information from several sources including a literature review; desk review of terminal evaluations, mid-term reviews and project implementation reports (PIRs); review of terminal evaluation templates, interviews; and, workshops. Five workshops with participation from the GEF Agencies were conducted to analyze the challenges related to Agency self-evaluation systems and to identify solutions through a design thinking-based participatory process.

Findings

5. The self-evaluation systems of GEF Agencies may be classified in three broad categories based on who manages terminal evaluations and whether the terminal evaluations are validated:

- (a) Managed by the project staff and validated by the Agency's evaluation unit (internal or independent unit)
- (b) Managed by the project staff but not validated by the Agency's evaluation unit or such a unit is absent
- (c) Managed by the Agency internal evaluation units.

6. Where an evaluation unit does not conduct or commission terminal evaluations but validates them, the unit spends less time on evaluations at the project level and more on drawing lessons at the portfolio level. In such situations the evaluation unit may not have as much confidence in the data quality of the terminal evaluation because the unit is a step removed from the data-gathering process. When an evaluation unit manages terminal evaluations, it ensures data quality more directly but has less time for portfolio-level and

thematic analysis. Where independent evaluation units, which report directly to the governing Board of the organization, have not yet been established, drawing lessons and learning at the portfolio and thematic levels is often difficult.

7. Policy frameworks and mechanisms in place within Agencies' self-evaluation systems support the provision of credible, quality, and timely information. Of the 18 GEF Agencies, 13 specify accountability and learning as the main purpose of self-evaluation. The Food and Agriculture Organization of the United Nations (FAO) and the Development Bank of Latin America (CAF) refer to the main purposes in similar terms emphasizing the need for a strong evidence-base for feedback into decision making. The United Nations Development Programme (UNDP) has capacity building as one of the added purposes of self-evaluation. All Agencies provide guidance on evaluation criteria and processes for terminal evaluations and quality assurance. The majority incorporate at least some elements of the GEF requirements and IEO guidance for terminal evaluations. Only a few agencies, which include the United Nations Environmental Programme (UNEP), UNDP, FAO, the United Nations Industrial Development Organization (UNIDO), the Foreign Economic Cooperation Office, Ministry of environmental Protection of China (FECO), and the World Wildlife Fund (WWF-US), has explicit guidance on mid-term reviews. The Agencies that joined the the Partnership during the second round of expansion are keen to have more guidance on conduct of mid-term reviews.

8. Arrangements in place within Agencies to harmonize their self-evaluation systems with GEF requirements vary among Agencies. Some provide a step-by-step guide, and some have mainstreamed these requirements within their standard guidelines. Development finance institutions such as the World Bank, Inter-American Development Bank (IDB), African Development Bank (AfDB), European Bank for Reconstruction and Development (EBRD), Asian Development Bank (ADB), and the International Fund for Agricultural Development (IFAD) follow their own evaluation guidelines for terminal evaluations and add GEF requirements to sections of the terminal evaluation reports or as additional material in an annex.

9. GEF Agencies use a variety of quality assurance practices. GEF Agencies, except for IFAD, have set up a GEF coordination unit. Depending on the Agency, the evaluation unit or the GEF coordination unit is responsible for quality assurance of terminal evaluations. Evaluation units of development finance institutions validate terminal evaluations. For IDB, the World Bank, and ADB these units validate terminal evaluations of only those projects that involve funding beyond a certain threshold. Evaluation units in UNEP, UNIDO, and FAO manage the evaluation process, including providing quality assurance. The GEF Coordination units of Conservation International (CI), the International Union for Conservation of Nature (IUCN), and WWF-US manage the conduct of terminal evaluations, provide feedback, and are responsible for quality assurance. Agency evaluation units generally do not validate project implementation reports and mid-term reviews. Only WWF-US and, upon exceptional request, FAO and IUCN validate mid-term reviews.

10. Ratings approaches of GEF Agencies are well aligned with the GEF IEO's approach in terms of what they aim to capture, and scales applied, but there are minor differences that affect comparability across Agencies. The diversity in Agency practices reflects an Agency's evaluation logic, prioritization of issues, and, to a minor extent, differences in interpretation

of the GEF guidelines. Even within the same Agency the same criteria and methodology may be applied with different levels of stringency at different points in time, making comparability across time periods challenging (GEF IEO 2014).

11. There are gaps in submission of project implementation reports and mid-term review reports, particularly in low-performing projects. Project implementation reports generally document challenges during implementation in a timely manner, although in some instances they may not adequately reflect the level of urgency required to meet the reported challenges. Compared to reports for projects that were rated satisfactory or highly satisfactory, project implementation reports were less likely to be available for projects that were rated unsatisfactory or highly unsatisfactory at implementation completion. Mid-term reviews are available for 43 percent of completed full-sized projects. Thus, the instrument is not being used as often as it should be.

12. GEF Agencies generally comply with the GEF IEO's terminal evaluation guidelines for full-sized projects. Compliance is good with requirements related to the timeliness of conduct of terminal evaluations; general information on the project; theory of change; reporting on outcomes; reporting on project monitoring and evaluation (M&E); consistency in performance ratings; and reporting of lessons rooted in project experience. However, compliance is weak with the requirements of reporting on application of social and environmental safeguards and on stakeholder consultations.

13. The quality of terminal evaluation reports is improving. Among the terminal evaluations that have been prepared in recent years, the percentage rated highly unsatisfactory or unsatisfactory for terminal evaluation quality is very small. The share of terminal evaluations that is rated in the satisfactory range for quality has increased, particularly those rated moderately satisfactory for quality.

14. Generally, self-evaluation systems of GEF Agencies provide support to learning on doing things right. Policies, guidance, and mechanisms are in place to ensure credible, quality, and timely information. Feedback loops are in place at the project and organization levels, and red flags in project implementation reports trigger corrective actions. The focus is on activities and outputs.

15. Self-evaluation systems place less emphasis on learning about doing the right things. In general, the development finance institutions have more robust mechanisms to generate and share insights about intervention effectiveness. They also have a tradition of assessing long-term effects and sustainability for a sample after projects have been completed for some time. There is variation among UN Agencies – some Agencies such as UNDP and IFAD have a strong emphasis on learning on doing right things, whereas some others have not given as much attention to it. Nongovernmental organizations (NGOs) have almost no systemic arrangements for supporting learning on doing the right things.

16. GEF Agencies generally do not incentivize candor in self-evaluation, though some are making efforts to encourage such an evaluation culture. The key measure of success in Agencies is project volume or deal flow. This organizational logic trickles down to project design and management. As a result, evaluation is mainly seen as a necessary requirement. Project staff are mainly interested in moving their projects along without issues and to get a good rating in the end. Due to the lack of systematized learning and exchange on “what

works,” there are no direct incentives for candor. As a result, opportunities for learning from challenges and failures of other projects may be limited.

17. Agencies are exploring approaches to incentivize candor. Some Agencies are changing the focus of performance ratings. For example, EBRD recently dropped the practice of rating performance to facilitate greater candor in, and learning from, self-evaluations. Others are reinforcing an evaluation culture based on quality and evidence. Some Agencies are investing in training project management staff. IDB, for example, has created a Development Effectiveness Unit, which support projects from design to post-evaluation.

Conclusions

The evaluation sought to answer four key questions.

- (a) How do policy frameworks in the GEF Agencies support their self-evaluation systems? The evaluation concludes that the policy frameworks of Agencies do support the self-evaluation systems. Their frameworks generally describe well how their self-evaluation systems should contribute to learning on doing things right. However, they do not adequately address how learning on doing right things should be enhanced. The self-evaluation system related frameworks of Agencies vary considerably in terms of the extent to which they address GEF requirements.
- (b) To what extent do the agency self-evaluation systems provide credible, quality, and timely information to support accountability and learning? The evaluation concludes that the self-evaluation systems generally provide credible information. The GEF Annual Performance Report 2020 (GEF IEO 2020) showed that project implementation reports generally assign overly optimistic performance ratings. The evaluation found that even though the ratings provided in the project implementation reports may be overly optimistic, their narratives adequately capture the challenges faced by the project. However, in a few instances the narratives may not reflect the level of urgency required to address challenges. Mid-term reviews are generally regarded as credible and useful, but these are not prepared for a majority of full-sized projects. Terminal evaluations are generally regarded as credible, and in general their quality is improving. Overall, the self-evaluation systems of all GEF Agencies support accountability well. However, some Agencies seem to be much better at deploying these systems for learning on doing the right things than others.
- (c) To what extent do the Agency self-evaluation systems meet the GEF requirements according to the relevant GEF policies and guidelines? Agencies self-evaluation systems generally meet the relevant GEF policies and guidelines, although there are compliance gaps in some areas. Agencies have either mainstreamed GEF requirements in their policies and guidance, or have put in place ad hoc arrangements to address the GEF requirements. They broadly use the same criteria and rating scales as used by the GEF IEO, although there are minor differences that pose challenges in comparisons across the Agencies. The mid-term reviews are mandatory for full-sized projects but are not prepared for a majority. The submission rate of project implementation reports is somewhat

lower for the low-performing projects. The quality of terminal evaluation reports is improving, although reporting on whether feedback from the Operational Focal Points has been sought and addressed and on application of safeguards, is generally weak.

- (d) What factors influencing the effectiveness of the self-evaluation systems could have an impact on the quality and timeliness of information provided to the GEF? Experienced, functionally independent evaluation units with strong capacities play an important role in promoting a robust self-evaluation culture. In operational terms Agencies that have mainstreamed GEF policies and guidance in their internal policies, guidance, and procedures, are able to prepare self-evaluation products that are compliant with GEF requirements.

Recommendations

18. The evaluation of the Agency Self-Evaluation Systems has two recommendations:

- (a) The GEF Secretariat and Agencies should strengthen the use of mid-term reviews for learning and adaptive management.
- (b) The GEF Secretariat, in collaboration with other partners, should strengthen learning through the systems that it manages, support for cross-Agency exchanges, and incentives for candor.

B. Evaluation of the GEF Portal

19. In 2018, the GEF transitioned from its Project Management Information System (PMIS) to the GEF Portal as the PMIS was increasingly unable to meet the evolving needs of the GEF partnership. The Portal is intended to provide a user-friendly online interface to submit, review, and approve project and program proposals and to store data and documents related to project approval, implementation, and results. This evaluation assesses the extent to which the GEF Portal is meeting its objectives and is an improvement over the PMIS. It presents lessons from the experience with the Portal's development and rollout.

20. Information was gathered from several sources. These include a review of relevant GEF Council documents, publications by the Secretariat, and evaluations by the GEF IEO; interviews with key informants; an online survey; and an observation-based survey of peer portals. The evaluation team also observed three workshops conducted by the GEF Portal team to train the Portal users on its new features. Along with the GEF Portal, the web pages of three peer portals including the Green Climate Fund Project Portfolio System, the International Renewable Energy Agency (IRENA) Project Navigator, and the United Nations Framework Convention on Climate Change Clean Development Mechanism (UNFCCC CDM) information system, were examined (as an external user) for comparison. .

Findings

21. The Portal has enhanced the online project proposal submission and review capabilities. The PMIS offered little support to the project proposal submission and review process as these activities were conducted offline and documents were uploaded to the

PMIS after the review. Now, Agencies submit proposals and the Secretariat conducts its reviews and takes decisions on the proposals on the Portal. The Portal creates a clear audit trail of the who, what, and when of a given action to facilitate accountability.

22. Overall, the Portal has contributed to improvement in data quality—especially of the more recent projects—through increased automation and arrangements to ensure data entry discipline, though some errors in data outputs were noted. The Portal has streamlined the process of submitting project proposals, project implementation reports, midterm reviews, and terminal evaluations, by helping the Agencies move from submission through emails to direct submission on the Portal. The Portal encourages discipline in data entry by requiring complete data for the preceding stage to move to the next stage and reduces errors through auto validation checks and data entry menu options. Concerns pertaining to historical data remain, despite efforts made to improve them, and glitches in data outputs continue to be discovered and addressed.

23. The Portal is easy to navigate, visually appealing, and accessible, and it compares well with its peers on these criteria. The Portal has a simplified professional design with a strong logic. It is easy for the user to identify an HTML link; web pages are well composed and with a clean layout. The icons used in the webpages are simple, elegant, and consistent. The evaluation assessed performance of the GEF Portal with three other peer portals that include the Green Climate Fund Project Portfolio System, IRENA Project Navigator, and UNFCCC CDM information system. The GEF Portal compares well with peer portals in criteria such as navigation, visual appeal, and accessibility. The user perception on ease of navigation and use of webpages is varied and appears to be linked with the frequency of usage—those who use it more often have a more favorable perception of it than those who don't—and based on user groups type. For example, a lower percentage of Agency staff assess performance of the Portal to be in the satisfactory range for ease in navigation and in use of webpages, than other users of the Portal.

24. The Portal is much more developed than when it was first launched, and user experience has improved; however, there are still several major gaps. Since its launch in 2018, several features have been added to the Portal. The Portal team is working on an improved reporting platform and developing a dashboard for reporting. However, there are several areas where the Portal needs to be developed further. These include development of a comprehensive system of alerts through emails, providing the ability to batch download documents, enhancing the capabilities of the search function, and ensuring that the calculations presented in the Portal data outputs and reports are correct. The Portal team links slow development of the Portal with limited resources budgeted for its development. Of the US \$ 922,229 cumulatively allocated for development of the Portal, at the start of fiscal year 2017 US \$658,650 was available for development of the Portal and at the end of fiscal year 2020 US \$ 299,000 remained unspent. These cost figures are a low-end estimate because these do not reflect full costs.

25. Users perceive a need for a transparent process to collect information on problems, prioritize problems, and report on the progress in addressing them. Although users' initial experience with the Portal was disappointing, several interviewees noted a marked change toward a more positive user experience with the Portal since mid-2019, when the ability to submit proposals was rolled out. From this point onward, they suggest that the Portal

provides at least the minimum required level of expected services. However, there is dissatisfaction among users with what they perceive as numerous low-level glitches and challenges. Users who have directly or indirectly contacted the Portal team to report glitches and request changes are generally satisfied with the responsiveness of the Portal team; however, some are unaware of the process for contacting the Portal team and/or are dissatisfied with what they perceive as a lack of clarity in the approach used to identify and prioritize problems in the Portal.

26. The Portal has contributed to the acquisition of more and better data, and access to data is more regulated than before. More data is being generated through the online processing of project proposals, reviews, uploading of documents, and entry of data on key milestones. However, availability of data to users is more regulated, which in part may be due to differences in the disclosure policies of the GEF and the GEF Agencies. Greater regulation has limited the real-time availability of data to external stakeholders and the public.

27. Connectivity is a major concern for many users. For many users, recurrent issues related to logging in, connection losses, and “silent logouts” caused by the page timing out, led to wasted effort. Earlier this problem was common across all users, but this was partially addressed when the Portal team enhanced the bandwidth available for the Portal. For many users that are based in the least developed countries and remote areas, it is difficult to use the Portal with ease because internet connectivity is a constraint for them.

Conclusions

28. The Portal is an improvement over the PMIS, and it also compares well with its peer portals on some of the technical parameters. The Portal has substantially achieved its objectives related to enhancement in project review and processing abilities, capturing of information in a consistent format, integration of GEF programming strategies and policies into the Portal, tracking of results of GEF activities, enhanced transparency, and safeguarding of confidential information. At the same time, its performance is mixed in terms of taxonomy and tagging, search and analytical abilities, and real-time availability of data to external stakeholders and the public. Other gaps in performance include a lack of ability to download batches of documents, lack of a comprehensive system of alerts on project cycle milestones, and errors in data outputs. These gaps need to be addressed urgently.

29. There is a need to give greater attention to addressing the needs of the users from the GEF Agencies, who perceive that the benefits of the Portal for them have not been commensurate with the effort that they need to put into using it. An overarching concern is slow development of the Portal: almost three years after its launch the Portal is perceived as a work in progress. The Portal team has linked slow development of the Portal to the limited resources that they have to work with. GEF Management needs to assess whether additional support is necessary and the extent to which speedier development of the Portal may be aided.

Recommendations

30. The Evaluation of the GEF Portal has two recommendations:
- (a) The GEF Secretariat should strengthen its process to address user feedback on the Portal.
 - (b) GEF Management should develop and implement a time bound plan to speed up the development of the Portal.

INTRODUCTION

31. As part of its ongoing work for the Seventh Comprehensive Evaluation (OPS-7) of the GEF, the GEF Independent Evaluation Office (IEO) has undertaken several evaluations and reviews that address different aspects of the GEF results-based management (RBM) system. This report brings together evidence from two evaluations: i) The Evaluation of the Agency Self-Evaluation Systems; and, ii) The Evaluation of the GEF Portal. The other reviews include the GEF Terminal Evaluation Validation Process, which was presented as an information document to the GEF Council during its December 2020 meeting; and the Evaluation of the GEF Results-Based Management (RBM) System – with focus on the arrangements managed by the GEF Secretariat – which is currently underway.

32. The Review of the GEF Terminal Evaluation Validation Process found that the GEF IEO's validation process was well established and the dataset from the process allows for comparisons across the GEF Partnership. It found several areas where the GEF IEO's approach to the validation process may be strengthened: greater sharing of information on the validation process with the Agencies; graduation of the Agencies that have a robust validation process in place and giving more attention to the newer Agencies. It also called for greater attention to strengthening community of practice on validation of terminal evaluations.

33. The two evaluations presented in this report further expand on this work and bring other areas relevant to RBM in focus. The Evaluation of the Agency Self-Evaluation Systems (Part A) assesses the extent to which these systems meet the GEF requirements and provide credible, quality, and timely, information to support accountability and learning. Through use of a systems and design thinking process it identifies bottlenecks and solutions.

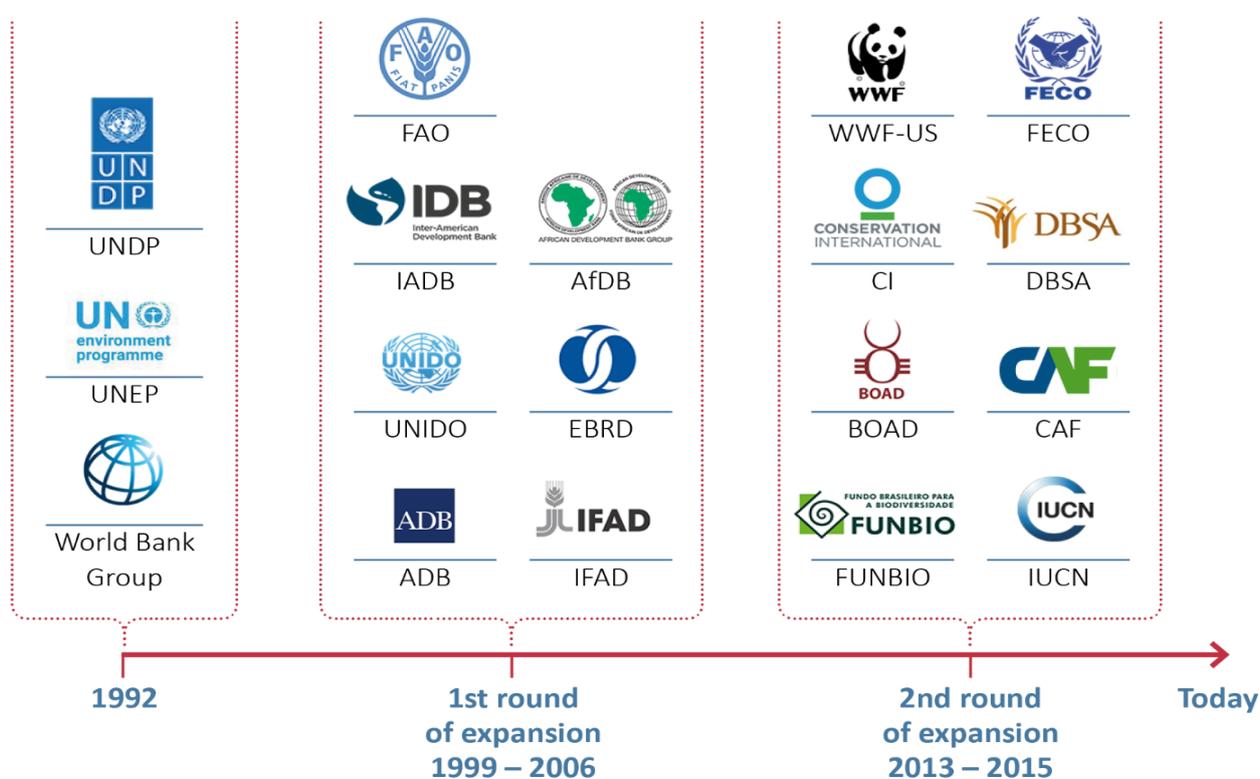
34. The Evaluation of the GEF Portal (Part B) assesses the extent to which the Portal has met its objectives to provide a user-friendly online interface to submit, review, and approve project and program proposals and to store data and documents related to project approval, implementation, and results. It identifies areas where the GEF Portal has performed well and those where there are gaps in performance.

PART A: EVALUATION OF AGENCY SELF-EVALUATION SYSTEMS

BACKGROUND

35. The GEF provides support to address global environmental concerns related to biodiversity, climate change, international waters, land degradation, the ozone layer, and persistent organic pollutants. Since its inception in 1992, the GEF has provided \$20 billion in grants to developing countries and countries with economies in transition. GEF grants are deployed by 18 accredited Agencies that joined the GEF Partnership at different points in time (figure 1). The GEF Agencies are the operational arm of the GEF. They work closely with project proponents — governments, civil society organizations, and other stakeholders — to design, develop, and implement GEF-funded projects and programs.

Figure 1: Timeline of GEF Agencies joining the GEF partnership



36. **The Agency self-evaluation systems are expected to facilitate learning and accountability across the GEF partnership.** At the project level, self-evaluations are reflective exercises that help implementors to learn from the experience. At the corporate level, these facilitate monitoring of the portfolio and may provide early warnings on trends that warrant attention, and contribute to learning through sharing of lessons. Although the GEF IEO had covered aspects of the Agency self-evaluation systems through its past work, it had not assessed the extent to which these systems were meeting the needs of the GEF Partnership and the areas where improvement was required. The Evaluation of the GEF Agency Self-Evaluation Systems is aimed at addressing this information gap.

37. The Development Assistance Committee of the Organisation for Economic Co-Operation and Development (OECD DAC) defines self-evaluation as “an evaluation by those who are entrusted with the design and delivery of a development intervention” (OECD 2002).

The definitions used across the United Nations system and across the multilateral development banks consider self-evaluation to be a management project/program management responsibility and primarily for the management's own use (UN 2018; UNEG 2017; ECG 2018). The term self-evaluation system includes components such as scope, responsibilities, methods, data, and products; they cover different stages of the project and program cycle. Both independent evaluation and self-evaluation are aimed at enhancing learning and accountability, although independent evaluation is generally regarded as more credible for accountability.

38. The expectations from the self-evaluation systems of the Agencies are outlined in several GEF policy documents and policies of the GEF Agencies. For example, the GEF Evaluation Policy (GEF IEO 2019), the Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects (GEF IEO 2017), and Minimum Fiduciary Standards for GEF Partner Agencies (GEF 2018) specify several requirements.

39. The GEF Evaluation Policy (2019) requires the Agencies to prepare terminal evaluations and mid-term reviews (where applicable), and monitor their respective GEF portfolios. The terminal evaluation guidelines for full-sized projects provide guidance on reporting on topics such as project details, stakeholder consultations, theory of change, methodology, outcome, sustainability, implementation, project monitoring and evaluation (M&E), safeguards, gender, and lessons. While the guidelines are applicable to the full-sized projects that involve more than \$2.0 million in GEF funding, the Agencies are encouraged to use these to develop terms of reference for evaluation of the GEF-funded medium-sized projects and programs.

1. The GEF Monitoring Policy (2019) addresses the guiding principles for monitoring, along with other requirements including reporting through project implementation reports, mid term reviews, and tracking tools. It requires that the Agencies submit to the Secretariat *“annual Project Implementation Reports, including information on project status; the amount of GEF Project Financing disbursed; the latest Development Objective Rating, Implementation Progress Rating, and Risk Rating...”* The Agencies are required to submit the project implementation reports no later than 75 calendar days after the end of the first full fiscal year (July 1–June 30). At project mid-term the Agencies are required to submit mid-term reviews for all full-sized projects, and for medium-sized projects where available, and tracking tools and information on achieved results across applicable core indicators and subindicators.

40. The GEF IEO undertook this evaluation to assess the extent to which Agency self-evaluation systems provide quality and timely information, comply with the GEF requirements, and facilitate learning across the GEF Partnership. Another motivation is to facilitate learning across the GEF Agencies through the evaluation process. The evaluation was conducted by the IEO and Endeava, a consulting firm with expertise in design thinking.

LITERATURE AND COVERAGE BY GEF AGENCIES

41. The literature on self-evaluation systems in international development organizations is not well developed. However, self-evaluation systems are getting increasing attention from the evaluation units of the international development organizations. The World Bank's Independent Evaluation Group (IEG) conducted the first detailed review of a GEF Agency self-evaluation system in 2016. Thereafter, the Independent Development Evaluation unit of the African Development Bank (AfDB) conducted an evaluation of its self-evaluation systems. The Independent Evaluation Department (IED) of the Asian Development Bank (ADB) recently completed a review of ADB's project level self-evaluation system (ADB IED 2020). The European Bank for Reconstruction and Development (EBRD) and the United Nations Development Programme (UNDP)¹ are also undertaking an assessment of their self-evaluation systems (EBRD 2019).

42. Several practitioners have discussed establishment of self-evaluation systems in international development organizations (Zall Kusek and Rist 2004; Bester 2012). Among the international development organizations, the stated purpose of self-evaluation is generally to foster learning and accountability (World Bank IEG 2016; ADB IED 2020); which is the same as the purpose of independent evaluation.

43. Self-evaluations serve many purposes for which independent evaluation may not be as well suited. Self-evaluations are useful in situations where decisions are urgent and require close synchronization (Picciotto, 1999). Self-evaluation also provides practitioners opportunities for conversion of tacit knowledge into explicit knowledge (Spender 1996; Nonaka 1994). Taut's (2007) 'action researcher' takes advantage of the rich information gained by being an insider and generates knowledge that facilitates adaptive management. This approach also allows for rapid feedback to others on lessons that may be applicable in other contexts with similar challenges (Taut 2007). Picciotto (1999) distinguishes self-evaluation from independent evaluation by noting that the former aims primarily at assisting decision makers, whereas the latter focuses primarily on accountability.

44. One of the major challenges with self-evaluation is the issue of credibility. Scriven (1975) argues that a self-evaluation has less credibility because of the perceived conflict of interest. This may be understood within the framework of the agent-principal problem as agents may lack incentives for candor (Ross 1973; Arrow 1984; and Grossman and Hart, 1992). Scriven (1975) notes that measures such as issuance of guidelines on conducting self-evaluations, use of checklists, and standardization of evaluation criteria and practices may enhance credibility. He argues that independent evaluation of at least some of the activities along with self-evaluations will enhance the credibility of the latter. Among international development organizations, validation of the terminal evaluations by the evaluation units is a commonly used approach to address credibility-related challenge.

45. In the evaluations conducted by the evaluation units of the GEF Agencies, there is a recognition that the self-evaluation systems are not being used for learning to the extent these should be. The evaluation of the World Bank's self-evaluation systems concluded that these systems complement the independent evaluation systems, but the focus of the former is more on results reporting and meeting the accountability needs and less on learning to enhance performance (World Bank IEG 2016). The review of ADB's project-level self-evaluation system concluded that the system is robust and credible and is useful for

¹ Communications with UNDP Independent Evaluation Office.

both learning and accountability (ADB IED 2020). However, it also assessed that reliability of gathered information is a concern and concluded that much of the tacit knowledge that is internal to the ADB staff remains untapped. Both evaluations stress the importance of incentives in ensuring utility of self evaluations and argue that without appropriate incentives it is likely that self-evaluation would become a bureaucratic requirement and less useful for learning (World Bank IEG 2016; ADB IED 2020).

KEY QUESTIONS

46. The evaluation seeks to answer the following questions:

- (a) **How do policy frameworks in the GEF Agencies support their self-evaluation systems?** The evaluation assesses how self-evaluation is addressed by Agencies through their policies. The assumption is that an enabling policy framework will lead to sound arrangements for self-evaluation, which will then lead to good quality self-evaluations. The evaluation assessed the extent to which policies explain the purpose and role of self-evaluations, provide guidance on how the self-evaluations ought to be conducted, and clarify relationship with independent evaluation.
- (b) **To what extent do the agency self-evaluation systems provide credible, quality and timely information to support accountability and learning?** The evaluation considers the arrangements within the Agencies to conduct self-evaluations (or equivalent²). It considers arrangements to address quality assurance, harmonization, information management and knowledge sharing, as they relate to self-evaluation of the GEF projects. The evaluation assesses the extent to which these arrangements allow for credible self-evaluation of GEF supported activities. The evaluation also assesses the opportunities and solutions for improving the self-evaluation systems.
- (c) **To what extent do the Agency self-evaluation systems meet the GEF requirements according to the relevant GEF policies and guidelines?** The evaluation assesses the quality of self evaluations and the extent to which GEF Agencies are meeting the relevant GEF requirements. These include the minimum requirements related to project M&E stipulated by the GEF Evaluation Policy (2019) and its predecessor the GEF M&E Policy (2010), the Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects (2017), and the GEF Monitoring Policy (2019).
- (d) **What are the factors that influence the effectiveness of the self-evaluation systems which could have an impact on the quality and timeliness of information provided to the GEF?** The evaluation assesses whether and how different variables such as policy frameworks, information management arrangements, incentives to promote candid reporting, quality assurance arrangements, level of resources provided, and capacities of the local partners-- affect the evaluation of self-evaluation systems.

² The term equivalent is used for Agencies where the evaluation unit may be involved in conducting evaluations such as terminal evaluation or mid-term reviews that are generally conducted by those involved in implementation of the activities.

METHODOLOGY

Evaluation Design

47. The evaluation incorporates the principles of design thinking in analyzing the factors that affect self-evaluation system performance and identifying potential solutions. The evaluation process included a participatory approach and captured the shared experience of the GEF Agencies through workshops wherein all GEF Agencies were invited to participate.

48. The evaluation used a multiple-case design and covers all the GEF Agencies (Yin 2018). The self-evaluation system of a GEF Agency – as it relates to the GEF supported activities – is the primary unit of analysis. Within each Agency information on the self-evaluation system was gathered through review of relevant policy documents of an Agency, interviews with the unit responsible for RBM and with senior management. The products of the self-evaluation system, e.g. project implementation reports, mid-term reviews, and terminal evaluations, were reviewed to determine the areas where these reports are strong and where these may be improved. In these analyses each product is treated as a unit of analysis, and – where possible – the performance on a specific dimension is aggregated and reported per Agency. The coverage for products is through a representative sample. Terminal evaluation templates used by GEF Agencies were reviewed to assess reporting practices.

49. **Systems thinking** was applied to the analysis to understand the interconnectedness and interdependencies that exist in self-evaluation systems across multiple levels. The evaluation also used systems thinking to develop an understanding of the dynamics that contribute to or inhibit effectiveness of the self-evaluation systems at different levels.

50. **Design thinking** was a key element of the evaluation approach. Design thinking considers the users' perspectives and needs. The evaluation sought collaboration with Agencies in an iterative process with continual input from them to identify opportunities for improvement and co-create solutions. This process also built community among the Agencies to facilitate cross-learning on self-evaluation system related experiences.

Sources of Information

51. **Literature Review:** The evaluation drew from the literature relevant to self-evaluation systems. Information gathered through the literature review was used in development of the evaluation design and approach. Reports prepared by United Nations Evaluation Group, Evaluation Cooperation Group, Multilateral Organisation Performance Assessment Network, and Joint Inspection Unit, that cover at least some aspects of self-evaluation in GEF Agencies were also reviewed. A more detailed discussion on the literature is provided in the approach paper of this evaluation (GEF IEO 2020a).

52. **Interviews:** Sixty-four interviews were conducted with staff from 16 Agencies including GEF coordinators of the Agencies and staff from their evaluation units.

53. **Workshops:** Five workshops were conducted. The first two of these workshops were conducted to identify opportunities for improvement. Representatives of 16 Agencies participated in at least one of these two workshops. Three design thinking workshops were held, with a total of 48 participants from 14 Agencies, to identify concrete solutions to improve self-evaluation systems at the level of Agencies and within the GEF partnership.

54. **Desk Reviews:** Several desk reviews were conducted as part of the evaluation. These include a review to assess Agency policies and guidelines relevant to self-evaluation; a review to assess compliance with the GEF IEO guidance on the preparation of terminal evaluations; a review of project implementation reports; and an assessment of availability of mid-term reviews.

55. The review of the Agency policies and guidelines relevant to self-evaluation assessed how the policy frameworks in the GEF Agencies support their self-evaluation systems. Of the 18 GEF Agencies, the review analyzed policy and guidance documents of 14 Agencies to assess their definition of self-evaluation system and how GEF requirements are reflected in these documents. Four Agencies did not provide relevant documents. Of the 18 Agencies, terminal evaluation templates for 12 Agencies – that had submitted terminal evaluations to the GEF IEO in the preceding four years – were analyzed.

56. The review of terminal evaluations to assess compliance with the GEF IEO guidelines covered 270 completed projects. The GEF IEO issued Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects in April 2017. The terminal evaluation guidelines (GEF IEO 2017) specifies reporting requirements on topics such as project identification, basic project details, the theory of change, feedback from stakeholders, methodology, outcome, sustainability, implementation, project M&E, environmental and social safeguards, gender concerns, lessons, and recommendations. All terminal evaluations that were prepared in October 2017 or later and submitted by the Agencies to the GEF IEO by August 2020 were covered. Although the guidelines are not applicable to the medium size projects and programs, the Agencies are encouraged to use these as inspiration to prepare terminal evaluations for these as well. Terminal evaluations for both FSPs and MSPs are covered in the review.

57. The review of project implementation reports assessed submission gaps, and consistency and timeliness in reporting. A stratified random sample of 150 completed projects approved from GEF-3 onwards was drawn. This includes three samples of 50 projects that represent different outcome rating categories: highly satisfactory and satisfactory; moderately satisfactory and moderately unsatisfactory; and, unsatisfactory and highly unsatisfactory. The ratings used are the validated outcome ratings provided by the GEF IEO or the evaluation units of the GEF Agencies.

58. The GEF Monitoring Policy (2019), along with the GEF Monitoring and Evaluation Policy (2010) require that the Agencies prepare and submit a mid-term review for full size projects at the mid-term of its implementation period, and submit a mid-term review for medium size projects where such review has been conducted. The evaluation assessed gaps in submission of these reviews for the completed full-sized projects and level of availability for the completed medium size projects. The assessment covered the same set of 150 projects (95 full size and 55 medium size projects) that were covered by this evaluation to assess project implementation report related submission gaps.

59. **Datasets:** The terminal evaluation review dataset maintained by the GEF IEO was used. The dataset contains validated ratings on performance dimensions such as project outcome, sustainability, M&E, implementation, execution, and quality of terminal evaluation. The dataset available as on February 2021 was used for analysis on trends in quality of terminal evaluation.

Conduct of the Evaluation

60. The evaluation was led by the GEF IEO. The IEO designed the overall approach of the evaluation. IEO conducted the review of the Agency compliance with the GEF IEO's terminal evaluation guidelines; trends in quality of terminal evaluations; and, review of reporting through project implementation reports. Endeava conducted the systems and design thinking workshops, interviews to assess Agency self evaluation systems, and review of the terminal evaluation templates.

61. The approach to the evaluation was developed during FY2020. The draft approach paper was shared across the GEF Partnership including the GEF Secretariat and Agencies. Two independent peer reviewers provided inputs on the draft approach paper. The feedback received from the peer reviewers, GEF Secretariat and the Agencies, was incorporated in finalizing the approach paper.

62. The evaluation was conducted during the period September 2020 to April 2020. Workshops were conducted from November 2020 to February 2021. The draft report of the evaluation was shared with the peer reviewers, the GEF Secretariat and Agencies, and their feedback has been addressed in finalization of the report.

FINDINGS

Structure of Self Evaluation Systems

63. The structure of the self-evaluation systems of GEF Agencies has variations. These systems differ primarily based on who commissions and conducts terminal evaluations. In most GEF Agencies terminal evaluations are commissioned and conducted by those that are involved – directly or indirectly – in project management. Of these Agencies, some such as ADB, AfDB, EBRD, IADB, IFAD, UNDP and World Bank, have a structurally independent evaluation unit, whose head reports directly to its governing board and is independent of the management. These independent evaluation units validate at least some of the terminal evaluations prepared by management. For example, evaluation units of ADB, IADB, and World Bank, validate terminal evaluations of those projects that involve funding beyond a threshold. In IFAD the evaluation unit provides feedback on the draft of the terminal evaluation reports for a few projects, whereas another unit which is not independent of the management conducts validations. UNDP's evaluation unit conducts validation for terminal evaluations of all GEF projects.

64. Another set of Agencies where management conducts terminal evaluations are those that do not have an evaluation unit. These include CI, WWF, and IUCN, that became part of the GEF Partnership during the second round of expansion. CI, WWF-US and IUCN have oversight arrangements to ensure that terminal evaluations are impartial, but these arrangements are not independent of policy and program departments. Without evaluation units, these Agencies do not tend to conduct meta, thematic or country reviews using self-evaluation results.

65. Evaluation units of UNEP, UNIDO, FAO, FECO, DBSA and CAF are responsible for conducting terminal evaluations. However, the evaluation units of these Agencies are not structurally independent of management. For example, the head of the evaluation unit of UNEP reports to the Executive Director of UNEP, and those of UNIDO and FAO to their respective Director General. The peer reviews of the evaluation function of UNEP, UNIDO, and FAO, have assessed their evaluation units to be functionally independent (MoFA

Denmark 2010; UNEG 2012a; UNEG 2012b). The evaluation units of FECO, DBSA and CAF have not undergone the peer review process.

66. The role of internal evaluation units in preparing terminal evaluations may involve a trade-off between learning more about individual projects and being able to conduct meta evaluations, and in-depth thematic and country/region portfolio reviews to facilitate learning at the corporate level. When internal evaluation units are not involved in conduct of terminal evaluations, they are able to dedicate more resources and time to learning focused activities at corporate level.

67. In addition to the evaluation units, there are other centralized structures within Agencies to facilitate self-evaluation. For example, several Agencies have distinct monitoring functions to coordinate work on project implementation report and mid-term reviews, reporting on core-indicators, and through tracking tools. Other centralized structures may also be involved in conducting and/or facilitate self-evaluations at the project and corporate levels. For example, the Development Impact Evaluation group at World Bank provides technical support to the projects that incorporate experimental approaches to assess impact of development interventions and provide lessons on what works and in what contexts. Similarly, the Quality Assurance Group (QAG) at World Bank, which operated till FY2010, provided early feedback on quality of project design and on quality of supervision to the management. See Annex A.1 for details on practice glimpses and A.2 for characteristics of self-evaluation systems of GEF Agencies.

Framework for Self-Evaluation in Agencies

68. **The policy frameworks of most GEF Agencies require self-evaluation of implemented activities and specify accountability and learning as the two main objectives.** Accountability is associated with finding out, documenting and sharing the actual results/performance of projects. The purpose of accountability is to enhance trust and credibility (figure 1). The purpose of learning is to enhance decision making and project performance within project and across projects. Yet, how the twin objectives of accountability and learning are to be met by the self-evaluation systems is usually not clearly defined. Survey of the policy documents of the GEF Agencies revealed the following:

- (a) Thirteen Agencies specify accountability and learning as the main purpose of self-evaluation. FAO and CAF refer to the main purposes in similar terms emphasizing the need for a strong evidence-base and for feedback into decision-making.
- (b) Only UNDP has capacity building as one of the added purposes of self-evaluation.
- (c) IUCN, UNIDO, ERDB, IADB and the World Bank include a definition of the primary users along with the purpose. Management and beneficiaries (and clients for DFIs) are identified as the main target audience of self-evaluation, followed by project and/or operation staff and development partners, with board, and the public being least mentioned.

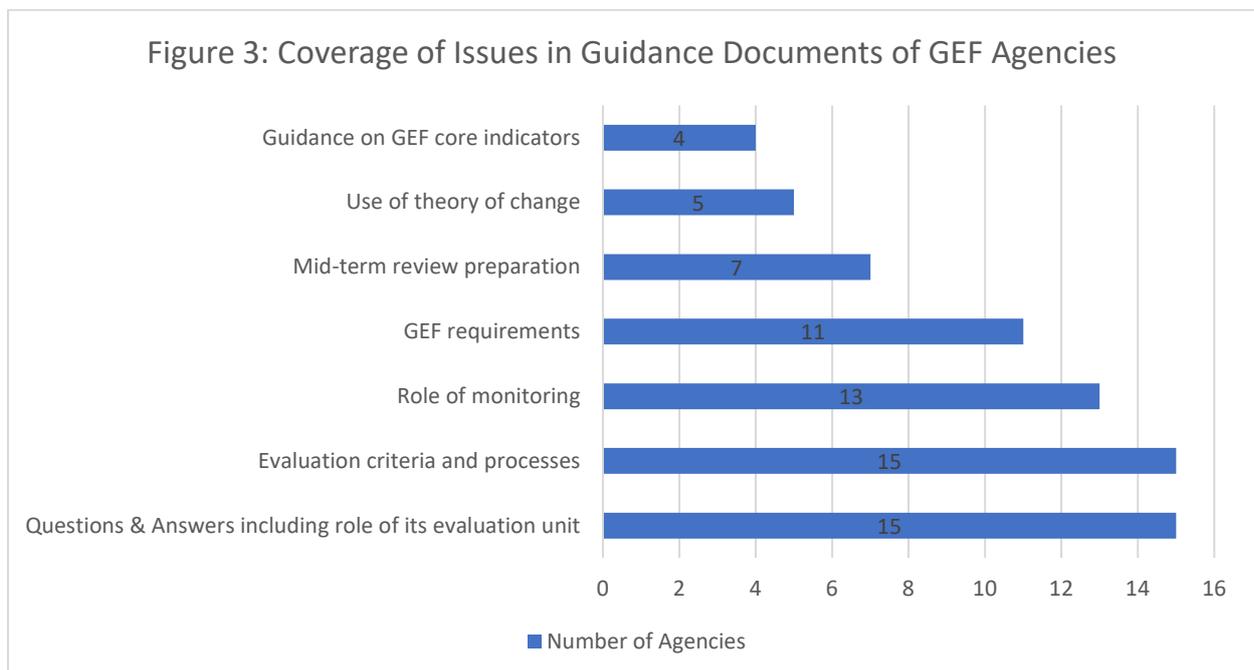
Figure 2: Purposes of Self Evaluation



Note: Evaluation purposes (green) linked with further objectives (blue). Bubble size represents word count in policy sections on M&E purpose. Data obtained from Agency policy documents.

69. All Agencies provide guidance on evaluation criteria and processes as well as on quality assurance (figure 2). The majority incorporate at least some elements of the GEF requirements and IEO guidance for terminal evaluations in their respective guidance. Compared to guidance for terminal evaluations, fewer Agencies, UNEP, UNDP, FAO, UNIDO, FECO, WWF-US and CI have explicit guidance on mid term reviews. This difference is because mid-term reviews are mandatory only for GEF supported programs and full-sized projects; in addition, the GEF Secretariat has not provided guidance on its preparation. During interviews GEF Agencies, especially those that joined the Partnership during the second round of expansion, were keen on having more guidance from the GEF Secretariat on conduct of mid-term reviews.

Figure 3: Coverage of Issues in Guidance Documents of GEF Agencies



70. **Guidance on how to harmonize the self-evaluation systems with GEF requirements varies among Agencies.** ADB, IADB, FAO, UNIDO, and FECO provide a step-by-step guide. CSOs and other UN Agencies except for IFAD have mainstreamed these requirements within their standard guidelines. Development finance institutions such as the World Bank, IADB, AfDB, EBRD, ADB and IFAD follow their own evaluation guidelines for terminal evaluations and add GEF requirements to sections of the terminal evaluation reports or as additional material in the annex. New Agencies such as WWF-US, CI and FECO have established self-evaluation procedures for GEF projects to meet requirements. CAF and DBSA are currently reviewing their policies and guidance for GEF project and programs. EBRD is currently redesigning its self-evaluation system altogether.

71. Most Agencies provide generic guidance on tools and data capture for evaluation. The World Bank, FAO, IFAD, UNEP and UNIDO provide guidance on the use of a theory of change in self-evaluation. UNEP has mainstreamed the GEF requirements in the ToRs for terminal evaluations and guidance notes on assessment of stakeholder consultations and tool kit on assessment of safeguards.

72. The GEF core indicators are specified in the guidance provided by UNDP, CAF, FAO and FECO. Some international development finance institutions face challenges in incorporating GEF core indicators, since they are not aligned with the Agencies' own results frameworks and GEF projects are usually add-ons to their larger conventional (baseline) development projects. Agencies may also monitor indicators that respond more directly to their mandate. For example, IFAD monitors several indicators for rural poverty impact in all its projects.

73. **Some Agencies are not clear about GEF requirements and about the good practices in implementing them.** Although GEF IEO has issued guidelines for conduct of terminal evaluations for full size projects, these have not been issued for medium-sized projects and for programs. Lack of GEF guidance on conduct of mid-term reviews also contributes to the perceived lack of clarity on GEF requirements. The staff of Agencies that joined the GEF Partnership during the second round of expansion and are yet to submit a terminal evaluation, i.e. BOAD, CAF, DBSA, FECO and FUNBIO, are less conversant with the GEF requirements related to terminal evaluations and need more support (GEF IEO 2020).

74. Agencies expressed an appetite for learning more on use of theory of change in design, implementation, and evaluation of GEF supported activities. They stressed the need to share good practice examples and opportunities for experience sharing on this topic. The Scientific and Technical Advisory Panel of the GEF has developed a primer on use of theory of change to support the design of GEF interventions (STAP 2019). The primer covers different stages of project cycle – design, implementation, at and after completion – and discusses how the approach may be applied and be useful. Of the GEF Agencies, five have their own guidance on theory of change may be applied in evaluations. It is still too early to empirically assess the extent to which the primer prepared by STAP has been effective. However, based on the present practice across the Partnership the evaluation finds that the theory of change approach is applied inconsistently and with different rigor because Agencies differ in their interpretation and implementation approaches.

Quality assurance

75. Quality assurance relates to arrangements in place to ensure that self-evaluation products provide credible and quality information to the GEF. GEF Agencies use a variety of

quality assurance practices. GEF coordination units support all self-evaluation procedures and provide support for quality assurance. All GEF Agencies, except IFAD, have set up such a coordinating unit. In many cases, Agencies have developed a step-by-step guide for the project teams, so that they follow GEF requirements. GEF coordination units also review reports before submitting them to the GEF through the GEF Portal.

76. Depending on the Agency, the evaluation unit or the GEF coordination unit plays a key role in quality assurance of terminal evaluations. Independent evaluation units of development finance institutions validate terminal evaluations although their approaches differ somewhat. ADB's evaluation unit validates terminal evaluations for all the GEF projects. The evaluation units of IADB and WBG only validate terminal evaluation above a certain funding threshold. In IFAD the evaluation unit provides feedback on the drafts of a few terminal evaluations and validations are conducted by another unit that is not independent of the management. UNDP's evaluation unit validates terminal evaluations for all GEF projects.

77. Evaluation units in UNEP, UNIDO and FAO, manage the evaluation process and provide quality assurance. Of these, the GEF IEO accepts the project performance ratings provided by UNEP's evaluation unit – for UNIDO and FAO the GEF IEO conducts terminal evaluation validations. In FECO, DBSA and CAF terminal evaluation quality related oversight and feedback is provided by the GEF coordination units. FAO is in the process of shifting from its present system of its evaluation unit conducting terminal evaluations to its project management structure taking over the responsibility of conducting terminal evaluations.

78. The respective GEF coordination units of CI, IUCN, and WWF-US manage the conduct of terminal evaluations. The coordination units provide feedback and are responsible for quality assurance. Altogether these Agencies have submitted five terminal evaluations so far

79. **Many Agencies have internal or external peer reviews of the terminal evaluation reports.** AfDB and World Bank undertake internal reviews and clearance processes for terminal evaluations that involve several review loops among different units and evaluation unit (see Box 1). EBRD's independent evaluation unit pre-reviews the terminal evaluations before the final submission and provides feedback to the team to ensure quality (Box 2 on ERBD). The independent evaluation units at IADB and ADB prepare a draft validation report that is reviewed internally by their independent evaluation unit staff and shared for comments with the project teams. Similarly, at World Bank, a terminal evaluation is reviewed by an experienced reviewer, and then further reviewed by a coordinator or manager in the evaluation unit. The validation report is then sent for comments to the regional practice group director and to the recipient country. Some GEF Agencies (FAO, UNDP, UNEP, UNIDO, FECO) recommend external reviews. Often this involves a reference group with internal and external stakeholders for the validation of preliminary findings and overall quality assurance of terminal evaluations (see Box 3 on FAO). UNDP requires the commissioning unit to present preliminary findings to the GEF Operational Focal Points in recipient countries.

Box 1: AfDB's triple loop review process

Project completion reports at AfDB go through a three-stage review process before being submitted to the Independent Development Evaluation (IDEV) unit. Firstly, a first draft of the report is submitted to the sector manager for clearance. Secondly, the draft is sent to peer

reviewers who should either be based in a field office or have previous field experience. Thirdly, the document is sent to the country or regional team for final clearance. At the end of the fiscal year, a sample of project completion reports is selected and sent to the bank's IDEV for validation and quality assessment. IDEV prepares its evaluation note based on all available documents and the discussions from the review process, as explained above. Additionally, the evaluation note itself is subject to peer review to ensure consistency in applying all guidelines and will be distributed to the operational department for a response who then can either agree or disagree with the findings. In case of the latter, the department will have to provide additional evidence to support their response.

This review process allows for extensive dialogue throughout the preparation and validation of the project completion report. While discussions among teams happen, especially on ratings, the overall process is perceived as highly useful as it helps the bank to come out with an independent, consensual, and reasonable assessment of a project's performance and additionally reducing bias stemming from the self-assessment of projects.

Sources: Interview with Guirane N'Diaye, Ayanleh Daher Aden, Hedi, Manai, Lafeta, Camera;
Documents: AfDB (2018)

Box 2: EBRD's IEU pre-views TEs before submission

To improve the quality of TEs, the independent evaluation office of EBRD (EvD) has introduced a feedback loop before validation. Project teams hand in a first draft. EvD reviews the draft by checking the quality and completeness e.g., whether outputs, outcomes, impacts, and lessons are reported on appropriately. Based on the feedback from EvD, operational teams are able to make adjustments before the final submission. An evaluator has 5 days to review the draft and the operational teams have another 3-5 days to incorporate the feedback and submission. Only after the final submission, the report undergoes validation by EvD. Information gaps or lack of evidence to support certain rating outcomes can be corrected. This review greatly improves the quality of reports, and ultimately also improves quality of terminal evaluation ratings.

Sources: Interview with Barry Kodolkin, EBRD (2019), EBRD (2018)

80. Some evaluation units conduct post completion project evaluations and/or impact evaluations for a set of projects. The evaluation units of the World Bank and ADB undertake field-based verifications for a sample of projects post-project closure (see Box 4 on World Bank). Four development finance institutions (WBG, ADB, IADB, AfDB) and four UN Agencies (UNEP, IFAD, FAO, UNIDO) conduct impact and thematic evaluations that may cover a sample of projects at implementation completion and/or post completion.

81. Project implementation reports and mid-term reviews are generally not reviewed by evaluation units. These reports are regarded as reports that provide real time information to the project management on implementation progress. Only WWF-US and, upon exceptional request, FAO and IUCN review mid-term assessments. FAO evaluation unit used to review all mid-term reviews up to 2018, when it shifted these responsibilities to their project management as part of decentralized evaluations. AfDB is piloting a process to review mid-term reviews. The evaluation units of the GEF Agencies do not conduct quality assurance activities for project implementation reports – GEF coordination units are generally involved in quality assurance of these reports.

Box 3: FAO's reference groups improve quality of, and learning from, evaluations

FAO's evaluation unit provides quality assurance for evaluations. In addition, it has developed a stakeholder approach to improve quality of reports as well as assess feasibility of recommendations and dissemination of lessons learned. A 'reference group' is formed early on in a TE process, at first with internal stakeholders and expanding to external stakeholders such as Operational Focal Points (OFP) and national counterparts at a later stage. The reference group support the design of the TE, review of ToR, data collection, reviews of various draft stages and advice on the feasibility of recommendations. Preliminary findings are then shared with internal and external stakeholders for validation insights and to pave a way towards the adoption of lessons learned.

Sources: Interviews with Rachel Bedouin, Team Lead Office of Evaluation (OED), Jenin Assaf, Evaluation Officer OED, Lavinia Monforte Evaluation Manager OED; FAO (2015)

Box 4: The World Bank Group conducts post project evaluations

The Independent Evaluation Group (IEG) of the World Bank is responsible for evaluating the overall development effectiveness of the World Bank Group. As part of its project level evaluations, IEG has a post project evaluation instrument, called Project Performance Assessment Report (PPAR). The PPAR is a field-based evaluation conducted for 20-25% of bank lending operations. Over the years, the World Bank Group has shifted the focus of PPARs from ratings, which are now placed into the appendix, to making the report more learning oriented. In doing so, the report should provide advice and feedback on what worked, what didn't work, and why. IEG is reviewing its approach to PPARs to further strengthen its learning related elements.

The PPAR is conducted a few years after project closure, when the effects of a certain intervention should be visible. The PPAR relies on a mixed method approach including direct evidence collection, site visits, and stakeholder consultation. The projects for which the PPAR is conducted are selected based on certain criteria that can be, but are not limited to, the existence of evidence gaps, alignment with IEG work program priorities (e.g., thematic evaluations), or the potential for learning from innovative projects. Usually, IEG seeks to examine clusters of similar projects to increase the operational significance for project teams and the learning potential of PPARs

Sources: Stephen Hutton (IEG); IEG WBG (2014), IEG WBG (n.d.)

82. GEF Agencies are expected to complete terminal evaluations within six months of implementation completion and submit the terminal evaluations to the GEF IEO within a year of implementation completion. In addition to their usual program management systems, in some Agencies GEF coordination units also trigger timely preparation of terminal evaluations. Most Agencies do not feel constrained by the time available for preparation and submission of terminal evaluations. ADB has an internal timeline for terminal evaluations that exceeds one year, and hence finds it difficult to meet the GEF deadline.

83. Timely submission of terminal evaluations may also be a challenge where GEF funding is a small part of a larger project and where the components funded by the GEF are completed at a different point in time than other components. For example, GEF funded activities are often a small component of a much larger project implemented by a development finance institution. For such projects, terminal evaluations are conducted at the end of the larger project, even if the GEF funded components were implemented much earlier. As a result, the deadline for submitting terminal evaluations within one year may not be met. This is also the case in the World Bank and UNDP – where GEF IEO accepts the terminal evaluation validations prepared by their respective independent evaluation units. The GEF IEO reports in submission gaps only when a submission is delayed for several years.

Approaches to Rating Project Performance

84. **Rating approaches of GEF Agencies are well aligned with the GEF IEO's approach in terms of what they aim to capture, and the scales applied. However, there are minor differences that affect comparability among Agencies' ratings.** Although Agencies are using the same criteria, there are differences in what is addressed within a criterion and how it is applied. The diversity in practices reflects an Agency's evaluation logic, prioritization of issues, and to a minor extent differences in interpretation of the GEF guidelines. Even within the same Agency the same criteria and methodology may be applied with different levels of stringency at different points in time making comparability across time periods challenging (GEF IEO 2014).

85. **Relevance:** The GEF defines relevance as the extent to which the intervention design and intended results are consistent with local and national environmental priorities and policies and to the GEF's strategic priorities and objectives, and remain suited to the conditions of the context, over time. Most of the Agencies address relevance to the local and national context. Main differences are:

- (a) Development finance institutions do not assess relevance of GEF projects using a GEF perspective. For example, EBRD only briefly addresses the projects relevance to the local context but does not rate the project's relevance. Further, ADB and EBRD do not address the project's relevance to the GEF.
- (b) There are some 'mandate' related interpretations of the criterion. For example, UNEP includes a section on gender and marginalized groups. UNIDO assesses and rates project ownership and relevance together.

86. **Effectiveness:** The GEF defines effectiveness as the extent to which the intervention achieved, or expects to achieve, results (outputs, outcomes, and impacts, including global environmental benefits). All Agencies assess effectiveness considering the extent to which the actual project outcomes are commensurate with those expected as defined in the project document. However, there are some differences in the sub-components assessed as part of effectiveness between Agencies. For example:

- (a) UNEP explicitly uses theory of change as part of the assessment of effectiveness. Other Agencies use theory of change implicitly and apply it inconsistently (GEF IEO, 2020b).
- (b) IFAD regards effectiveness as project performance in terms of its outcomes and outputs. It also measures several indicators of rural poverty impact. However, these are considered as indicators of project impact and not project effectiveness.

87. Most agencies rate the overall effectiveness of the project except EBRD, which rates the achievement of each project component separately but does not indicate an overall effectiveness rating.

88. Other minor differences refer to the achievement of Global Environmental Objectives as part of the WBG and FAO effectiveness assessment. UNIDO includes project management as one component of effectiveness.

89. **Efficiency:** The GEF defines efficiency as the extent to which the intervention achieved value for resources, by converting inputs (funds, personnel, expertise, equipment, etc.) to results in the most timely and cost effective way possible, compared to alternatives.

Most Agencies focus on financial resources and/or time when assessing efficiency. There are only minor differences in efficiency, which include:

- (a) UNEP includes a section on the use of pre-existing resources.
- (b) FAO includes a section on the monitoring and evaluation of the project.
- (c) IFAD includes project management.
- (d) Most Agencies rate the overall efficiency of the project except for WWF-US

90. **Sustainability:** The GEF defines sustainability the “*continuation/likely continuation of positive effects from the intervention after it has come to an end, and its potential for scale-up and/or replication; interventions need to be environmentally as well as institutionally, financially, politically, culturally and socially sustainable*” (GEF IEO 2019, pg. 13). Agencies broadly follow this guidance, with some differences:

- (a) UNEP does not consider environmental risks, and EBRD, CI, FAO and IUCN do not explicitly address the four aspects.
- (b) The WBG do not rate or assess sustainability because in their assessment there is not enough evidence at project closure to report meaningfully on this criterion. Instead, they assess risk to development outcomes, which is also the approach used by the GEF IEO to assess likelihood of sustainability.

91. There are minor differences how Agencies assign the overall outcome rating. According to GEF requirements, the overall outcome rating should include the three criteria – effectiveness, efficiency, and relevance. Sustainability should not be part of the overall outcome rating (GEF, 2017). The GEF IEO provides an overall outcome rating after considering relevance, effectiveness, and efficiency, of outcomes. Of the three criteria, relevance and effectiveness are considered critical – without ratings in these two criteria being in the satisfactory range, the overall outcome rating may not be in the satisfactory range. The GEF IEO adopted this approach for the outcome rating so that the process for arriving at the overall outcome rating is transparent. There are minor differences in the GEF IEO’s approach and the approaches followed by the GEF Agencies:

- (a) ADB includes sustainability in its overall outcome rating.
- (b) WWF includes achievement of project objectives in its overall outcome rating.
- (c) UNEP, IFAD, and IUCN do not provide an overall outcome rating as defined by the GEF.

92. Most Agencies use the rating scales following the GEF requirements. GEF IEO uses a six-point rating scale, from highly satisfactory to highly unsatisfactory to assess effectiveness, efficiency, and relevance. Sustainability is rated on a four-point rating scale ranging from likely to unlikely (GEF, 2017). Some exceptions are notable. For the rating scale for sustainability, most Agencies are aligned with the GEF using a four-point rating scale ranging from likely to unlikely EBRD and UNEP uses the same six-point rating scale as is used for effectiveness, efficiency, and relevance. ADB uses its own four-point rating system to assess sustainability.

Self-Evaluation Products: Submission, Compliance, and Quality

93. **Project implementation reports document challenges faced during project implementation and identify corrective actions; however, there are gaps in submission of these reports, particularly in low-performing projects.** Project implementation reports provide an account of project implementation progress on an annual basis. This study assessed the extent to which these reports document challenges during project implementation. The analysis showed that these reports generally document challenges during implementation in a timely manner, although in some instances they may not adequately reflect the level of urgency required to meet the challenge. Overall, this suggests that overall the project implementation reports are providing information that may be useful for corrective actions. However, the study found that compared to reports for projects that were rated satisfactory or highly satisfactory, project implementation reports were less likely to be available for projects that were rated unsatisfactory or highly unsatisfactory at implementation completion (table 1). It also showed that the mid-term reviews are being prepared for only a limited number of full-sized projects. The percentage is much lower for medium-sized projects, because Agencies are encouraged – but not required – to prepare mid-term reviews for these projects. Overall, the mid-term review – though noted as a useful instrument for learning by the Agencies – is not being used as often as it should be.

Table 1: Availability of project implementation reports and mid-term reviews for completed projects

Project Information Available	Projects Rated U-HU (n = 50)	Projects Rated MS-MU (n = 50)	Projects Rated HS-S (n = 50)
At least one PIRs submitted for a project	88%	90%	98%
A PIR submitted for each year a project was under implementation ³	62%	68%	80%
Projects with a mid -term review	26%	36%	38%
<i>Full size Projects with a mid-term review *</i>	<i>36% (28)</i>	<i>44% (32)</i>	<i>49% (35)</i>
<i>Medium size projects with a mid-term review*</i>	<i>14% (22)</i>	<i>22% (18)</i>	<i>13% (15)</i>

Source: review of project implementation reports carried out as part of this evaluation and Annual Performance Report 2020. *Number of observations in parentheses.

94. **GEF Agencies generally comply with the GEF IEO’s terminal evaluation guidelines for full-sized projects.** Compliance with the terminal evaluation guidelines was assessed for terminal evaluations that had been prepared in or after October 2017, six months after issuance of the guidelines. A total of 270 terminal evaluations, of which 185 were for full-sized projects and 85 for medium-sized projects, were covered. Agencies generally comply with requirements related to the timeliness of conduct of terminal evaluations; general information on project; theory of change; reporting on outcomes; reporting on project M&E; consistency in performance ratings; and reporting of lessons rooted in project experience. However, compliance with the requirements related to reporting on application of social and environmental safeguards and stakeholder consultation in preparation of the terminal evaluation is weak (table 2). Similar patterns of compliance were observed in terminal evaluations of medium-sized projects.

95. **Review of terminal evaluations prepared and submitted by Agencies from October 2017 onward shows that all Agencies have gaps in compliance with the terminal evaluation guidelines for full-sized projects.** However, the extent to which they comply and the areas where they have gaps differ. Some gaps are minor. For example, the terminal evaluations prepared by the World Bank

³ Gaps in the first year of implementation are not considered as projects may not complete PIRs for this year.

generally do not provide the GEF project ID number but it is easy to ascertain this based on the World Bank's project ID number which is linked with the GEF project ID. Other gaps are more substantial. For example, most terminal evaluations prepared by Agencies do not report on whether the evaluation team sought and incorporated inputs from the respective GEF Operational Focal Points during preparation of the terminal evaluation. As a result, it is difficult to ascertain the extent to which the GEF Agencies are complying with the Minimum Requirement 4 – regarding engagement of the Operational Focal Point in terminal evaluation preparation. Feedback from other stakeholders of the project is sought and reported on in most of the terminal evaluations. Overall, terminal evaluations prepared by the Food and Agriculture Organization of the United Nations (FAO), UNDP, the United Nations Environment Programme (UNEP), the United Nations Industrial Development Organization (UNIDO), and the World Bank were generally compliant in most of the dimensions on which compliance is assessed. As a group, the Agencies that joined during the second round of expansion appear to comply less with the guidelines. Similarly, the Agencies that have prepared only a few terminal evaluations since October 2017 as a group appear to be less compliant. Some Agencies, for example UNEP, have recently updated their instruments to mainstream GEF requirements in their guidance related to reporting on stakeholder consultation and safeguards. However, it will take time before the results of this update of guidance will be evident.

Table 2: Level of compliance on reporting requirements for terminal evaluations of full-sized projects

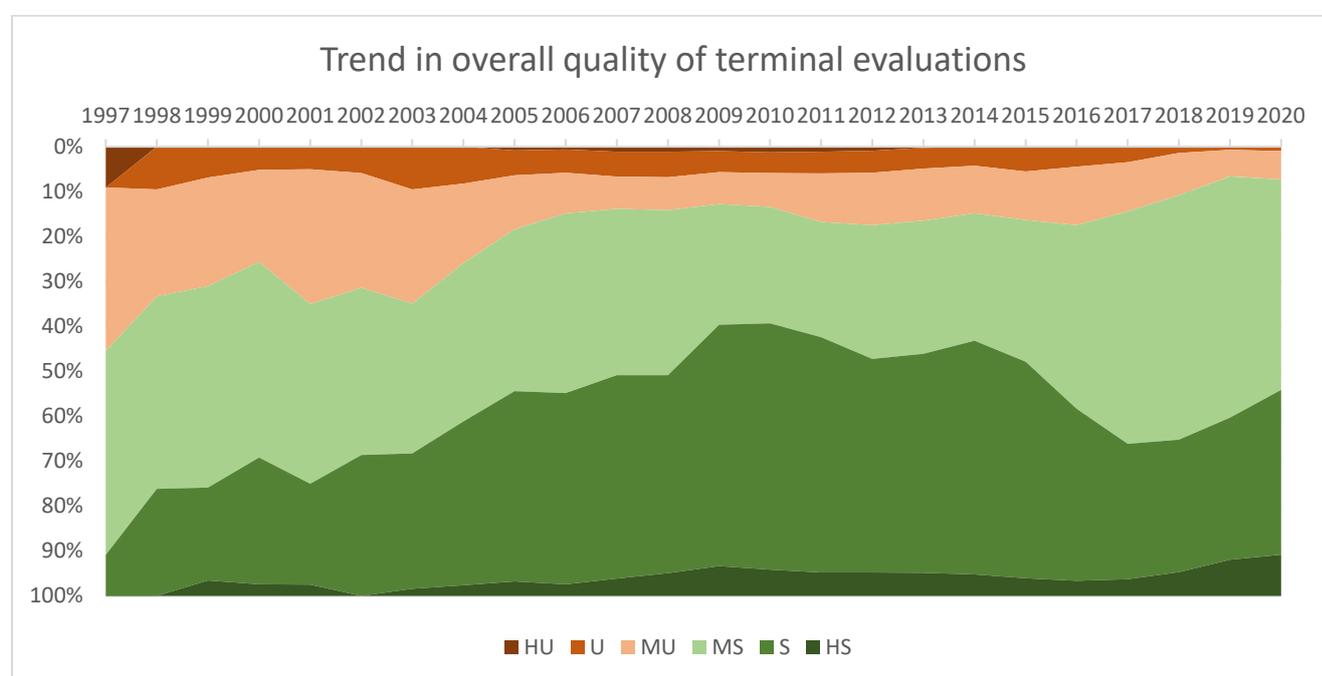
Compliance dimensions	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
Terminal evaluations reviewed	19	73	25	15	33	20	185
Timeliness of conduct	√√	√√√√	√√	√√√	√√√	√√	√√√
General information	√√√√	√√√√	√√√√	√√√√	√√√	√√√√	√√√√
OFP Feedback (national projects)	√	√√	√√	√	√√	√	√√
Feedback of other Stakeholders	√√√	√√√	√√√√	√√√	√√√	√√√	√√√
Theory of change	√√√	√√	√√√√	√√√	√√√	√√√	√√√
Sources of info. (methodology)	√√√√	√√√√	√√√√	√√√√	√√	√√√	√√√
Reporting on Outcome	√√√	√√√	√√√√	√√√	√√√√	√√√	√√√
Reporting on Sustainability	√√√	√√√√	√√√√	√√√√	√√	√√	√√√
Reporting on Project M&E	√√√	√√√√	√√√√	√√√√	√√√√	√√	√√√
Reporting on Agency performance	√√	√√√√	√√	√√√	√√√√	√√√	√√√
Co-financing	√√√	√√√	√√√	√√	√√	√√	√√√
Social and environ. safeguards	√√	√√	√√	√	√√√√	√√	√√
Lessons and recommendations	√√√	√√√	√√√	√√√	√√√	√√√	√√√
Provide performance ratings	√√√√	√√√√	√√√	√√√√	√√	√√	√√√
Consistency of ratings	√√√	√√√	√√√√	√√√	√√√√	√√√	√√√

Source: review carried out as part of this evaluation and APR2020. OFP= GEF Operational Focal Point.

Note: √√√√ = fully compliant, √√√ = mostly compliant; √√ = partially compliant; √ = noncompliant.

96. **The quality of terminal evaluation reports is improving.** Analysis of the terminal evaluation quality ratings provided through the terminal evaluation validation process of the GEF IEO and Agency evaluation units shows that there has been a steady improvement in the quality of terminal evaluations (figure 3). Very few terminal evaluations are rated highly unsatisfactory or unsatisfactory for quality. The share of terminal evaluations that are rated in the satisfactory range for quality has increased, particularly those rated as moderately satisfactory for quality. Among the GEF Agencies, compared to the period before 2010, terminal evaluations prepared by UNEP have improved in quality (table 3). Overall, the terminal evaluations for full-sized projects and medium-sized projects are of similar quality. The quality of terminal evaluations of World Bank–implemented medium-sized projects was assessed to be somewhat lower than for its full-sized projects. The latter have the advantage of being validated by its evaluation unit.

Figure 4: Three-year rolling average of terminal evaluation quality ratings presented—based on year of terminal evaluation completion (N=1,774)



Source: GEF IEO Terminal Evaluation Review Data 2021.

Note: HS= Highly Satisfactory; S=Satisfactory; MS= Moderately Satisfactory; MU= Moderately Unsatisfactory; U = Unsatisfactory; and, HU = Highly Unsatisfactory.

Table 3: Terminal evaluations rated in the satisfactory range for quality – by Agency and project type

GEF Agency	Full-Sized Projects			Medium-Sized Projects		
	Through 2009	2010 Onwards	All Years	Through 2009	2010 Onwards	All Years
ADB	—	93 (15)	93 (15)	0 (3)	—	0 (3)
FAO	—	96 (25)	96 (25)	—	82 (11)	82 (11)
IDB	—	70 (17)	70 (17)	—	60 (5)	60 (5)
IFAD	—	80 (20)	80 (20)	—	75 (4)	75 (4)
UNDP	80 (116)	85 (363)	84 (479)	83 (71)	83 (248)	83 (319)
UNEP	81 (21)	97 (75)	94 (96)	84 (43)	97 (77)	93 (120)

UNIDO	100 (1)	97 (30)	97 (31)		85 (33)	85 (33)
World Bank	88 (136)	89 (265)	88 (401)	76 (79)	66 (44)	72 (123)
Joint/others	75 (12)	88 (51)	86 (63)	100 (1)	90 (10)	91 (11)
All Agencies	84 (286)	88 (861)	87 (1,147)	79 (197)	84 (432)	83 (629)

Source: GEF IEO Terminal Evaluation Review Data 2021.

Note: Satisfactory range includes Highly Satisfactory, Satisfactory, and Moderately Satisfactory rating. Number of observations are given in the parentheses.

Self-Evaluation Systems and Learning

97. The types of learning taking place in Agencies may be broadly classified as: learning about doing things right and learning about doing the right things. Doing things right relates to whether projects are being implemented as planned, while doing the right things relates to the effectiveness of interventions and understanding what interventions work, how, and why.

98. **In general, self-evaluation systems provide support to learning on doing things right.** Policies, guidance, and mechanisms are in place to ensure credible, quality, and timely information. Agency evaluation units play an important role in ensuring the quality of terminal evaluations. In addition, most Agencies also have internal and/or external peer review mechanisms in place. For project implementation reports and mid-term reviews, GEF coordination units within Agencies ensure quality and credibility. Feedback loops are in place at the project and organization levels. At the project level, red flags in project implementation reports trigger corrective actions. The focus is on activities and outputs. Mid-term reviews also consider outcomes (to the extent possible) and review the project logic to provide recommendations. Mid-term reviews are an opportunity to adjust the approach of the project midcourse; however, their timing is mostly driven to comply with requirements, which is not always ideal for enabling adaptive management. Terminal evaluations help with management learning because they are synthesized by most Agencies per year to review the project portfolio performance against defined objectives and to provide recommendations. These annual performance reports typically also reflect on the self-evaluation system itself and identify areas for improvement. NGOs do not have these arrangements for systematic reporting at the portfolio level.

99. Self-evaluation systems place less emphasis on learning on doing the right things. Development finance institutions have the most advanced mechanisms in place to generate and share insights about intervention effectiveness. Some of them (ADB, AfDB, EBRD, International Fund for Agricultural Development [IFAD], and the World Bank) have or are currently reviewing their self-evaluation systems to get more insights about what works. Other structures, such as the Development Impact Evaluation group at the World Bank, which works with project teams to assess impacts of development interventions, may also be useful for learning on doing right things. There is variation among UN Agencies – some Agencies such as UNDP and IFAD have strong emphasis on learning on doing right things, whereas some others have not given so much attention to it. NGOs have almost no systemic arrangements for supporting learning on doing the right things.

100. Terminal evaluations are not used adequately to support learning on doing the right things. Although terminal evaluations are intended to fulfil this purpose, the format is not

adequate for this type of learning. Information provided in the “lessons learnt” section is often too generic or contextual to be meaningful for others. Documents are too long and inaccessible to be useful for others. Information gathered through interviews and workshops suggests that mid-term reviews have some potential to facilitate learning on doing the right things. However, these are not conducted for the majority of projects. Project implementation reports play only a minor role in facilitating learning on doing things right. These implementation reports are prepared annually as a routine product to provide feedback on whether things are on right track.

101. Evaluation units of Agencies support learning on doing the right things by synthesizing data from self-evaluation products and complementing it with field verifications and data. Cross-cutting analyses of self-evaluation results via meta reviews, thematic reviews, and country reviews are important Agency tools for generating insights from multiple projects on a specific area of intervention or approach. Terminal evaluations are the main source of information, but mid-term reviews are also used. Independent evaluation units of development finance institutions and UN agencies have these reviews as an important part of their mandate. Impact evaluations and technical reviews also help with learning “what works,” because they include outcome and impact data collected after project end.

102. Databases of lessons learnt are available but are reported by the Agencies to be hard to search. The Evaluation of the Knowledge Management conducted by the GEF IEO (2020) and the Evaluation of the GEF Portal presented in Part B of this report identified gaps in the search-related abilities of the Portal and makes a case for upgrading its capabilities for knowledge sharing and learning.

103. Learning practices play an important role in learning on doing right things, because these practices create experiences for staff to update their assumptions and mental models on “what works.” Stakeholder workshops and face-to-face debriefs between project staff, evaluation staff, and national counterparts are common practice at UNIDO, FAO, UNDP, and the Foreign Economic Cooperation Office in the Ministry of Environmental Protection of China (FECO). The World Wildlife Fund (WWF-US) run regular learning workshops to update of the project’s theory of change (box 5). IDB organizes an end-of-the-year open house where all the evaluations that have been produced throughout the year are presented. ADB hosts events to discuss evaluations based on “what works and what doesn’t.” At EBRD, the independent evaluation unit produces a continuous stream of webinars to share best practices. Some Agencies like ADB, IDB, and UNIDO have brown bags/green bags to share information on good practices and failures from projects.

104. **Learning across the GEF Partnership is a major lever for progress but self-evaluation system products are not adequately leveraged for cross-Agency learning.** The main mechanisms for cross-Agency learning are thematic reviews by GEF IEO and the Scientific and Technical Advisory Panel (STAP). These reviews are valued highly by Agencies. IW-Learn is recognized as highly effective in enabling learning across the Partnership. Yet, Agencies called for more opportunities to learn directly from each other’s project experiences and insights. This would also create a strong motivation for them to improve their own self-evaluation products. Similarly, the GEF Partnership currently does not offer a repository to share good practice examples, theory of change examples, templates, and project documentation. Agencies suggested that such a repository may be of great value for mutual learning.

Box5: WWF-US uses Theory of Change for learning

The World Wildlife Fund (WWF) network has a long history of applying conceptual diagrams and results chains, encouraged by WWF's Project and Programme Management Standards, based on the Conservation Measures Partnership Conservation Standards. It is encouraged that the situation diagram and result-chain diagram are developed with the software program called, Miradi. Miradi helps project managers go through the steps of the WWF Programme Standards and can help to produce key outputs such as situation analysis diagrams, results chains (theory of change), action plans, and monitoring plans.

In the context of GEF projects, two results-based management specialists serving an advisory role to WWF-US and the GEF support all GEF projects, as needed, with the development or review of situation diagrams and results chains during the planning stages of a project. Those two specialists also support GEF project teams to update the theory of change during annual reflection and learning workshops. The development of situation diagrams and result chains has helped WWF with good project design, the selection of appropriate indicators for the result framework, and learning practices within projects.

Sources: Interview with Amelia Kissick (Results-based Management Specialist, WWF-US)

105. Candor is an important factor that facilitates learning, but it is rarely incentivized. The key measure of success in Agencies is project volume or deal flow. This organizational logic trickles down to project design and management. As a result, evaluation is mainly seen as a necessary requirement. Project staff are mainly interested in moving their projects along without issues and in getting a good rating in the end. Because of the lack of systematized learning and exchange on "what works," there are no direct incentives for candor. Difficulties and failures in projects tend to be hidden rather than used as an opportunity for learning.

106. Some Agencies are changing the focus of performance ratings; for example, ERBD recently dropped the practice of rating performance to facilitate greater candor in, and learning from, self-evaluations. Some Agencies are reinforcing an evaluation culture based on quality and evidence. For example, the validation process of the evaluation units of most development finance institutions is about applying an impartial perspective to assess whether available evidence justifies the ratings. The re-rating ensures harmonized ratings across various terminal evaluations before submission to the GEF IEO, encouraging project teams' greater candor in reporting rather than a focus on good ratings.

107. Other Agencies are investing in training project management staff. IDB, for example, has created a Development Effectiveness Unit, which supports projects from design to post-evaluation. Investment in capacity building and skill development in quality evaluation have helped to build an evaluation culture among project and program staff. This means creating an environment in which inquiry, evidence, and learning are valued as essential to good management, but not limited to its use mainly to reporting. UNDP evaluation unit has developed an evaluation certification training course which is mandatory for all UNDP staff whose work involves M&E.

108. Development Effectiveness Unit. Its focus is on ensuring that evaluation results are used to inform country strategies and project cycles. Such a strong link between evaluation results and design encourages a learning from evaluation culture. ADB, the World Bank Group, and UNDP, have instituted evaluation awards to give recognition to teams and departments that have excelled in producing high-quality self-evaluation reports (box 6). The example of collaborative preparation of project implementation reports at UNDP also illustrate that the right incentive in management can promote an evaluation culture (box 7).

Box 6: The Asian Development Bank's excellence in evaluation award

The Independent Evaluation Department (IED) of the Asian Development Bank (ADB) introduced the annual IED Awards in 2014. They recognize ADB teams and departments that have excelled in producing high-quality self-evaluation reports. The candidates are drawn from the terminal evaluations that were validated by IED. The assessed criteria include adherence to evaluation guidelines, adequacy of evidence and analysis, and quality of lessons and recommendations. Awards are given out based on five categories including best reports for sovereign operations, best reports for nonsovereign operations, and the best operation department on self-evaluation. For the ceremony, all ADB staff, board, management, and consultants are invited. The award is given out by IED management, preceded by a justification for the nomination of the reports that is read out. The winners are given a trophy, plaques, and certificates. IED extensively promotes them on ADB's internal communication platform as well as on social media. Additionally, an award ceremony video is produced and disseminated both internally and externally.

The awards have increasingly gained traction within ADB. This can be seen in the number of ceremony attendees that has been constantly increasing over the past years, reaching more than 100 staff and consultants who attended the 2020 virtual ceremony. Among the participants were the board including the Vice Presidents and Director Generals of most operational departments. The awards have positively affected the overall quality of terminal evaluations within ADB. This has been accompanied by an increased interest and inquiries from operations departments on exemplary self-evaluation reports.

Sources: Interview with Garrett Kilroy.

Box 7: Collaborative Project Implementation Review at the United Nations Development Programme improves candor

The United Nations Development Programme (UNDP) puts a strong emphasis on stakeholder-led approaches in evaluation. One mechanism to achieve this is its collaborative reporting and rating of Project Implementation Review (PIR). The reporting for the PIR in UNDP is conducted in collaboration between the GEF technical advisor, the Program Manager, the Country Office Focal Point, the GEF Operational Focal Point, and sometimes also with the Regional Manager. The collaborative reporting and rating of the PIR helps UNDP ensure coherence and coordination within UNDP's decentralized model of self-evaluation. It requires extensive dialogue throughout the preparation and rating of the PIR. More evidence can be built on progress through this collaboration. It also enables a consensual assessment of a project's performance and issues. The overall effect of a collaborative process is greater candor in reporting. This is substantiated by the GEF IEO Annual Performance Report, which notes that 'for UNDP-implemented projects, either country officers or regional technical advisors or both tend to raise implementation challenges in PIRs with clarity, forcefully, and in good time, more often than the respective project managers' (GEF 2020b).

Sources: GEF IEO (2020); Interview with Pascale Bonzom.

Conclusions

109. The evaluation sought to answer four key questions. The findings help in answering these questions substantially.

110. **How do policy frameworks in the GEF Agencies support their self-evaluation systems?** The evaluation found that the policy frameworks of Agencies do support the self-evaluation systems. Their frameworks generally describe well how their self-evaluation systems should contribute to learning on doing things right. However, they do not adequately address how learning on doing right things should be enhanced. The self-evaluation system–related frameworks of Agencies vary considerably in terms of the extent to which they address GEF requirements.

111. **To what extent do the agency self-evaluation systems provide credible, quality, and timely information to support accountability and learning?** The evaluation found that

the self-evaluation systems generally provide credible information. Annual Performance Report 2020 (GEF IEO 2020) showed that project implementation reports generally give overly optimistic performance ratings. This evaluation finds that even though the ratings provided in the project implementation reports may be overly optimistic, their narratives adequately capture the challenges faced by the project. However, in a few instances the narratives may not reflect the level of urgency required to tackle a challenge. Mid-term reviews are generally regarded as credible and useful but these are not prepared for a majority of full-sized projects. Terminal evaluations are generally regarded as credible, and in general their quality is improving. Overall, the self-evaluation systems of all GEF Agencies support accountability well. However, some Agencies seem to be much better than others at deploying these systems for learning on doing right things.

112. To what extent do the Agency self-evaluation systems meet the GEF requirements according to the relevant GEF policies and guidelines? Agencies' self-evaluation systems generally meet the relevant GEF policies and guidelines, although there are compliance gaps in some areas. Agencies have either mainstreamed GEF requirements in their policies and guidance or have put in place ad hoc arrangements to address the GEF requirements. They broadly use the same criteria and rating scales as used by the GEF IEO, though there are minor differences that pose challenges in comparisons across the Agencies. The mid-term reviews are mandatory for full-sized projects but are not prepared for a majority of the others. The submission rate of project implementation reports is somewhat lower for the low-performing projects. The quality of terminal evaluation reports is improving, though reporting on whether feedback from the Operational Focal Points has been sought and addressed, and on application of safeguards, is generally weak.

113. Even though mid-term reviews are required for full-sized projects, guidance provided to the Agencies on how these should be conducted is inadequate. This is a barrier particularly for the newer Agencies in the GEF Partnership. Increasing reliance on programmatic approaches – especially impact programs – to deliver GEF support, makes the gaps in guidance for conduct of mid-term review even more salient. These programs are more likely to benefit from mid-term reviews, given their long duration, greater complexity, and high level of GEF funding. Guidance on how such reviews ought to be conducted and the issues these should address, along with outreach to the Agencies, will make mid-term reviews an effective instrument and will probably increase its usage. With the switch to the GEF Portal during GEF-7, it should be easier to track submission gaps. However, the GEF Secretariat may also need to follow up with the Agencies. Similarly, GEF IEO will need to work more proactively with the Agencies in providing them feedback on the quality of terminal evaluation reports and on compliance gaps.

114. What factors influencing the effectiveness of the self-evaluation systems could have an impact on the quality and timeliness of information provided to the GEF? Experienced, functionally independent evaluation units with strong capacities play an important role in promoting a self-evaluation culture. In operational terms, Agencies that have mainstreamed GEF policies and guidance in their internal policies, guidance, and procedures are able to prepare self evaluation products that are compliant with GEF requirements.

115. Some Agencies such as the World Bank, WWF-US, ADB, and UNDP have experimented with approaches that may facilitate learning. Their experiences may be useful

for other Agencies and may be supported by the GEF Secretariat and the GEF IEO through interagency workshops and other instruments for knowledge sharing.

116. A low level of candor in self-evaluations is a major barrier to learning both within and across Agencies. Project managers are under pressure to get good ratings and to have many projects. This presents a barrier to risk taking and learning. Platforms for sharing of project experiences and learning seem to incentivize candor in reporting and may help in harmonization. There is scope for the GEF Secretariat and GEF IEO to play a role in creating and strengthening such platforms, because these issues may be difficult to resolve at the intra-Agency level.

Recommendations

117. The evaluation has two recommendations:

- (a) **The GEF Secretariat and Agencies should strengthen use of mid-term reviews for learning and adaptive management.** The evaluation shows that despite their potential, mid-term reviews are conducted in a limited number of instances and the guidance on mid-term reviews is inadequate. The Secretariat should provide more guidance to the Agencies on conduct of the mid-term reviews, should share good practice examples, and should track timely conduct and submission of mid-term reviews. The Agencies should conduct the mid-term reviews for GEF-supported projects, as mandated by the GEF Monitoring Policy (2019).
- (b) **The GEF Secretariat, in collaboration with other partners, should strengthen learning through the systems that it manages, support for cross-Agency exchanges, and incentives for candor.** The Secretariat needs to play a greater role in facilitating learning across the GEF Partnership. Inter-Agency meetings and extended constituency workshops may be used to strengthen peer exchange on self-evaluation-related topics such as use of theory of change, and design and implementation of monitoring and evaluation plans. Similarly, enhancement of search and analysis capabilities of the Portal may be useful in strengthening learning across the Partnership. The Secretariat may also need to rope in Agencies, GEF IEO, and/or the STAP, based on the specific knowledge management challenge that needs to be addressed. For example, GEF Agencies that have experimented with incentives to enhance candor may be encouraged to share their experiences; similarly, the STAP may be drawn upon for use of theory of change, and GEF IEO on guidance on mid-term reviews.

PART B: EVALUATION OF THE GEF PORTAL

118. The Global Environment Facility (GEF) Portal is intended to provide *“a user-friendly on-line interface to allow direct entry and review and approval of projects and programs proposed for funding by the GEF”* and *“to store data and documents related to their implementation”* (GEF Secretariat 2018, pg. 1). The GEF transitioned from its Project Management Information System (PMIS) to the Portal in 2018. The shift was driven by the need to upgrade the GEF’s system as the PMIS was increasingly unable to meet the needs of the GEF Partnership (GEF IEO 2017b). The Evaluation of the GEF Portal is an input to the Seventh Comprehensive Evaluation (OPS-7) of the GEF. The OPS-7 approach paper specifically notes that OPS-7 will cover *“the redesigned results portal”* (GEF IEO 2020, pg.10).

119. Concerns related to PMIS had been noted in several evaluations. OPS-2 noted concerns related to data on co-financing (GEF IEO 2002). OPS-3 noted that the PMIS did not capture information systematically and did not facilitate monitoring of the GEF portfolio, and that the GEF Secretariat and monitoring and evaluation staff do not find its data to be fully reliable (GEF IEO 2005). OPS-4 noted some improvements in the PMIS but also highlighted weaknesses such as a low level of automation and poor quality of historical data (GEF IEO 2010). It observed that the pace of improvement in the PMIS has been slow because of the cost-cutting measures adopted by the Secretariat and that the choice of delayed development was not optimal. OPS-6 noted that the availability and quality of information provided by the PMIS remain an area of concern (GEF IEO 2018). It further noted that the primary reason for low quality of data is that they are being entered manually and not updated regularly. It highlighted the need for improving systems for project management data, monitoring, and knowledge sharing.

120. In 2010–11 Deloitte & Touche LLP was engaged by the GEF Secretariat to conduct an Independent Review of the GEF Systems (GEF Secretariat 2011). The review identified 23 critical issues; and the majority of these pertained to the weaknesses in the information technology systems and databases used to record and manage the GEF project pipeline and to reporting and data sharing across the GEF Partnership. Based on the analysis prepared by Deloitte & Touche LLP, the GEF Secretariat prepared an information document listing options for strengthening the GEF systems (GEF Secretariat 2012).

121. The GEF then decided to be a part of the GEF Trustee’s Financial Intermediary Funds (FIF) Information Technology (IT) Project. After two years, the GEF dropped out of the effort. It then engaged with the World Bank-ITS to develop an entirely new system to replace PMIS. The efforts begun in 2017 and the Portal was launched at the start of GEF-7 in July 2018. The design of the Portal and procedures and rules regarding its use are aimed at delivering the following benefits (GEF Secretariat 2018):

- (a) Enhanced project review and processing abilities to improve operational efficiency.
- (b) Improved accessibility of data through capturing of information in a consistent format, use of taxonomy to tag projects based on their characteristics, entry of georeferenced information, and improved search and analytical abilities.
- (c) Integration of key GEF programming strategies and policy requirements in the Portal to facilitate submission and review of project and program proposals;

tracking of actual results against targeted results; and a dashboard view of status and trends across the GEF project portfolio.

- (d) Enhanced transparency through increased real-time availability of data to external stakeholders and the public and safeguarding confidential information.

122. At its launch the GEF Portal had limited capabilities. Thereafter, several features have been rolled out; several features are under development. This evaluation assesses the extent to which the GEF portal is meeting its objectives, and the extent to which it is performing as per the ex-ante expectations. It documents lessons from the experience of the Portal's development and rollout. This draft report of the evaluation is being shared across the GEF Partnership for comments and feedback.

BACKGROUND

123. The PMIS was the predecessor of the GEF Portal. The PMIS started in 1998, when the GEF Council requested a list of projects the GEF had funded. From that dataset it became a web interface-based system in 2009. During GEF-4 (2006–10) \$700,000 were earmarked for the special initiative of development of an upgraded PMIS. However, this budget was not used during GEF-4 because of cost-related concerns. Nonetheless, the PMIS had outlived its usefulness; not because it was static, but because it was unable to keep up with increasing expectations.

124. PMIS was not on cloud, was not designed for management of the GEF project cycle, was not accessible to GEF Agencies to input data, and was not adequately secure. These weaknesses ruled it out as the system for the future.

125. The Independent Review of the GEF Systems recommended that the GEF PMIS be integrated into the GEF Trustee's Financial Intermediary Funds (FIF) Information Technology (IT) Project (GEF Secretariat 2011). Although project management information systems of four funds—the Global Partnership for Education, Global Agriculture and Food Security Program, Climate Investment Fund, and Adaptation Fund—were eventually integrated into the Trustee's FIF IT project, the GEF dropped out of the initiative after two years because the progress in the Trustee-led development of the system was slower than expected and did not account for the complexity of the GEF project cycle modalities and workflow.

126. After the decision to drop out of the Trustee-led initiative, the GEF continued to explore other options. With time there was greater support for developing a new system instead of revamping the PMIS. The Secretariat entered discussions with the Information and Technology Solutions (ITS) unit of the World Bank for development of the new system. Subsequently, the GEF Secretariat signed a contract with the ITS unit to develop the GEF Portal. In all, at the start of FY2017, \$658,650—which included the remaining balance of the amount approved in 2005 and \$222,229 repurposed from the remaining balance of the allocation for the 'Performance Based Allocation System'—was available for development of the Portal (GEF Secretariat 2017). This does not include the full costs of staff time involved in the development.

127. The GEF Portal task team within the Secretariat led the process of conceptualization of the Portal. The conceptualization process involved: documentation and analysis of the activities conducted by the PMIS, and identification of the functions of the Portal. The Secretariat team, with support from World Bank ITS, held consultations with Secretariat

staff from the program and policy and operations units, GEF Agencies, GEF IEO and Scientific and Technical Advisory Panel (STAP). The team that managed the PMIS also provided inputs to this conceptualization process.

128. The ITS focused on developing a minimally viable product – within the resource envelope provided to it – as a basis for the discussion and soliciting feedback. The Portal was rolled out in July 2018 to coincide with the start of the GEF-7 replenishment period. The calendar leading up to the launch included a compressed period of testing. The product rolled out in July 2018 offered limited functionality. Subsequently, several additional features have been rolled out. These rollouts are accompanied with online workshops to introduce the new features and to also get feedback on areas where the Portal may improve further.

KEY QUESTIONS AND METHODOLOGY

Key questions

129. The key questions of the evaluation are as follows:

- (a) To what extent has the Portal facilitated enhanced project review and processing abilities and contributed to improvement in operational efficiency?
- (b) To what extent has the Portal improved access to data for key stakeholders in the GEF partnership?
- (c) How well are the key GEF programming strategies and policy requirements integrated in the Portal?
- (d) To what extent has the Portal enhanced transparency through increased real-time availability of data?
- (e) To what extent are the users of the Portal satisfied with its performance?

Methodological approach

Sources of information

130. Information has been gathered from several sources. These include review of documents related to the GEF Portal such as relevant Council documents, publications by the Secretariat, evaluations by the GEF IEO; survey of the Portal; interview of key informants; and an online survey. The evaluation team also observed three workshops conducted by the GEF Portal team to train Portal users on its new features.

Survey of the Portal

131. The survey of the GEF Portal aimed at documenting the browsing experience for a new user. The survey focused on the navigation, design, and comprehensiveness of the Portal, the proposition being that a user-friendly design will have features that are intuitive and do not rely on previous experience or knowledge. A new user perspective may be useful in assessing user-friendliness and identifying areas where the user interface may be improved. The evaluation team conducted this survey by visiting individual pages of the web interface, by running different queries and data downloads, and assessing the capabilities of the Portal. The survey was conducted during October 2020.

Stakeholder interviews

132. During the interview with the GEF Portal team, key stakeholder and user categories were identified. For each category, key informants were identified through consultations with the respective contact persons. For example, the evaluation team contacted the Agency contacts for the GEF to identify staff who were active users and well informed on issues related to use and performance of the Portal. These were then contacted and interviewed (Annex B.1). In all, 33 key informants representing different categories of GEF Portal stakeholders were interviewed (Table 4). Each key informant was asked questions to capture their unique perspective on design, implementation, and use of the Portal and/or issues relevant to assessing its performance.

Table 4: Coverage through interviews

Stakeholder category	Total interviewed
GEF Secretariat	15
<i>Program managers</i>	7
<i>Portal team</i>	3
<i>RBM team</i>	3
<i>Management</i>	1
<i>Country Relations</i>	1
World Bank ITS	2
GEF IEO	4
STAP Secretariat	1
GEF Agency Staff	8
Peer portal (Green Climate Fund)	3
Grand Total	33

Note: IEO = Independent Evaluation Office; ITS = Information and Technology Solutions; STAP = Scientific and Technical Advisory Panel.

Online survey

133. The online survey focused on gathering information on the experience of the active users of the GEF Portal. The survey was administered from November 12–24, 2020. It consisted of questions that categorized the respondents, screened them based on frequency of usage of the Portal, and captured user perceptions of the Portal’s performance in different areas. The evaluative questionnaire consisted of both open-ended questions and questions that used a rating scale. It covered topics such as: website navigation, completeness of web pages, information accessibility, process of proposal submission and appraisal, uploading and downloading of documents, and transparency.

134. The online survey was administered to active users of the Portal. A list of the email IDs of the users who have access to the GEF Portal was acquired from the ITS unit responsible for managing the access. The list was cleaned by excluding the repeated and disbanded email addresses.⁴ After these exclusions, the number of unique recipients of the invitation to participate in the online survey was estimated to be 845 (Table 5). However, all those who have access are not active users; the number of active users who use the GEF

⁴ The prevalence of disbanded email addresses was assessed for the emails in the GEF and World Bank domain, and this data was extrapolated to estimate the disbanded addresses in the list acquired from the Information and Technology Solutions unit. This information was then used to assess the response rate of the active users.

Portal at least once a month is estimated to be 667. Overall, 158 active users (24 percent) responded to, and 133 (20 percent) answered all the questions of, the online survey.

Table 5. Participation figures for survey

Total email addresses listed as users having access to the Portal	1001
Estimated email addresses listed after excluding repeats and disbanded addresses ⁵	845
Estimated active users that use the Portal at least once a month or more	667
Total respondents to the online survey (as % of the estimated survey recipients)	205 (24%)
Active users that responded to the online survey (as % of estimated active users)	158 (24%)
Active users that answered the survey fully (as % of estimated active users)	133 (20%)
Active users that answered the survey fully as % of those that started the survey	84%

135. For the questions that involved use of a rating scale, six-point scales were used without a neutral option, but with a seventh option generically labelled “do not know/unable to assess.” Depending on the question and attribute evaluated, bipolar or unipolar response scales were used. Bipolar scales were used for questions where answers may belong to one of the two sides of neutrality. For example, when assessing difficulty or ease, a response could range between “very difficult” and “very easy.” In contrast, unipolar scales were used for attributes where measuring two sides of neutrality makes little sense. For example, when assessing the importance of an attribute the possible responses ranged from “not important at all” to “extremely important.” The survey results presented in this report exclude the “do not know/unable to assess” responses.

Comparative analysis

136. A comparative analysis was conducted to assess quality and usefulness of the Portal vis-a-vis other peer portals. The evaluation team first searched evaluations of other similar portals that covered their performance, including evaluations that covered these portals as one of the topics.⁶ This search did not yield enough data for analysis; therefore, an alternate approach was adopted. The evaluation team examined three portals—the Green Climate Fund Project Portfolio System; IRENA Project Navigator; and the UNFCCC CDM information system—as external users. These are used for management of information on projects that are focused on addressing environmental concerns, though the scope of activities carried out in these portals varies. The evaluation team visited them and through observation and testing documented their respective features. The evaluation team approached those involved in management of these portals to get their perspective through an interview. The team was able to interview three key informants related to the GCF portal but did not get a response from the teams managing the IRENA Project Navigator and the UNFCCC CDM. Therefore, data gathered for the three peer portals is primarily based on observations made as an external visitor to the three portals.

⁵ This is an estimate because disbandment could be ascertained only for the GEF and World Bank domain-based email addresses. For other Agencies, the rate of disbandment found for the GEF and the World Bank was prorated.

⁶ The search covered organizations such as the European Bank for Reconstruction and Development, the United Nations, Inter-American Development Bank, International Finance Corporation, European Investment Bank, Nordic Development Fund, IRENA, and the Green Climate Fund.

Limitations

137. Some important perspectives are not fully covered in the evaluation. The team that managed the PMIS was also identified as a key stakeholder category and contacted for interview. However, after initially indicating willingness to be interviewed, the informant eventually dropped out. As a result, the perspective of the PMIS team is not reflected in this evaluation.

138. The evaluation team had limited access to the World Bank ITS team. The evaluation team did conduct a one-hour interview with two members of the ITS team. Two other interviews with the team members had to be canceled because of no-shows and non-responses despite follow-up. As a result, the evaluation team is not able to fully address issues related to ITS role and performance and fully reflect their perspective in this evaluation.

139. The Portal is evolving. Much of the analysis presented in this report is based on the status of the Portal in November 2020. Although several new features were introduced after November 2020, these are not reflected in the observation-based analysis presented in this report.

140. The development of the Portal is supported through resources allocated for the special initiative for GEF Management Information System. However, the accounting and reporting of utilization of the budgeted amount does not consider the full costs of development of the Portal. The evaluation was unable to ascertain the full costs of development of the Portal. Therefore, it notes the reported costs as a lower bound of the actual costs.

Findings

Key Finding 1: The Portal has enhanced the online project proposal submission and review capabilities of the GEF Partnership.

141. **The process for submission and review of project proposal has shifted to the Portal.** The PMIS offered little support for the project proposal submission and review activities (GEF IEO 2017b). The proposals were submitted through emails and the documents were then uploaded to the PMIS by Secretariat staff. During the appraisal process it was difficult to keep track of the email exchanges between the Secretariat and the Agencies. Projects were approved/endorsed offline, and the printed versions of documents were signed and then uploaded. The Portal, in comparison, has substantially reduced the need for offline actions. Now the Agencies submit proposals, the Secretariat conducts its reviews and takes decisions on the proposals at the Portal. The Portal creates an audit trail providing information on decisions taken, the persons who took it, and when it was taken. This audit trail and chain of custody facilitates audit of actions taken—or not taken—at the Portal and establishes accountability. During the ongoing COVID-19 pandemic, the ability to submit project proposals and review submitted proposals online has been instrumental in facilitating smooth conduct of project appraisals. The process is not as seamless as it could be because it is not yet supported by a comprehensive system of automatic alerts on changes in a proposal's status and on approaching project cycle related deadlines.

Key Finding 2: Overall, the Portal has contributed to improvement in data quality especially on the recent projects through increased automation and arrangements to ensure data entry discipline, although errors in data outputs are also found.

142. **The Portal has streamlined the process of submitting project implementation reports, mid-term reviews, and terminal evaluations, which may help in reducing submission gaps.** Before the GEF shifted to the Portal, Agencies submitted these documents through emails. The chances of documents getting misplaced and uploaded incorrectly to the PMIS were higher. Further, it was difficult to have an overview of the submission gaps in real time and would require considerable time and effort to ensure compliance. These submissions are now made by the Agencies directly at the Portal, and it is possible to track submission of these documents more effectively.

143. **The Portal encourages discipline in data entry and reduces errors.** The Portal integrates the GEF-7 strategies and core indicators in its data entry interface. It uses an approach where tasks related to data entry are linked with further movement of a given project in the project cycle. For example, project submitters (Agencies) and project reviewers (GEF Secretariat) are unable to move to the next step without entering complete information for the preceding step. This guides those responsible for entering the information—who also have inbuilt stakes in moving the process forward—to enter it for all the fields. Auto-validation is inbuilt for many fields so that the person entering information has real-time feedback on potential errors in the entered data that helps them in entering the correct information. For example, Agencies are required to provide information on core results indicators for GEF-6 and GEF-7 projects at different points in the activity cycle. Earlier, it was difficult to ensure compliance with such requirements without time-consuming back-and-forth. Now the Portal provides real-time feedback to the Agencies on whether the relevant field for the data on core indicator has been filled and allows the Agencies to reach the next-stage web page only after the relevant information is entered. Validation checks incorporated in the Portal design also cover financial envelopes, GEF STAR limits, and Impact Program set-asides. Use of a menu of response options also limits the scope for errors in data entry – it limits a response to what is possible and logically consistent, e.g., responses for project modality, focal area, programming objectives, recipient country, etc. However, several key informants from Agencies noted some technical challenges because the Portal is not well adapted to address needs of different types of projects. For example, the Portal is still not able to split data entry requirements for projects implemented jointly by two or more Agencies. Thus, for such projects, even though the project is being implemented by two or more Agencies, the data may be entered only by the lead Agency. This means that for joint projects, the respective Agencies will need have some off-line back and forth on the information that the lead Agency needs to enter on behalf of other Agencies.

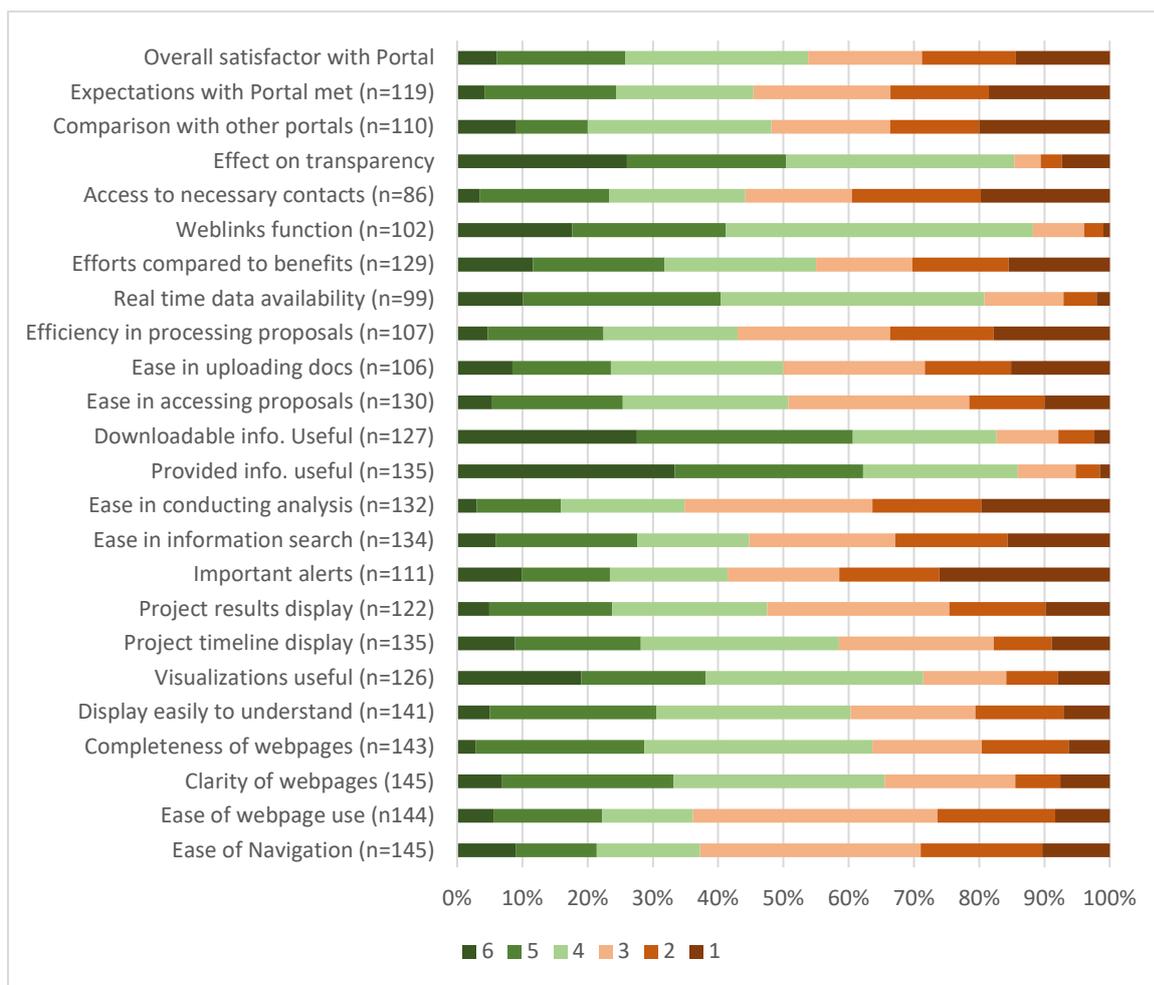
144. **Availability of real-time data on project cycle status has improved, though concerns pertaining to historical data remain and glitches in data outputs continue to be discovered.** Eighty one percent of the responses to the online survey (ignoring doesn't know and unable to assess responses) assessed that real-time availability of data on GEF activities has improved a result of the GEF Portal (figure 1 and Annex B.4). Many actions related to the project life cycle, such as project proposal submission, project appraisal, and submission of project implementation reports, mid-term reviews, and terminal evaluations, are conducted directly on the Portal. There is an audit trail of action taken, the identity of the action taker, and the time stamp of the action taken. This is improving the quality of data on project cycle milestones, project financials, and the status of project proposals, and providing it in real time. Quality of historical data remains a concern despite efforts to improve it. The Portal team filled some of the gaps in the PMIS project data through

collaboration with the Agencies. However, several challenges remain. Some problems cropped up because of data migration from the PMIS to the Portal. For example, original document upload dates got automatically revised by the system: the date of the document is the date of migration, and every time the system gets refreshed, a new date based on that latest system update gets listed. Documents that used to span several years suddenly had the same date, making document-tracking by dates difficult. Several users who were interviewed reported errors in data outputs. These get fixed when brought to the notice of the Portal team, but the turnaround time is perceived to be long, which discourages some from reporting such issues.

145. Portal users have the option to keep versions as draft up until they deem the document as final. Once final, Agencies need to go through GEF Secretariat and ITS to allow for editing data entry. This ensures data integrity and avoids changes on already final submissions. At the same time, it disincentivizes corrections in the uploaded information. Sometimes, when the wrong report is uploaded, it is followed by uploading of the correct copies without the Agency going through the GEF Secretariat and ITS to get the incorrect version removed. Later, when these documents need to be used, the presence of several copies is confusing to a user who may need to spend additional effort in determining which of the versions are correct.

146. **The Portal is more secure than the PMIS.** Management of the PMIS was dependent on the knowledge and experience of the PMIS manager, who had also developed the system. Problems encountered in using the PMIS and new data products required could be addressed by the manager. Dependence on any one person to this extent is a big risk. The PMIS manager and another member in his team were empowered to make changes in the data and related information. While this ability had tactical utility in addressing the glitches, it also meant that it was difficult to track who did what, when, and why. The Portal is more secure in these areas. It is managed by a larger team and is not critically dependent on any one person. It also regulates who enters data, about what, and when. It also has an auditable trail of who did what, when, and why.

Figure 5: User perceptions on Portal performance, on a scale of 6 to 1



Source: Source: Online survey. Note: Actual response options and scales have been condensed to facilitate presentation. All responses were on a six-point scale with 6 representing high performance and 1 representing low performance; 6–4 represents high performance range and 3–1 the low performance range.

Key Finding 3: Technically the Portal is easy to navigate, visually appealing, and accessible and compares well with its peers on these criteria. However, users identify several areas where the design performance needs to be improved.

147. **The Portal design is user friendly.** The Portal has a simplified professional design with a strong logic. The layout, color, and fonts are user friendly (Table 6). It is easy for the user to identify an HTML link; web pages are well composed and with a clean layout. The icons used in the web pages are simple, elegant, and consistent. The Portal does not have a site map, but it works because its design is simple. Most respondents of the online survey were satisfied with the visualizations, displays, completeness, and clarity of web pages (figure 1).

148. **The user perception on ease of navigation is varied.** Of the online survey respondents, only a minority felt that the it was easy to navigate the website and that the web pages of the Portal were easy to use (see figure 1 and Annex B.4). Thirty-seven percent of the respondents assessed navigation of the Portal to be somewhat to very easy. There is some divergence in responses based on frequency of use. While 58 percent of those using the Portal daily assess navigation to be somewhat to very easy (n=38), 30 percent of those that use it at a lower frequency find it to be so (n=107). Sixty-four percent of the respondents assessed that the arrangement of the web pages makes it somewhat to very

difficult for them to use the Portal. Thus, part of the difficulty in navigation is related to irregular use of the Portal and consequently low familiarity with it, and part of the difficulty may be due to the arrangement of pages.

149. The GEF Portal compares well with peer portals in criteria such as navigation, visual appeal, and accessibility. The evaluation assessed the design and performance of the GEF Portal with that of its peer portals on several criteria (see Table 6 and Annex B.3). The comparison shows that the GEF Portal design and performance are comparable to those of peer portals. The design of the GEF Portal is clean—it is simple and has an understandable page layout and topic classification—and similar to the design used by the GCF and IRENA portals. It has a simpler design than the UNFCCC CDM portal, which contains more information but requires users to spend more time in understanding the options and using the portal.

Table 6: Comparison of the GEF Portal with good practices in portal design⁷

Good practices	GEF Portal
Navigation	
Clear page layout through use of proper headings and topic clusters	Menu icons are used. Some of the icons used match with their function. Others are not well matched – may lead a user to expect other type of information than provided.
Clear location on the sitemap through use of breadcrumbs	No.
Ease of understanding the components/features of the site	Menu icons are used. Buttons with different functions share color, screen location, shape, and size. There is potential to improve the visual guidance on the intent for the user.
Ease of finding the information needed	<ul style="list-style-type: none"> • The filter menu may be made more user friendly with hierarchical organization. • The Portal provides the option to reverse filter selections – currently it is the “x” button on filters applied. There may be other ways to help users navigate filtering as well. • Icons leading to pop-ups are not intuitive.
Avoid analysis paralysis by reducing the number of choices	Number of menus offered is limited.
Visual Appeal	
Consistent branding	Yes.
Appealing color scheme/Useful color contrast	<ul style="list-style-type: none"> • Clean title bar. • The green color and blue are highly saturated, and it may be that they compete for a user’s attention. • The logo and green text disappear from the rest of the navigation experience. There may be ways to leverage this color and design to guide users back to the top level. • Buttons with different functions share color. This should be changed to guide the user better.

⁷ Also See Morgeson, Forrest V. (2012)

Content easy to read by being concise, minimizing text and/or using short paragraphs	Minimal text.
Content easy to read through proper text layout	Pop-ups with some distracting text-wrap.
Content easy to read through block reading	Not applicable since the amount of text presented is minimal.
Content easy to read through use of bullet points	Yes. Use of alternating gray bars used for project lists. It would be interesting to check if this style can be pushed in other areas where things are listed.
Use of negative spaces to avoid crowded pages	Clean design with minimal text and negative spaces well distributed.
Use of quality images/photos and balance between words and images	No. Only icons used. More images could make the site more visually appealing.
Use of human face aiming at “putting a face to the name”	No.
Accessibility	
Accessibility for the public	<ul style="list-style-type: none"> Account needed: Only persons/organizations legitimate through the World Bank Group can access. The pertinent result from the web search leads to a blank page. When the “/App” at its end is deleted, the link works.
Functional links	<ul style="list-style-type: none"> The home button leads to a currently empty page. Too many pop-ups. Not only does that disable the back buttons, but it is likely to run into problems with secure browser setups that disable pop-ups.
Optimization for mobile	<i>Unable to assess</i>
Optimization for speed (website speed test performed using the Pingdom Website Speed Test)	<ul style="list-style-type: none"> Load time: 982 ms Overall performance grade A (99/100)
Prioritization of SEO (Website SEO check performed using the SEO Checker from IONOS.)	<ul style="list-style-type: none"> Low optimization, 90 percent optimization potential. The GEF Portal URL (gefportal.worldbank.org) redirects to the Microsoft Online Login page for the World Bank. The first results from web search are documents or news about the Portal rather than the Portal itself.
Strong security (Website security check performed using the SSL Server Test from Qualys SSL Labs.)	<ul style="list-style-type: none"> Overall rating B (satisfactory). The server does not support Forward Secrecy with the reference browsers.

150. **The user assessment of the quality of design of the GEF Portal is mixed.** Fifty-two percent of the online survey respondents—setting aside the “unable to assess” responses—felt that the quality of the GEF Portal was slightly to a lot worse than other portals that they were familiar with (figure 1 and Annex B.4). Among the survey respondents, the perception of GEF Agencies was somewhat different than those of other stakeholders. Only 42 percent of the respondents from GEF Agencies rated the quality of the GEF Portal to be slightly to much better than other portals. In comparison, 61 percent of the respondents from other groups (excluding the GEF Agencies) described the quality of the GEF Portal to be slightly to

much better than other portals; the difference is statistically significant at 90 percent significance level. Based on the supporting evidence gathered through the interviews, it appears that the difference in perspective is driven by inadequate prioritization of the needs of the GEF Agencies in the Portal design.

Key Finding 4. The Portal is much more developed than when it was first launched, and user experience has improved. However, several major gaps remain, and the drawn-out development process is a source of frustration to users.

151. **The Portal was an incomplete product when it was launched, and this led to challenges in transitioning to the new system.** The Portal offered limited functionality and had not undergone structured testing other than checking formats of reports and checking for data accuracy. Several key informants noted that development of the Portal during 2017–18 was on a “compressed” schedule, and that therefore the time devoted to testing was limited. A couple of weeks prior to the launch, some users were provided early access to the Portal to check how it worked, but this experience was insufficient and there was little time to address the identified problems before the launch. Initially, not all user groups were provided access. The Portal also faced connectivity issues: it was difficult to access and its web pages would freeze during use, leading to loss of time and work for the users. This issue was subsequently resolved by the ITS by increasing the bandwidth available for the Portal. For a long time before the reporting function was upgraded, those who could use the Portal were unable to get reports they needed. Even though data in the PMIS were not updated after the launch of the Portal, it was commonly used to generate reports pertaining to GEF-6 and preceding periods. Several respondents report that they still use the PMIS for several functions because the Portal is not fully function and the PMIS seems to be better suited for some tasks. A shared frustration among several respondents was that there is no clear indication as to when the Portal will achieve full functionality.

152. **Since its launch in 2018, several features have been added to the Portal, but there are several areas where the Portal needs to be developed further.** The Portal offered limited function at its launch. In the past two and a half years several additional features have been rolled out. The rollouts are accompanied with workshops and training on use of the new features. The Portal team notes that it is working on an improved reporting platform and developing a dashboard for improved reporting. The users generally appreciate the new additions that have made the Portal more useful to them. At the same time, the users regard the Portal to be a work in progress and identify several areas where it ought to be developed and glitches need to be fixed. These include development of a comprehensive system of alerts on status changes and approaching deadlines/events, providing the ability to batch download documents, enhancing the capabilities of the search function, and ensuring that the calculations presented in the Portal data outputs and reports are always correct. Several interviewees from the GEF Secretariat and from the Agencies noted their frustration at the long drawn out process of the Portal’s development, with no clear end date in sight. The Portal team and World Bank ITS attribute the slower development to the limited resources that they work with. During the GEF-4 period (2006-2010) the GEF Council allocated US \$ 700,000 for the Portal⁸ through a special initiative (GEF Secretariat, 2017). Subsequently, US \$ 222,229 was repurposed from the special initiative for Performance Based Allocation System to the budget for the Portal. Of the cumulative

⁸ Referred to as “GEF Management Information System” in GEF Secretariat (2017).

total of US \$ 922,229 allocated, at the start of fiscal year 2017 US \$658,650 was available for development of the Portal and at the end of fiscal year 2020 \$ 299,000 remained unspent. These cost figures are a low bound and these do not reflect full costs incurred in development of the Portal.

153. The present search function is inadequate to facilitate users in accessing information from the Portal. The current search engine within the Portal offers the possibility of filtering and aggregating (sum, average, etc.) the variables: for example by country, by focal area, etc. Additionally, the material being searched is up to date. Yet, many users reported having difficulty searching for information within the portal. Fifty-five percent of the online survey respondents assessed that it was slightly to very difficult to search for information at the Portal and 65 percent found it slightly to very difficult to analyze content within the Portal. Users seek a feature to facilitate sorting and filtering of the documents within the Portal. For example, they should be able to filter by country or funding source rather than having to comb through reports one by one. Some of the areas where users find the search function to be deficient include:

- (a) Project Implementation Reports, midterm, reviews and terminal evaluations are difficult to find.
- (b) Documents that are saved as annexes are not well integrated with the main document and are difficult to find.
- (c) The Portal is expected to provide the ability to use georeferenced data: so far it allows for only upload of maps, text files and pictures, but does not upload georeferenced shape files that digitally demarcate geographical areas covered by the GEF project.
- (d) Content search is not enabled; for example, a user may not ask how many hectares have been funded by the GEF via land degradation projects and get a list or tabulation.
- (e) Search by indicators and project components is not enabled. For example, a user is not able to know which projects within a portfolio were aimed at capacity building.

154. Search and identification are difficult because of challenges in labeling and tagging. The Portal's lack of categorization makes it difficult for users to sort, filter, and aggregate data. Because of weaknesses in the naming convention followed for the Portal, in most instances documents cannot be identified without opening them. It would be useful if, when searching for documents, the names of the documents were made more obvious so the user could easily interpret what they are looking at. Developing stricter naming conventions would help alleviate this problem. Dates are also a challenge because original document dates are getting automatically revised by the system. The date of the documents is the date of migration, and every time the system gets refreshed, they receive a new date. Documents that used to span several years suddenly all have the same date. This represents another issue, particularly when tracking published documents.

155. The Portal does not allow bulk download of documents, which is a major barrier for users who analyze content. The currently available download function enables the download of data in multiple formats as well as of core indicator reports. The users are generally appreciative of this ability. However, bulk download of documents, which is an important function for users who conduct a portfolio-level analysis, is an area for improvement. In the PMIS, bulk downloading was enabled. In the new Portal, this feature is not enabled: the users are required to download files one by one. Downloading documents

one by one is especially difficult when the documents are located on different web pages. This is a challenge for the users from the GEF IEO who need to analyze a huge volume of documents. For example, for an ongoing review one of the evaluation teams had to download nearly 1,000 reports one by one. It required more than five workdays just to download the set!

156. The alerts provided by the Portal need to be better tailored to the diversity of users and action takers. Presently, the Portal provides alerts to the Agencies through emails on some of the fixed project cycle deadlines. Because GEF Agencies are interested in timely submission and speedy movement through project preparation, these alerts are well aligned with their interests. However, there are areas where the system needs to develop further. It needs to provide alerts that are targeted not only at the Agency staff, but also other action takers at the Portal. In addition to alerts provided close to deadline, some alerts may be provided as different points as the deadline approaches. The Portal also needs to incorporate a system of alerts and reminders for activities that need to be completed within a time band instead of an exact date, e.g., terminal evaluations. An improvement of this nature in the alert system would provide information on status update and would also affect the users' perceived quality of the workflow within the portal.

157. The dashboard is useful for those who need reports from the Portal. Several interviewees emphasized the importance of a dashboard for their work. Currently, the Portal only allows table-style reports, but a dashboard is under development. The dashboard would enable an immediate view of progress, trends, and performance in certain areas from different perspectives, e.g., progress over time, regional, theme, and focal areas. Along with an easy view of the current portfolio based on real time data, it will facilitate a user-friendly project data access.

158. Despite efforts to improve data quality, challenges remain and errors in data outputs and calculations are a concern. Before the migration of the past data from the PMIS to the new portal, the Portal team filled in data gaps with the help of the GEF Agencies. After the migration of data, the Portal team conducted verification checks to assess whether data had been migrated free of errors. Through this process, several issues were found and corrected. However, several challenges related to data quality remain. For example, in the absence of guidelines on the categorization of data, often there is inconsistency in the use of categories to classify data. In several instances, data given in the CEO letter for endorsement/approval do not match with the corresponding data provided in the Portal. The quality of historical data, despite efforts to improve, remains poor and is perceived to be unreliable. Several users reported instances where the calculations/totals provided by the Portal were incorrect and had to be recalculated manually. In some instances, the categories of input data and data output were misaligned. For example, the category of countries that were SIDS reported as least developed countries and vice versa. These issues do get fixed when they are brought to the notice of the Portal team, but they also mean that the data users have to make extra effort to check whether the downloaded data are correct and can be used for further analysis.

Key Finding 5. Users perceive a need for a transparent process to collect information on problems, prioritize problems, and report on the progress in addressing them.

159. Although initially the experience of the users with the Portal was underwhelming, it improved over time. Several interviewees noted a marked change toward a more

positive user experience with the Portal since mid-2019, when users were able to gain full access to the Portal (submit, draft proposals), arguing that since then the Portal has broadly served its needs. However, there is dissatisfaction among users—including users from the Agencies, GEF IEO, and the GEF Secretariat—with what they perceive as numerous low-level glitches and challenges. Several interview respondents noted that the new Portal should have represented dramatic improvements in operational efficiency over the PMIS, and its performance closely aligned with industry standards for such a Portal, but their expectations have not been met.

160. Although users are generally satisfied with the responsiveness of the Portal team, several users are dissatisfied with what they perceive as a lack of clarity in the approach used to identify and prioritize problems in the Portal. A system that is fully responsive to user feedback should meet the following criteria: a clear process for recording complaints; tracking of progress toward resolution; 100 percent or almost 100 percent response rate; and a commitment to a timeframe for providing a response. While there is a process in place for users to report issues that encounter in using the Portal, the arrangements are not as systematic as they could be. Therefore, the experience of users on the Portal team’s performance in addressing their feedback varies. In general, interviewees felt that the number of times they need to reach out and the back-and-forth required for getting glitches fixed was frustrating, though most respondents found the Portal team to be responsive. The feedback to the Portal team, including ITS staff designated for the purpose, is usually through emails. The users from Agencies send feedback to their GEF coordination unit which then forwards it to the GEF Portal team for action. Several online survey respondents felt that the Portal ought to provide an option for direct feedback without having to write separate emails. Forty-one percent of the online survey respondents reported having provided feedback to the Portal team (n=136). Of the respondents that provided substantive details on their interaction (n=36), 83 percent were generally satisfied with their experience (including instances where the team was not able resolve the underlying issue because of technical and or resource limitations). Some respondents noted that the responsiveness of the Portal team was poor up to 2020 but improved thereafter. Several respondents found the process of reporting problems or requesting changes to be long drawn out and were unclear as to what happens when they submit a request and when it will be dealt with. A few felt discouraged from reporting problems.

161. Lack of clarity in communications is often reported as a concern. Several interviewees noted that even though at its launch the Portal was an incomplete product, they did not recall being told so by GEF Management. This gap is perceived to also extend to management of deficiencies: there is little communication acknowledging the shortcomings that persist, and of a process and time frame for fixing them. Portal users express the need for more proactive communication by the Portal team. Several respondents state that there should be at least an overall user manual for the portal and that it should include a troubleshooting section for the most common issues. Similarly, although the Portal is in part meant to already function as an open data source, several interviewees mentioned the need for additional efforts to make data transparent and easily distributed, e.g., by giving users more detailed guidance on the data being presented. This could be partly accomplished through a codebook for spreadsheets describing what can be found in the different columns.

162. The Portal has shifted some of the work burden to the Agencies. Analysis of the online survey data shows that GEF Agencies account for 65 percent of the active users and

54 percent of the frequent⁹ users of the Portal. Agencies are responsible for submitting project proposals, project implementation reports, midterm reviews, and terminal evaluations, to the GEF. Earlier, they used to submit these documents through emails, which were then uploaded to the PMIS by the GEF Secretariat and GEF IEO (for terminal evaluations) staff. Since the advent of the Portal, almost all these exchanges have moved to the Portal. For the Secretariat, the Portal has removed the need to upload project documents. However, Agency staff spend more time in uploading the documents than before. Though new approach reduces errors in submission, the Agencies feel that the additional benefits at their end have not been commensurate with the additional effort.

Key Finding 6. The Portal has contributed to the acquisition of more and better data. However, there are barriers to seamless entry and accessing data.

163. **The Portal is contributing to acquisition of more and better data, but availability is restricted.** More data are being generated through the online processing of project proposals, reviews, uploading of documents, and entry of data on key milestones. However, availability of data to users is more restricted. During interviews, users expressed that sometimes the data are simply not available to the public. The portal feeds some of the information to the GEF website, which is accessible to the public. Limited data access impedes education-related research by scholars. In addition, Agency staff often receive requests from countries concerning the status of their funding but do not have the relevant information available to them. This may be driven in part by the differences in the disclosure policies of the GEF and the GEF Agencies, resulting in different levels of access to data for various stakeholders.

164. **Users are generally happy with the level of transparency the Portal provides.** Eighty-five percent of the active users of the Portal agreed, somewhat to completely, with the statement that “the GEF Portal has contributed to increased transparency in GEF operations.” Non-GEF-affiliated users of the Portal were generally happy that they have the same access to data as the GEF Secretariat.

165. **Connectivity is a major concern for many users.** Earlier most users faced difficulty in working at the Portal due to challenges in logging in, connection losses, and “silent logouts” caused by the page timing out leading to wasted effort. These challenges become more acute when there was heavy use of the portal, such as around deadlines. Much of this problem was due to relatively narrow bandwidth assigned to the Portal – it was subsequently addressed by the ITS and now sufficient bandwidth is provided to the Portal. However, at the user end the challenges remain in countries that have poor internet connectivity. This disproportionately affected users in least developed countries and remote areas. Concerns at the user end also need to be addressed because with decentralization of data entry to Agencies—especially project managers who are posted in recipient countries—the Portal needs to be accommodate the technical constraints of most, if not all, of its users.

CONCLUSION

166. The Portal has substantially achieved its objectives related to enhancement in project review and processing abilities, capturing of information in a consistent format, integration of GEF programming strategies and policies into the Portal, tracking of results of

⁹ At least a few times a week.

GEF activities, enhanced transparency, safeguarding of confidential information, and ensuring data integrity. As a result of the progress in achieving its objectives, the Portal is an improvement over the PMIS. It also compares well with its peer portals on some of the technical parameters on which their performance was compared. At the same time, its performance is mixed in terms of taxonomy and tagging, search and analytical abilities, and real-time availability of data to external stakeholders. Other gaps in performance include lack of ability to download batches of documents, lack of capability to send project cycle-related auto-alerts through emails, and errors in data outputs.

167. Several interview respondents and online survey participants expressed frustration of the slow and prolonged process of development. Although those who brought deficiencies and glitches to the notice of the Portal team were largely satisfied with their interaction, a sizable proportion of respondents were either not sure what the process was to provide feedback to the Portal team or how the Portal team will address their requests. This suggests scope for more proactive communication by the Portal team with the users so that they are better aware of the process and have a clearer idea of the road map to further development of the Portal.

168. The advent of the Portal shifted the burden from the Secretariat to the GEF Agencies, but the Portal development process has not prioritized their needs. The Portal team has linked slow development of the Portal to the limited resources that they have to work with. At the same time there are substantial resources for the special initiative for the Portal that have not been fully utilized. The GEF Management needs to assess whether additional support is necessary and how speedier development of the Portal can be aided.

RECOMMENDATIONS

169. The evaluation has two recommendations:

- **The GEF Secretariat should strengthen its process to address user feedback on the Portal.** The evaluation found that the Portal team has been readily available to address user needs. But the present process for addressing user feedback needs to be strengthened so that it fully meets user needs. The strengthened process should enable direct feedback through the Portal along with the options that are presently available. It should also record user feedback/complaints, require a response within a committed time frame, and track progress towards resolution.
- **GEF Management should develop and implement a time bound plan to speed up the development of the Portal.** The users of the GEF Portal perceive that despite significant progress, the development of Portal has continued for long. For some, especially Agencies, this perceived delay in completion of the development phase is a source of frustration and it may be causing some inefficiencies across the GEF Partnership. Several gaps need to be addressed and it may be more cost effective to address them sooner rather than later. The GEF Management should assess how best it may speed up the process and implement its plan.

REFERENCES

Part A: References for Evaluation of the Agency Self Evaluation Systems

- ADB IED. 2020. Annual Evaluation Review 2020. ADB's Project Level Self Evaluation System. Independent Evaluation Asian Development Bank, March 2020. Available at: <https://www.adb.org/sites/default/files/evaluation-document/539791/files/aer2020.pdf>
- AfDB IDEV. 2018. Evaluation of Quality Assurance across the Project Cycle of the African Development Bank Group (2012–2017). Synthesis Report. Available at: <https://idev.afdb.org/sites/default/files/Evaluations/2020-03/Quality%20Assurance%20Synthesis%20Evaluation%20Report.pdf>
- Arrow, Kenneth. 1984. The economics of agency. No. TR-451. Stanford University California Institute for Mathematical Studies in the Social Sciences, 1984. Available at: <https://apps.dtic.mil/sti/pdfs/ADA151436.pdf>
- Bester, A. 2016. Results-Based Management in the United Nations Development System. United Nations Development Programme. Available at: <https://www.un.org/en/ecosoc/qcpr/pdf/sgr2016-studies-rbm-8jan2016.pdf>
- EBRD. 2019. Independent external evaluation of EBRD's evaluation system. Final Report: Main Report. Available at: <https://www.ebrd.com/what-we-do/evaluation-full-report.pdf>
- ECG. 2018. ECG Practice Note: Self-evaluation in ECG member institutions. Available at: <https://www.ecgnet.org/documents/46856/download>
- GEF. 2019a. Minimum Fiduciary Standards for GEF Partner Agencies. GA/PL/02. Available at: https://www.thegef.org/sites/default/files/documents/gef_minimum_fiduciary_standards_partner_agencies_2019.pdf
- . 2019b. Policy on Monitoring. GEF Secretariat, GEF/C.56/03/Rev.01, June 2019. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.C.56.03.Rev_01_Policy_on_Monitoring.pdf
- GEF IEO. 2010. The GEF Monitoring and Evaluation Policy 2010. GEF Independent Evaluation Office, November 2010. Available at: <https://www.gefio.org/sites/default/files/documents/reports/gef-me-policy-2010-eng.pdf>
- . 2014. GEF Annual Performance Report 2013. GEF Independent Evaluation Office, September 2014. Available at: <https://www.gefio.org/sites/default/files/documents/reports/apr-2013.pdf>
- . 2017. Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects. Available at: <https://www.gefio.org/sites/default/files/ieo/evaluations/files/gef-guidelines-te-fsp-2017.pdf>
- . 2019. The GEF Evaluation Policy. Available at: <http://www.gefio.org/sites/default/files/ieo/evaluations/files/gef-me-policy-2019.pdf>
- . 2020a. Evaluation of the Agency Self Evaluation Systems. Revised Approach Paper. Available at: http://www.gefio.org/sites/default/files/ieo/documents/files/agency-self-evaluation-2020-approach-paper_0.pdf
- . 2020b. Annual Performance Report. Available at: <https://www.gefio.org/sites/default/files/ieo/evaluations/files/apr-2020.pdf>
- . 2020c. Report on the Review of the GEF Terminal Evaluation Validation Process. GEF Independent Evaluation Office. November 2020. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.E_C59_Inf_01_report_of_GEFIEO_TE_validation_process.%20docx_0.pdf

- — —. 2020d. Evaluation of Knowledge Management in the GEF. GEF Independent Evaluation Office, GEF/E/C.59/04, November 2020. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.E_C59_04_evaluation_of_KM_GEF_2020.pdf
- GEF STAP. 2019. Theory of Change Primer. A STAP document. Scientific and Technical Advisory Panel of the Global Environment Facility. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF_STAP_C.57_Inf.04_Theory%20of%20Change%20Primer_0.pdf
- IDB. 2008. Development Effectiveness Framework. Available at: <https://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=1775544>
- MoFA Denmark. 2010. Peer Review of the Evaluation Function of UNIDO. Evaluation Department, Ministry of Foreign Affairs of Denmark, March 2010. Available at: [file:///C:/Users/wb269966/Downloads/UNIDO%20final%20peer%20review%20\(1\).pdf](file:///C:/Users/wb269966/Downloads/UNIDO%20final%20peer%20review%20(1).pdf)
- OECD. 2002. Glossary of Key Terms in Evaluation and Results Based Management. Available at: <http://www.oecd.org/development/peer-reviews/2754804.pdf>
- Picciotto, Robert. "Towards an economics of evaluation." *Evaluation* 5, no. 1 (1999): 7-22.
- Ross, Stephen A. "The economic theory of agency: The principal's problem." *The American economic review* 63, no. 2 (1973): 134-139.
- Scriven, Michael. *Evaluation bias and its control*. Kalamazoo, MI: Evaluation Center, Western Michigan University, 1975.
- Spender, J-C. "Making knowledge the basis of a dynamic theory of the firm." *Strategic management journal* 17, no. S2 (1996): 45-62.
- WBG IEG. 2014. Learning and Results in World Bank Operations: How the Bank Learns. Available at: https://ieg.worldbankgroup.org/sites/default/files/Data/reports/chapters/learning_results_eval.pdf
- WBG IEG. 2016. Behind the Mirror: A Report on the Self-Evaluation Systems of the World Bank Group. Available at: <https://ieg.worldbankgroup.org/evaluations/ar2016>
- WBG IEG. 2021. The Project Performance Assessment Report. [online] Available at <https://ieg.worldbankgroup.org/methodology/PPAR> [Accessed 18 January 2021].
- OECD. 2005. Knowledge Management. Glossary of statistical terms. [online] Available at: <https://stats.oecd.org/glossary/detail.asp?ID=6878> [Accessed 20 January 2021]
- Patton, M.Q. 2010. Future trends in evaluation. From policies to results. *Developing capacities for country monitoring and evaluation systems*, pp.44-57.
- Senge, P. 1990. *The Fifth Discipline, the Art and Practice of the Learning Organization*.
- Taut, Sandy. "Studying self-evaluation capacity building in a large international development organization." *American Journal of Evaluation* 28, no. 1 (2007): 45-59.
- UNEG. 2012a. Professional Peer Review of the Evaluation Function of United Nations Environment Programme. United Nations Evaluation Group, March 2012. Available at: https://wedocs.unep.org/bitstream/handle/20.500.11822/272/Professional_Peer_Review_of_the_Evaluation_Function_of_UNEP.pdf?sequence=1&isAllowed=y
- — —. 2012b. Professional Peer Review of the Evaluation Function of Food and Agriculture Organisation. United Nations Evaluation Group, July 2012. Available at: <http://www.uneval.org/document/download/1560>
- — —. 2016. Norms and Standards for Evaluation. New York: United Nations Evaluation Group. Available at: <http://www.unevaluation.org/document/download/2787>
- Yin, Robert K. 2017. *Case study research and applications: Design and methods*. Sage publications.
- Zall Kusek, Jody, and Ray Rist. 2004. *Ten steps to a results-based monitoring and evaluation system: a handbook for development practitioners*. The World Bank, 2004.

Part B. References for Evaluation of the GEF Portal

- GEF (Global Environment Facility) Council. 2020. Summary of the Chair - Decision on Agenda Item 07. 58th GEF Council Meeting, June 2020. GEF, Washington, DC. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.C.58_Summary_of_the_Chair_0.pdf
- GEF IEO (Global Environment Facility Independent Evaluation Office). 2002. The First Decade of the GEF – Second Overall Performance Study of the GEF. January. GEF, Washington, DC. Available at: <https://www.gefio.org/sites/default/files/documents/reports/ops2.pdf>
- . 2005. Progressing Toward Environmental Results – Third Overall Performance Study of the GEF. June. GEF, Washington, DC. Available at: <https://www.gefio.org/sites/default/files/documents/reports/ops3.pdf>
- . 2010. Progress towards Impact – Fourth Overall Performance Study of the GEF. April. GEF, Washington, DC. Available at: <https://www.gefio.org/sites/default/files/documents/reports/ops4.pdf>
- . 2017a. The GEF in the Changing Environmental Landscape – Sixth Comprehensive Evaluation of the GEF. 2018 Report number 110. GEF, Washington, DC. Available at: <https://www.gefio.org/sites/default/files/documents/reports/ops6-report.pdf>
- . 2017b. Review of Results-Based Management in the GEF. Evaluation Report number 112. GEF, Washington, DC. Available at: <https://www.gefio.org/sites/default/files/documents/reports/rbm-study-2017.pdf>
- . 2020. Seventh Comprehensive Evaluation of the GEF (OPS-7) Approach Paper. May. GEF, Washington, DC. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.E_C58_02_OPS7_Approach_paper_June_2020_0.pdf
- GEF Secretariat. 2011. Independent review of GEF Systems. GEF Council Document GEF/C.41/Inf.12. October 2011. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/C.41.Inf_12_Independent_review_of_GEF_Systems_0_4.pdf
- . 2012. Options for Strengthening GEF Systems: Addressing the Findings and Recommendations of the Independent Review of GEF Systems. GEF Council Document GEF/42/Inf.12. May 2012. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/C.42.Inf_12_Options_for_Strengthening_GEF_Systems_4.pdf
- . 2017. GEF Business Plan and Corporate Budget for FY18. GEF Council Document GEF/C.52/06. May. GEF, Washington, DC. Available at: http://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.C.52.06_GEF_Business_Plan_and_Corporate_Budget_for_FY18_0.pdf
- . 2018. Progress Report on the Development of the New GEF Portal. Council Document GEF/C.54/Inf.06. June. GEF, Washington, DC. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.C.54.Inf_06_GEF_Portal.pdf
- . 2019. GEF Business Plan and Corporate Budget for FY20. GEF Council Document GEF/C.52/06. May. GEF, Washington, DC. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF_C.56_06_GEF%20Business%20Plan%20and%20Corporate%20Budget%20for%20FY20.pdf

———. 2020. GEF Business Plan and Corporate Budget for FY21. GEF Council Document GEF/C.52/06. May. GEF, Washington, DC. Available at: https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF_C.58_03_GEF%20Business%20Plan%20and%20Corporate%20Budget%20for%20FY21.pdf

Morgeson, Forrest V. 2012. "E-Government performance measurement: A citizen-centric approach in theory and practice." In *Electronic governance and cross-boundary collaboration: Innovations and advancing tools*, pp. 150-165. IGI Global.

ANNEXES

Annexes - Part A. Evaluation of Agency Self Evaluation Systems

Annex A.1 Practice Glimpses

Practice Glimpse 1: IADB enhances development effectiveness through an extensive guidance & support infrastructure

	<p>Context:</p> <p>Being a development institution, IDB is constantly faced with the challenge of accurately assessing the bank's performance and ensuring its development effectiveness. While already following international standards and good practices for evaluation of the Evaluation Cooperation Group (ECG), IDB acknowledged that there is still more work that can be done internally in order to fully overcome deficiencies and improve its development effectiveness. In 2008, IDB introduced the Development Effectiveness Framework (DEF) which is embedded into the project cycle of all IDB projects. With the purpose of setting clear standards for the evaluation of bank interventions, providing guidance for project teams, and monitoring the bank's progress in key development effectiveness indicators, the DEF encompasses various tools offering a comprehensive guide for project teams.</p>
<p>Good Practice:</p> <p>The bank established the Office of Strategic Planning and Development Effectiveness (SPD) to provide essential support to teams throughout the project cycle. This includes support on the framework, evaluation guidelines and their application. Core products featured in the DEF and supported by SPD are the Development Effectiveness Matrix (DEM) at design stage of a project and the Project Completion Report (PCR) to assess quality at exit of interventions. The DEM is a checklist comprising best practice standards to design and monitor evidence-based projects. The dimensions of the matrix are evaluability, strategic alignment with the Corporate Results Framework, project risks and additionality of the project to ensure and control the quality of the bank's operations. SPD offers on demand trainings on the preparation of high-quality PCRs. These sessions are organized around presentations on the core and non-core criteria of the PCRs, followed by breakout work groups where project teams can work on their individual PCRs with the support of an SPD facilitator. Additionally, SPD offers support in conducting impact evaluations; up to date, IDB has carried out more than 600 impact evaluations</p>	<p>Effects:</p> <p><i>"PCRs prepared at IDB are a credible telling of a project's history, relevance to the development challenges facing the country, the delivery of results and the efficiency and sustainability of these (Leonardo Corral)."</i></p> <p>As a result, the SPD regularly receives requests for training sessions from all sectoral divisions within IADB with project teams appreciating the support. Furthermore, the department has become the "go-to" department for project teams for all questions related to the Development Effectiveness Framework and its tools, including questions on impact evaluations, evaluability at entry, monitoring during execution, and PCRs at closure.</p>

which function as a valuable source for determining institutional effectiveness.

Sources: Interview with Leonardo Corral, IADB (2008)

Practice Glimpse 2: CAF and FECO value GEF standards and Agency practices as inspiration for their own evaluation procedures



Context:

Having entered the GEF partnership only in 2015, the Development Bank of Latin America (CAF) is one of the newest GEF Agencies. CAF currently has six GEF projects in execution with the first terminal evaluation for a GEF project yet to be due. Similarly, FECO joined the GEF partnership only a few years ago and has not had any Terminal Evaluations yet. Therefore, both institutions are particularly eager to exchange with and learn from more experienced Agencies within the partnership on best practices and past experiences.

Good Practice:

“Since CAF is a relatively new Agency, we are prompt and willing to learn from past experiences and at the same time to coordinate future actions for the upcoming GEF agenda (René Gomez-Garcia Palao).”

Being a new GEF Agency, CAF highly values GEF standards and policies on self-evaluation processes as they help the Agency to identify potential flaws in its own system and foster internal procedures to continuously improve.

Likewise, FECO has systematically leveraged its pre-accession harmonization to review and compare GEF requirements and international good practice with its own approaches. This way, FECO refined its own evaluation system for GEF projects with an extensive review function by its GEF coordination unit and learning meetings for various stakeholders. Within the GEF ecosystem, CAF particularly appreciates platforms such as the inter-Agency retreats, that are organized by the GEF secretariat as a space to create an open dialogue among Agencies and to discuss key issues around projects, evaluation, and learning experiences.

Effects:

Besides positive spillover effects on CAF’s and FECO’s self-evaluation systems, inter-Agency exchanges such as the retreats are seen as a valuable channel to maintain a fluid communication, among the entire GEF partnership ecosystem. The presentations and discussions as part of the retreats allow for learning from past experiences and at the same time help CAF become fit for the future. This includes discussions about upcoming trends and coordination of future actions that are aligned with the GEF agenda. Lastly, various Agencies have been experiencing an openness among the partnership to support each other. The retreats help in establishing working relationships across Agencies that might even include discussions about future project collaboration in specific countries or regions.

Sources: Interviews with René Gomez-Garcia Palao (CAF), Liu Lei (FECO)

Practice Glimpse 3: EBRD's IEU pre-views terminal evaluations before submission

	<p>Context:</p> <p>The independent evaluation office of EBRD (EvD) has been facing challenges with the quality of reports produced by operational teams. One key reason has been the bank's overall results architecture. Because it is a bank, management signals stress the importance of project approvals and loan processing; but project evaluation, producing high quality reports, and learning from past experiences are less incentivized. Additionally, ratings at EBRD as part of self-evaluation products are currently not part of the bank's corporate scorecard; therefore, they are not tied to consequences in case of poor project performance—which operational teams are aware of. Regarding the overall self-evaluation system at EBRD, there is a need for the bank to reposition its self-evaluation system, shifting the responsibility and ownership from EvD toward management.</p>
<p>Good Practice:</p> <p>It is important to note that EBRD is in the process of revising its self-evaluation system, particularly because of the time intensity of the current process. However, to ensure the effectiveness of the validation exercise given the current situation, EvD has included an additional review layer in the process prior which requires project teams to hand in a first draft of the terminal evaluation prior to validation. EvD reviews the draft by checking the quality and completeness e.g., whether outputs, outcomes, impacts, and lessons are reported on appropriately. Based on the feedback from EvD, operational teams are able to make adjustments before the final submission. An evaluator has 5 days to review the draft, and the operational teams have another 3–5 days to incorporate the feedback and submission. Only after the final submission, the report undergoes validation by EvD.</p>	<p>Effects:</p> <p><i>“Usually, when ratings are off by a large order of magnitude, the review process will correct that” (Barry Kodolkin)</i></p> <p>While this review process might change in the course of EBRD's development of its new self-evaluation system, the prevalidation review has helped EvD in making the validations of reports more effective. The review has a large impact on the quality of reports produced by operational teams and supports the flagging of incompleteness and lack of evidence to support certain rating outcomes. Instead of dealing with these issues at validation stage when it is almost too late, the process allows them to be addressed at an earlier stage.</p>

Sources: Barry Kodolkin, EBRD (2019), EBRD (2018)

Practice Glimpse 4: AfDB's triple loop review process

	<p>Context:</p> <p>Institutional effectiveness requires a good quality management system that ensures the reliability, credibility and quality of information being produced as part of monitoring and evaluation (M&E) processes. While the validation of reports by an independent entity is probably the most common arrangement put into place to ensure quality, there are some mechanisms that can enhance the quality of evaluation products even before the final validation. Such mechanisms may include peer reviews and feedback loops as part of the process of preparing terminal evaluations.</p>
<p>Good Practice:</p> <p>AfDB has a Quality Assurance Framework in place that covers all phases of a project cycle. As part of the so-called Quality Assurance at Exit (QaE) of public sector operations, the framework foresees the preparation of a project completion report which must go through an extensive review process before being submitted to the independent evaluation office (IDEV). This review and clearance process includes three instances. First, a first draft of the report is submitted to the sector manager for clearance. Second, the draft is sent to peer reviewers who should either be based in a field office or have previous field experience. And third, the document is sent to the country or regional team for final clearance. At the end of the fiscal year, a sample of project completion reports is selected and sent to IDEV for validation and quality assessment. IDEV prepares its evaluation note based on all available documents and the discussions from the review process, as explained above. Additionally, the evaluation note itself is subject to peer review to ensure consistency in applying all guidelines and will be distributed to the operational department who then can either agree or disagree with the findings. In case of the latter, the department will have to provide additional evidence to support their response.</p>	<p>Effects:</p> <p>The review processes as part of the Quality Assurance framework allow for extensive dialogues and discussions throughout the preparation and validation of the project completion report. This also includes an exchange on the ratings assigned to a project. While discussions among teams happen, the overall process is perceived as highly useful because it helps the bank to come out with an independent, consensual, and reasonable assessment of a project's performance and additionally reducing bias stemming from the self-assessment of projects.</p>

Sources: Interview with Guirane N'Diaye, Ayanleh Daher Aden, Hedi, Manai, Lafeta, Camera; Documents: AfDB (2018)

Practice Glimpse 5: FAO's Reference Groups ensure Quality Assurance and learning from evaluations



Food and Agriculture Organization
of the United Nations

Context:

Independent evaluation units (IEUs) of many GEF Agencies are seeking ways to increase the adoption of lessons learned, including on the national level. Meanwhile a trusted and collaborative relationship with the national counterparts may improve candor and the reliability of information shared. At the same time, resources are scarce and the IEUs are required to ensure their various functions including in most cases Quality Assurance (QA) and dissemination of lessons learned on terminal evaluations are satisfied.

Good Practice:

FAO has developed a stakeholder approach to support QA as well as feasibility of recommendations and dissemination of lessons learned. It applies this stakeholder approach by developing a 'reference group,' which can be formed early on in a terminal evaluation process, at first with internal stakeholders while extending the group to external stakeholders such as Operational Focal Points (OFPs) and national counterparts at a later stage. The QA function of the reference group includes support in the design of the terminal evaluation, review of terms of reference, data collection, reviews of various draft stages, and advice on the feasibility of recommendations. Preliminary findings are then shared with internal and external stakeholders in the interest of validating insights and paving a way for the adoption of lessons learned.

Effects:

The highly consultative process applied by FAO supports the reliability and quality of its evaluation reports. The inclusion of internal and external stakeholders such as the management counterpart and external parties such as OFPs supports capacity building. Moreover, it may improve the feasibility of recommendations and make the adoption of lessons learned more likely.

Sources: Interviews with Rachel Bedouin, Team Lead Office of Evaluation (OED), Jenin Assaf, Evaluation Officer OED, Lavinia Monforte Evaluation Manager OED; FAO. 2015. OED Evaluation Manual. Office of Evaluation. Available at: http://www.fao.org/fileadmin/user_upload/oed/docs/OED_Evaluation_Manual_April_2015_new.pdf



Context:

“At the end of the day it is the culture of the institution needing to value the M&E [...]” (Garrett Kilroy) emphasizing the importance of creating and nurturing an evaluation culture.

In the past, self-evaluation was not a priority for task team leaders at the ADB. Because it is a bank, the focus of the organization is on lending approvals and loan processing; hence the critical achievement is to get deals through the council. ADB's Independent Evaluation Department (IED), reflecting on this inherent management logic, decided to change the game. The key question they asked themselves was ‘how could they highlight the relevance and potential of self-evaluation and reward high quality evaluations?’

Good Practice:

To mitigate the underlying imbalance of priorities, IED introduced the annual IED Awards in 2014 as part of which ADB teams and departments are awarded that have excelled in producing high-quality self-evaluation reports. The candidates are drawn from the terminal evaluations that were validated by IED. The assessed criteria include adherence to evaluation guidelines, adequacy of evidence and analysis, and quality of lessons and recommendations. Awards are given out based on five categories including best reports for sovereign operations, best reports for non-sovereign operations, and the best operation department on self-evaluation. For the ceremony, all ADB staff, board, management, and consultants are invited. The award is given out by IED management, preceded by a justification for the nomination of the reports that is read out. The winners are given a trophy, plaques, and certificates. IED extensively promotes them on ADB's internal communication platform as well as on social media. Additionally, an award ceremony video is produced and disseminated both internally and externally.

Effects:

While there is still more work to be done especially on management side to send signals on the importance of M&E throughout the whole project cycle, the awards have increasingly gained traction within ADB.

“Over the years the interest in the awards has increased and IED Awards have now become very ‘aspirational’.” (Garrett Kilroy)”

This can be seen in the number of ceremony attendees that has been constantly increasing over the past years, reaching more than 100 staff and consultants that attended the 2020 virtually ceremony. Among the participants were the board including the Vice Presidents and Director Generals of most operational departments. As reported by IED's project validation team, the awards have positively impacted the overall quality of terminal evaluation's prepared at ADB. This has been accompanied by an increased interest and inquiries from operations departments on sample best self-evaluation reports.

Sources: Interview with Garrett Kilroy

Practice Glimpse 7: Collaborative Project Implementation Review at UNDP improves candor



Context:

UNDP puts a strong emphasis on stakeholder-led approaches as part of these three goals of evaluation: learning, accountability, and capacity building. One of the mechanisms to ensure that the evaluation function is a stakeholder-led process, also given its decentralized model of operation, is through its collaborative reporting and rating of Project Implementation Review (PIR) between the different Agency and country offices.

Good Practice:

The reporting for the PIR in UNDP is conducted in collaboration between the GEF technical advisor, the Program Manager, the Country Office Focal Point, GEF Operational Focal Point, and sometimes also with Regional Manager; it therefore entails a huge effort in collaboration between different offices and functions. However, depending on projects, the above parties will play greater or lesser roles in the PIR reporting and rating.

The format of UNDP's PIR involves an assessment of progress of the result framework, including the progress in achieving indicator targets, from baselines to midterm target levels and finally end of project target levels. In addition, there is an 'Overall Development Objective Rating.' This rating is substantiated by the evidence presented in the progress in achieving indicator targets and by the UNDP-GEF Technical Adviser, who provides a summary narrative to explain the rating. There is also an 'Overall Implementation Progress Rating, which is the cumulative disbursement against the approved budget, as well an 'Overall Risk Rating.' The PIR also provides a summary of how well gender mainstreaming is happening in the project, how well social and environmental safeguards are being considered, its communication impact, and how its leveraging partnerships.

Effects:

The collaborative reporting and rating of PIR helps UNDP ensure coherence and coordination within UNDP's decentralized model of self-evaluation.

This process allows for extensive dialogues and discussions happening throughout the preparation and rating of the PIR. More evidence can be built also on progress with indicators through this collaboration. The collaborative process can also help the project come out with a consensual assessment of a project's performance and issues, additionally reducing bias stemming from the self-assessment of the project only by the project manager. The overall effect of a collaborative process is a greater candor in reporting as well as better reporting of oversight.

The effect of collaborative PIR on candor is substantiated by the GEF IEO annual report assessment, which notes that 'for UNDP-implemented projects, either country officers or regional technical advisors or both tend to raise implementation challenges in PIRs with clarity, forcefully, and in good time, more often than the respective project managers' (GEF APR 2019).

Sources: GEF IEO (2020); Interview with Pascale Bonzom

Practice Glimpse 8: Post-project evaluation at the World Bank Group for an evidence base to learn from projects

 <p>WORLD BANK GROUP</p>	<p>Context:</p> <p>The Independent Evaluation Group (IEG) of the World Bank is responsible for evaluating the overall development effectiveness of the World Bank Group. IEG carries out a range of evaluation products on several levels, supporting the bank by strengthening its accountability function and generating lessons from past operations.</p>
<p>Good Practice:</p> <p>As part of its project-level evaluations, IEG has a post-project evaluation instrument, called the Project Performance Assessment Report (PPAR). The PPAR is a field-based evaluation conducted for 20–25 percent of bank lending operations. Over the years, the World Bank has shifted the focus of PPARs from ratings, which are now placed into the appendix, to making the report more learning oriented. In doing so, the report should provide advice and feedback on what worked, what didn’t work, and why.</p> <p>Preparing a PPAR requires the IEG reviewer to be engaged from the very beginning of the evaluation to provide advice on evaluation questions and methods. Being a project evaluation and not a validation, the PPAR relies on a mixed-method approach including direct evidence collection, site visits, and stakeholder consultation. While validation reports (ICRRs) as part of IEG’s standard validation process are based on evidence that is available at project closure, the PPAR is done after sufficient time has passed, which additionally allows for capturing evidence-based long-term effects. The projects for which the PPAR is conducted are selected based on certain criteria that can be, but are not limited to, the existence of evidence gaps, alignment with IEG work program priorities (e.g., thematic evaluations), or the potential for learning from innovative projects. Usually, IEG seeks to examine clusters of similar projects to increase the operational significance for project teams and the learning potential of PPARs.</p>	<p>Effects:</p> <p>The PPAR adds an additional layer to quality-assuring results and verifying the development performance of the Agency. The collection of direct evidence in particular to address evidence gaps, the consultation of stakeholders, and also the timing of the PPAR, sometimes several years after project closure, allow for a more rigorous assessment of a project.</p>

Sources: Stephen Hutton (IEG); IEG WBG (2014), IEG WBG (n.d.)

Practice Glimpse 9: Three Agencies innovate their self-evaluation systems with the application of Theory of Change



Context:

Despite moves from logical framework to Theory of Change (ToC) models in the field of development, monitoring and evaluation (M&E) design remains a linear logic: construct models and diagrams in which inputs lead to activities, activities lead to outputs, and outputs lead to outcomes. Patton (2010) discussing the future of evaluation emphasizes that we will need increasingly to look at projects and programs from a systems perspective which focuses on the interdependent configuration of factors that lead to outcomes rather than a simple cause-effect model.

Some GEF Agencies are, however, innovating their self-evaluation systems with the application of ToC as a methodology— as both systems thinking process and construction of systems models. Here we give the examples of how three Agencies are using ToC and the effects its application is having on quality of self-evaluation.

Good Practice:

World Bank Group: ToCs were made mandatory in 2017 as part of the World Bank’s reform of project terminal self-evaluations. ToC was introduced for all newly approved projects and for all project terminal evaluations. The goal is to tell a more accurate performance story that aligns to real project intentions. Leadership from the Bank Group’s operational policy vice presidency, with support from the independent evaluation function, encouraged the use of ToC among operations staff. Further, enforcement of the mandatory requirement also scaled up the uptake of ToC as a tool.

UNEP: UNEP evaluations are structured around a project or programme’s Theory of Change to assess the causal logic of the intervention and

Effects:

World Bank Group: The application of ToC was perceived by IEG to be generally successful, though there has been variation in practice across sectors/regions. In other words, quality of ToC can greatly vary. The World Bank Group’s IEG, as a result, frequently offers training clinics on ToC and gives support to teams on creating their ToC. The use of ToC helps teams in various project phases. At project design, it is helping to improve clarity and communication around project logic as well internal quality review processes to debate design logic; it is especially a valuable tool for engaging clients through participatory process. For evaluation, it is helping teams to debate about what worked and did not work from results (outcomes and impacts) rather than delivery perspective (outputs). Over the long term, the World Bank Group would like to encourage teams to use ToC to articulate longer term goals/higher level outcomes, beyond the accountability function. It would like to see its application to understanding impacts beyond project (cumulative effects, replication/scale-up effects, etc.) as well as identifying the type of indicator data to collect to track higher-level country outcomes.

UNEP: Standardized use of ToC together with a characterization of standard project performance ratings has improved the consistency of evaluative

determine whether all external factors affecting outcomes, impact, sustainability, and up-scaling have been carefully considered. This means that evaluation at UNEP not only reconstructs a ToC for the terminal evaluation, but the evaluation method uses the reconstructed ToC to assess effectiveness of the intervention as a guiding framework. However, it can also be used to help inform other evaluation criteria e.g. relevance and M&E. Of interest, the reconstructed ToC in the terminal evaluation may add intermediate outcomes that were missing in the logic in order to better assess the validity of the causal pathways.

WWF-US: The WWF network has a long history of applying WWF Project and Programme Management Standards, based off of the Conservation Measures Partnership (CMP) Conservation Standards. As such, a few tools are used to develop a strong theory of change, including situation analysis diagrams, threats ranking and results chains. These tools can be developed through a software tool (Miradi) that has been designed for use by WWF and other conservation organizations. Miradi helps project managers go through the steps of the WWF Programme Standards and can help to produce key outputs such as situation diagrams, results chains (theory of change), action plans, monitoring plans, etc.

In the context of GEF projects, two result-based management specialists serving an advisory role to WWF-US GEF support all GEF projects, as needed, with the development or review of situation diagrams and results chains during the planning stages of a project. Those two specialists also support GEF project teams to update the ToC during annual reflection and learning workshops.

judgment and the quality of evaluation reports. Specifically, use of ToC allows for a more robust, logical, and deductive assessment of the strengths or weaknesses in the causal pathways affecting the intervention. It also enables greater comparability, reliability, and candor in evaluation. Experience gained by the UNEP's independent evaluation office with use of ToC has been fed back into formal Agency project and program design requirements – leading to improvements in design quality.

WWF-US: The application of the CMP Conservation Standards, including the development of situation analysis diagrams and results chains has helped this GEF Agency with good project design, the selection of appropriate indicators for the result framework, and learning practices within projects. The long history of WWF-US developing results chains, in particular, for the development of ToC could be harnessed for the GEF partnership. Miradi Share, a web-based software tool and the Conservation Actions Measures (CAM) library are additional tools that can be utilized by the GEF partnership for project and program design, management, monitoring, learning and knowledge sharing.

Sources: Patton (2010), Interviews with: Stephen Hutton (IEG for the WBG), Michael Spilsbury (IEO of UNEP), Amelia Kissick (RBM Specialists, WWF-US)



Context:

“International Waters (IW): Learn is a unique niche in the GEF, it has been around long enough to prove itself as a viable and useful means, strengthening the performance of the focal area”
(Andrew Hudson)

IW: Learn was brought to light when the focal area and the GEF made a virtue out of necessity. In contrast to the remaining GEF focal areas, IW did not have an overarching convention or Conference of the Parties back in 1997. In the absence of an overarching convention, a vision of a global knowledge network for the focal area International Waters was elaborated by UNDP and met with positive resonance by the GEF. Ever since, IW: Learn has been financed and supported not only by the GEF, but an array of its Agencies and partners. As intended since its inception in 2000, IW: Learn is not only a convention of the focal area but serves as a dynamic platform for knowledge sharing and inter-project collaboration. The broad participatory foundation of IW: Learn is ensured by the mandatory inclusion of each project and by the unique niche and value proposition of the approach.

Good Practice:

IW: Learn hosts biennial conferences since 2000 as well as thematic and regional capacity building events gathering a wide range of stakeholders. Emphasizing smaller scale group discussions and going beyond scientific aspects allows for fruitful deep dives and the emergence of a ‘portfolio experience’. Over the years, IW: Learn’s website grew into a one-stop-shop for best practice, project data and knowledge products. The latter include experience notes providing hands-on support for preparing GEF project documents. Last but not least, IW: Learn facilitates inter-project collaboration and learning in the form of project twinning. This has been found to be highly effective for overcoming shared project management challenges and is enabled by matching similar projects.

Effects:

The various opportunities of face-to-face exchanges and the functionalities of IW: Learn’s website foster learning among project managers, country officials, Agencies, and other partners. While this already helps reduce silos within and between organizations, IW: Learn also integrates both formerly separated domains of freshwater and marine ecosystems. Through means of knowledge production and sharing around projects, results and best practice are made explicit and accessible, hence contributing to overall portfolio performance and learning in the GEF partnership.

Sources: Interview with Andrew Hudson

Practice Glimpse 11: UNIDO undertakes systematic ‘meta’ reviews for cross-project learning from TEs



Context:

Agencies Independent Evaluation Units (IEU) are in many cases balancing their resources between project evaluations and strategic evaluations. In order to satisfy accountability needs, GEF requirements for Agencies mainly focus on project evaluations. However, there is a lot of value for future project design and other learning needs stemming from a comparison of project lessons learned. These are called systematic or ‘meta’ reviews. However, in contrast to strategic evaluations, systematic reviews are so far done on a more ad-hoc basis, focus on a comparative perspective and draw insights around existing lessons learned.

Good Practice:

UNIDO conducted such a meta review on its projects in the area of renewable energy (RE)-based mini-grid projects in rural areas with the two-fold aim to review its experience as an institution and to evaluate the lessons learned in each project to refine the design of future projects. For the working paper “Renewable energy-based mini grids: The UNIDO Experience. 2017”, several projects and their evaluations were reviewed according to five key criteria determining long-term sustainability of mini-grid projects. The report was written in collaboration between staff from various departments further contributing to breaking silos for future project design.

Effects:

Various Agencies voiced that the potential of project evaluations is not fully realized for cross-project learning. However, generating cross-cutting lessons learned is an important step in feeding learnings from the project level, to project design and management as well as to the institutional level. Systematic reviews directly contribute to this goal and can help identify factors contributing to project impact, sustainability, and scalability. The exercise itself fosters learning between the IEU and project managers involved in synthesizing lessons learned from projects and their evaluations. However, in order to scale this good practice from an ad-hoc to a regular exercise, a dedicated resource allocation will be required.

Sources: Interviews with Rana Ghoneim, Division Chief in the Energy Department & Jossy Thomas (Industrial Development Officer, UNIDO); document review: UNIDO (2017)

Agency Profiles

1. Agency profile World Bank Group (WBG)

<p>purpose of self-evaluation</p> <ul style="list-style-type: none"> • accountability • learning <p>management support</p> <p>capacity building</p>		 WORLD BANK GROUP		
<p>relationship with independent evaluation</p> <ul style="list-style-type: none"> • independent evaluation unit <p>semi-independent evaluation unit</p> <p>manages TEs</p> <ul style="list-style-type: none"> • validates TEs 		<p>arrangements ensuring:</p>	<p>credibility and quality</p> <ul style="list-style-type: none"> • QA by GEF unit or program managers on PIR and MTR • internal peer review on TE <p>external consultation on TE</p> <ul style="list-style-type: none"> • field based/ technical reviews for a sample of TEs • updating ToC at regular intervals • use of ToC for evaluation and learning 	
<p>specific guidance provided on</p> <ul style="list-style-type: none"> • evaluation criteria and processes • quality assurance including role of independent evaluation office <p>preparation of MTR</p> <p>how to comply with GEF as step-by-step guide</p> <ul style="list-style-type: none"> • role of monitoring in evaluation <p>role of GEF core indicators in evaluation</p> <ul style="list-style-type: none"> • role of ToC in evaluation 			<p>information management</p> <ul style="list-style-type: none"> • aggregation and synthesis in annual reports • evaluation and learning database 	
			<p>Learning</p> <ul style="list-style-type: none"> • learning practices • thematic reviews <p>meta reviews</p> <ul style="list-style-type: none"> • impact evaluations 	

good practice example

Good Quality Assurance mechanisms
 Useful system for accountability and generating information for learning
 Extensive peer review opportunities

areas for improvement

More use of self-evaluation for learning, particularly, extracting and utilizing lessons learned to inform future decisions; improve the sharing of lessons learned

2. Agency profile European Bank for Reconstruction and Development (EBRD)

purpose of self-evaluation		
•	accountability	
•	learning	
	management support	
	capacity building	
relationship with independent evaluation		
•	independent evaluation unit	
	semi-independent evaluation unit	
	manages TEs	
•	validates TEs	
specific guidance provided on		
•	evaluation criteria and processes	
•	quality assurance including role of independent evaluation office	
	preparation of MTR	
	how to comply with GEF as step-by-step guide	
•	role of monitoring in evaluation	
	role of GEF core indicators in evaluation	
	role of ToC in evaluation	
arrangements ensuring:		
		• QA by GEF unit or program managers on PIR and MTR
		• internal peer review on TE
		external consultation on TE
		field based/ technical reviews for a sample of TEs
		updating ToC at regular intervals
		• use of ToC for evaluation and learning
		information management
		• aggregation and synthesis in annual reports
		• evaluation and learning database
		Learning
		• learning practices
		• thematic reviews
		meta reviews
		impact evaluations

good practice example

Pre-Validation Feedback to increase the efficiency of the current system

In the process of moving away from a focus on ratings as part of evaluation products

areas for improvement

Shifting responsibility of evaluation more to management

3. Agency profile Asian Development Bank (ADB)

<p>purpose of self-evaluation</p> <ul style="list-style-type: none"> • accountability • learning <p>management support</p> <p>capacity building</p>			
<p>relationship with independent evaluation</p> <ul style="list-style-type: none"> • independent evaluation unit <p>semi-independent evaluation unit</p> <p>manages TEs</p> <ul style="list-style-type: none"> • validates TEs 			<p>arrangements ensuring:</p>
<p>specific guidance provided on</p> <ul style="list-style-type: none"> • evaluation criteria and processes • quality assurance including role of independent evaluation office <p>preparation of MTR</p> <ul style="list-style-type: none"> • how to comply with GEF as step-by-step guide • role of monitoring in evaluation <p>role of GEF core indicators in evaluation</p> <p>role of ToC in evaluation</p>			<p>credibility and quality</p> <ul style="list-style-type: none"> • QA by GEF unit or program managers on PIR and MTR • internal peer review on TE <p>external consultation on TE</p> <ul style="list-style-type: none"> • field based/ technical reviews for a sample of TEs <p>updating ToC at regular intervals</p> <ul style="list-style-type: none"> • use of ToC for evaluation and learning
		<p>information management</p> <ul style="list-style-type: none"> • aggregation and synthesis in annual reports • evaluation and learning database 	
		<p>Learning</p> <ul style="list-style-type: none"> • learning practices • thematic reviews <p>meta reviews</p> <ul style="list-style-type: none"> • impact evaluation 	

good practice example

Good Quality Assurance mechanisms
Useful system for accountability and generating information for learning

areas for improvement

Project monitoring and evaluation (M&E) systems during implementation
Improve the sharing of lessons learned

4. Agency profile Inter-American Development Bank (IADB)

purpose of self-evaluation	<ul style="list-style-type: none"> • accountability 																							
<ul style="list-style-type: none"> • learning 	management support																							
<ul style="list-style-type: none"> • learning 	capacity building	<table border="1"> <tr> <td data-bbox="743 524 932 600">arrangements ensuring:</td> <td data-bbox="932 524 1437 600">credibility and quality</td> </tr> <tr> <td data-bbox="743 600 932 676"> <ul style="list-style-type: none"> • independent evaluation unit </td> <td data-bbox="932 600 1437 676"> <ul style="list-style-type: none"> • QA by GEF unit or program managers on PIR and MTR </td> </tr> <tr> <td data-bbox="743 676 932 752"> <ul style="list-style-type: none"> • validates TEs </td> <td data-bbox="932 676 1437 752"> <ul style="list-style-type: none"> • internal peer review on TE </td> </tr> <tr> <td data-bbox="743 752 932 828"> <ul style="list-style-type: none"> • evaluation criteria and processes </td> <td data-bbox="932 752 1437 828"> <ul style="list-style-type: none"> • information management </td> </tr> <tr> <td data-bbox="743 828 932 904"> <ul style="list-style-type: none"> • quality assurance including role of independent evaluation office </td> <td data-bbox="932 828 1437 904"> <ul style="list-style-type: none"> • aggregation and synthesis in annual reports </td> </tr> <tr> <td data-bbox="743 904 932 981"> <ul style="list-style-type: none"> • how to comply with GEF as step-by-step guide </td> <td data-bbox="932 904 1437 981"> <ul style="list-style-type: none"> • evaluation and learning database </td> </tr> <tr> <td data-bbox="743 981 932 1057"> <ul style="list-style-type: none"> • role of monitoring in evaluation </td> <td data-bbox="932 981 1437 1057"> <ul style="list-style-type: none"> • learning </td> </tr> <tr> <td data-bbox="743 1057 932 1133"> <ul style="list-style-type: none"> • role of GEF core indicators in evaluation </td> <td data-bbox="932 1057 1437 1133"> <ul style="list-style-type: none"> • learning practices </td> </tr> <tr> <td data-bbox="743 1133 932 1209"> <ul style="list-style-type: none"> • role of ToC in evaluation </td> <td data-bbox="932 1133 1437 1209"> <ul style="list-style-type: none"> • thematic reviews </td> </tr> <tr> <td data-bbox="743 1209 932 1285"> <ul style="list-style-type: none"> • role of ToC in evaluation </td> <td data-bbox="932 1209 1437 1285"> <ul style="list-style-type: none"> • meta reviews </td> </tr> <tr> <td data-bbox="743 1285 932 1361"> <ul style="list-style-type: none"> • role of ToC in evaluation </td> <td data-bbox="932 1285 1437 1361"> <ul style="list-style-type: none"> • impact evaluation </td> </tr> </table>	arrangements ensuring:	credibility and quality	<ul style="list-style-type: none"> • independent evaluation unit 	<ul style="list-style-type: none"> • QA by GEF unit or program managers on PIR and MTR 	<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • internal peer review on TE 	<ul style="list-style-type: none"> • evaluation criteria and processes 	<ul style="list-style-type: none"> • information management 	<ul style="list-style-type: none"> • quality assurance including role of independent evaluation office 	<ul style="list-style-type: none"> • aggregation and synthesis in annual reports 	<ul style="list-style-type: none"> • how to comply with GEF as step-by-step guide 	<ul style="list-style-type: none"> • evaluation and learning database 	<ul style="list-style-type: none"> • role of monitoring in evaluation 	<ul style="list-style-type: none"> • learning 	<ul style="list-style-type: none"> • role of GEF core indicators in evaluation 	<ul style="list-style-type: none"> • learning practices 	<ul style="list-style-type: none"> • role of ToC in evaluation 	<ul style="list-style-type: none"> • thematic reviews 	<ul style="list-style-type: none"> • role of ToC in evaluation 	<ul style="list-style-type: none"> • meta reviews 	<ul style="list-style-type: none"> • role of ToC in evaluation 	<ul style="list-style-type: none"> • impact evaluation
arrangements ensuring:	credibility and quality																							
<ul style="list-style-type: none"> • independent evaluation unit 	<ul style="list-style-type: none"> • QA by GEF unit or program managers on PIR and MTR 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • internal peer review on TE 																							
<ul style="list-style-type: none"> • evaluation criteria and processes 	<ul style="list-style-type: none"> • information management 																							
<ul style="list-style-type: none"> • quality assurance including role of independent evaluation office 	<ul style="list-style-type: none"> • aggregation and synthesis in annual reports 																							
<ul style="list-style-type: none"> • how to comply with GEF as step-by-step guide 	<ul style="list-style-type: none"> • evaluation and learning database 																							
<ul style="list-style-type: none"> • role of monitoring in evaluation 	<ul style="list-style-type: none"> • learning 																							
<ul style="list-style-type: none"> • role of GEF core indicators in evaluation 	<ul style="list-style-type: none"> • learning practices 																							
<ul style="list-style-type: none"> • role of ToC in evaluation 	<ul style="list-style-type: none"> • thematic reviews 																							
<ul style="list-style-type: none"> • role of ToC in evaluation 	<ul style="list-style-type: none"> • meta reviews 																							
<ul style="list-style-type: none"> • role of ToC in evaluation 	<ul style="list-style-type: none"> • impact evaluation 																							
relationship with independent evaluation	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							
<ul style="list-style-type: none"> • validates TEs 	<ul style="list-style-type: none"> • independent evaluation unit 																							

good practice example

Good Quality Assurance mechanisms
 Useful system for accountability and generating information for learning
 Office of Strategic Planning and Development Effectiveness (SPD) that supports project teams with M&E throughout the project cycle.
 IADB frequently conducts impact evaluations for projects

areas for improvement

More use of self-evaluation for learning between its IEU, project staff, stakeholders, and beneficiaries.

5. Agency profile African Development Bank (AfDB)

purpose of self-evaluation	<ul style="list-style-type: none"> • accountability 	
<ul style="list-style-type: none"> • learning 	management support	
<ul style="list-style-type: none"> • learning 	management support	

<table border="1"> <tr> <td data-bbox="193 174 248 215"></td> <td data-bbox="248 174 695 215">capacity building</td> </tr> <tr> <td colspan="2" data-bbox="193 215 695 456"></td> </tr> </table>		capacity building																										
	capacity building																											
<table border="1"> <tr> <th colspan="2" data-bbox="193 456 695 533">relationship with independent evaluation</th> </tr> <tr> <td data-bbox="193 533 248 607">•</td> <td data-bbox="248 533 695 607">independent evaluation unit</td> </tr> <tr> <td data-bbox="193 607 248 647"></td> <td data-bbox="248 607 695 647">semi-independent evaluation unit</td> </tr> <tr> <td data-bbox="193 647 248 687"></td> <td data-bbox="248 647 695 687">manages TEs</td> </tr> <tr> <td data-bbox="193 687 248 763">•</td> <td data-bbox="248 687 695 763">validates TEs</td> </tr> </table>	relationship with independent evaluation		•	independent evaluation unit		semi-independent evaluation unit		manages TEs	•	validates TEs	arrangements ensuring:	<table border="1"> <tr> <th colspan="2" data-bbox="963 456 1437 533">credibility and quality</th> </tr> <tr> <td data-bbox="963 533 1007 607">•</td> <td data-bbox="1007 533 1437 607">QA by GEF unit or program managers on PIR and MTR</td> </tr> <tr> <td data-bbox="963 607 1007 647">•</td> <td data-bbox="1007 607 1437 647">internal peer review on TE</td> </tr> <tr> <td data-bbox="963 647 1007 687"></td> <td data-bbox="1007 647 1437 687">external consultation on TE</td> </tr> <tr> <td data-bbox="963 687 1007 763"></td> <td data-bbox="1007 687 1437 763">field based/ technical reviews for a sample of TEs</td> </tr> <tr> <td data-bbox="963 763 1007 804"></td> <td data-bbox="1007 763 1437 804">updating ToC at regular intervals</td> </tr> <tr> <td data-bbox="963 804 1007 913">•</td> <td data-bbox="1007 804 1437 913">use of ToC for evaluation and learning</td> </tr> <tr> <td colspan="2" data-bbox="963 913 1437 958"></td> </tr> </table>	credibility and quality		•	QA by GEF unit or program managers on PIR and MTR	•	internal peer review on TE		external consultation on TE		field based/ technical reviews for a sample of TEs		updating ToC at regular intervals	•	use of ToC for evaluation and learning		
relationship with independent evaluation																												
•	independent evaluation unit																											
	semi-independent evaluation unit																											
	manages TEs																											
•	validates TEs																											
credibility and quality																												
•	QA by GEF unit or program managers on PIR and MTR																											
•	internal peer review on TE																											
	external consultation on TE																											
	field based/ technical reviews for a sample of TEs																											
	updating ToC at regular intervals																											
•	use of ToC for evaluation and learning																											
<table border="1"> <tr> <th colspan="2" data-bbox="193 763 695 840">specific guidance provided on</th> </tr> <tr> <td data-bbox="193 840 248 949">•</td> <td data-bbox="248 840 695 949">evaluation criteria and processes</td> </tr> <tr> <td data-bbox="193 949 248 1025">•</td> <td data-bbox="248 949 695 1025">quality assurance including role of independent evaluation office</td> </tr> <tr> <td data-bbox="193 1025 248 1066"></td> <td data-bbox="248 1025 695 1066">preparation of MTR</td> </tr> <tr> <td data-bbox="193 1066 248 1142"></td> <td data-bbox="248 1066 695 1142">how to comply with GEF as step-by-step guide</td> </tr> <tr> <td data-bbox="193 1142 248 1218">•</td> <td data-bbox="248 1142 695 1218">role of monitoring in evaluation</td> </tr> <tr> <td data-bbox="193 1218 248 1294"></td> <td data-bbox="248 1218 695 1294">role of GEF core indicators in evaluation</td> </tr> <tr> <td data-bbox="193 1294 248 1346"></td> <td data-bbox="248 1294 695 1346">role of ToC in evaluation</td> </tr> </table>	specific guidance provided on		•	evaluation criteria and processes	•	quality assurance including role of independent evaluation office		preparation of MTR		how to comply with GEF as step-by-step guide	•	role of monitoring in evaluation		role of GEF core indicators in evaluation		role of ToC in evaluation	<table border="1"> <tr> <th colspan="2" data-bbox="963 1025 1437 1066">information management</th> </tr> <tr> <td data-bbox="963 1066 1007 1142">•</td> <td data-bbox="1007 1066 1437 1142">aggregation and synthesis in annual reports</td> </tr> <tr> <td data-bbox="963 1142 1007 1218">•</td> <td data-bbox="1007 1142 1437 1218">evaluation and learning database</td> </tr> </table>	information management		•	aggregation and synthesis in annual reports	•	evaluation and learning database					
specific guidance provided on																												
•	evaluation criteria and processes																											
•	quality assurance including role of independent evaluation office																											
	preparation of MTR																											
	how to comply with GEF as step-by-step guide																											
•	role of monitoring in evaluation																											
	role of GEF core indicators in evaluation																											
	role of ToC in evaluation																											
information management																												
•	aggregation and synthesis in annual reports																											
•	evaluation and learning database																											
	<table border="1"> <tr> <th colspan="2" data-bbox="963 1218 1437 1294">learning</th> </tr> <tr> <td data-bbox="963 1294 1007 1335">•</td> <td data-bbox="1007 1294 1437 1335">learning practices</td> </tr> <tr> <td data-bbox="963 1335 1007 1375">•</td> <td data-bbox="1007 1335 1437 1375">thematic reviews</td> </tr> <tr> <td data-bbox="963 1375 1007 1415"></td> <td data-bbox="1007 1375 1437 1415">meta reviews</td> </tr> <tr> <td data-bbox="963 1415 1007 1458">•</td> <td data-bbox="1007 1415 1437 1458">impact evaluation</td> </tr> </table>	learning		•	learning practices	•	thematic reviews		meta reviews	•	impact evaluation																	
learning																												
•	learning practices																											
•	thematic reviews																											
	meta reviews																											
•	impact evaluation																											

good practice example

Extensive peer review opportunities as part of the Quality Assurance Framework

areas for improvement

More use of self-evaluation for learning between its IEU, project staff, stakeholders, and beneficiaries. Shifting current focus on ratings and disconnect of ratings between project teams and IEU more towards learning

6. Agency profile International Fund for Agricultural Development (IFAD)

purpose of self-evaluation		
•	accountability	
•	learning	
	management support	
	capacity building	
relationship with independent evaluation		
•	independent evaluation unit	
	semi-independent evaluation unit	
	manages TEs	
•	validates TEs	
specific guidance provided on		
•	evaluation criteria and processes	
•	quality assurance including role of independent evaluation office	
	preparation of MTR	
	how to comply with GEF as step-by-step guide	
•	role of monitoring in evaluation	
	role of GEF core indicators in evaluation	
•	role of ToC in evaluation	
arrangements ensuring:		
		• QA by GEF unit or program managers on PIR and MTR
		• internal peer review on TE
		external consultation on TE
		field based/ technical reviews for a sample of TEs
		updating ToC at regular intervals
		• use of ToC for evaluation and learning
		information management
		• aggregation and synthesis in annual reports
		evaluation and learning database
		learning
		• learning practices
		• thematic reviews
		meta reviews
		• impact evaluation

good practice example

Advanced data capture as part of its SES, with a range of indicators for rural poverty on all its project. A Development Effectiveness Unit dedicated in applying SE results for informing project design.

areas for improvement

More use of self-evaluation for learning between its IEU, project staff, stakeholders, and beneficiaries.

7. Agency profile United Nations Environment Programme (UNEP)

<p>purpose of self-evaluation</p> <ul style="list-style-type: none"> • accountability • learning <p>management support</p> <p>capacity building</p>		
<p>relationship with independent evaluation</p> <ul style="list-style-type: none"> • independent evaluation unit <p>semi-independent evaluation unit</p> <p>manages TEs</p> <ul style="list-style-type: none"> • validates TEs 		<p>arrangements ensuring:</p>
<p>specific guidance provided on</p> <ul style="list-style-type: none"> • evaluation criteria and processes • quality assurance including role of independent evaluation office • preparation of MTR <p>how to comply with GEF as step-by-step guide</p> <ul style="list-style-type: none"> • role of monitoring in evaluation <p>role of GEF core indicators in evaluation</p> <ul style="list-style-type: none"> • role of ToC in evaluation 	<p>information management</p> <ul style="list-style-type: none"> • aggregation and synthesis in annual reports • evaluation and learning database 	
	<p>learning</p> <ul style="list-style-type: none"> • learning practices • thematic reviews <p>meta reviews</p> <ul style="list-style-type: none"> • impact evaluation 	

good practice example

Very robust guidance for evaluation (see Annex A.2 on policy documents)
 Very robust peer reviews for evaluation

areas for improvement

More use of self-evaluation for learning between its IEU, project staff, stakeholders, and beneficiaries.

8. Agency profile United Nations Development Programme (UNDP)

<p>purpose of self-evaluation</p> <ul style="list-style-type: none"> • accountability • learning <p>management support</p> <ul style="list-style-type: none"> • capacity building 		
<p>relationship with independent evaluation</p> <ul style="list-style-type: none"> • independent evaluation unit <p>semi-independent evaluation unit</p> <p>manages TEs</p> <p>validates TEs</p>		<p>arrangements ensuring:</p>
<p>specific guidance provided on</p> <ul style="list-style-type: none"> • evaluation criteria and processes • quality assurance including role of independent evaluation office • preparation of MTR <p>how to comply with GEF as step-by-step guide</p> <ul style="list-style-type: none"> • role of monitoring in evaluation <p>role of GEF core indicators in evaluation</p> <p>role of ToC in evaluation</p>	<p>information management</p> <ul style="list-style-type: none"> • aggregation and synthesis in annual reports • evaluation and learning database 	
	<p>learning</p> <p>learning practices</p> <ul style="list-style-type: none"> • thematic reviews <p>meta reviews</p> <p>impact evaluations</p>	

good practice example

Excellent collaborative PIR reporting which adds value to collaboration, evidence, and oversight.

areas for improvement

Lack of validation by UNDP's IEU before GEF submission may compromise quality and consistency of reports.
 Lack of involvement of UNDP's IEU in managing SE also influence impartiality of SE.
 More capacity building and investment in the use of ToC for understanding project's causal pathways.
 More use of self-evaluation for learning between its IEU, project staff, stakeholders, and beneficiaries.

9. Agency profile Food and Agriculture Organization (FAO)

purpose of self-evaluation	<ul style="list-style-type: none"> • accountability 			Food and Agriculture Organization of the United Nations
<ul style="list-style-type: none"> • learning 	management support			
	capacity building	arrangements ensuring:	credibility and quality	<ul style="list-style-type: none"> • QA by GEF unit or program managers on PIR and MTR
relationship with independent evaluation	<ul style="list-style-type: none"> • independent evaluation unit 			<ul style="list-style-type: none"> • internal peer review on TE
	semi-independent evaluation unit		<ul style="list-style-type: none"> • external consultation on TE 	
<ul style="list-style-type: none"> • manages TEs 	validates TEs		field based/ technical reviews for a sample of TEs	
			updating ToC at regular intervals	
specific guidance provided on	<ul style="list-style-type: none"> • evaluation criteria and processes 		<ul style="list-style-type: none"> • use of ToC for evaluation and learning 	
<ul style="list-style-type: none"> • quality assurance including role of independent evaluation office 	<ul style="list-style-type: none"> • preparation of MTR 		field based/ technical reviews for a sample of TEs	
<ul style="list-style-type: none"> • how to comply with GEF as step-by-step guide 	<ul style="list-style-type: none"> • role of monitoring in evaluation 		updating ToC at regular intervals	
<ul style="list-style-type: none"> • role of GEF core indicators in evaluation 	<ul style="list-style-type: none"> • role of ToC in evaluation 		<ul style="list-style-type: none"> • use of ToC for evaluation and learning 	
			<ul style="list-style-type: none"> • use of ToC for evaluation and learning 	
		field based/ technical reviews for a sample of TEs		
		updating ToC at regular intervals		
		information management		
		<ul style="list-style-type: none"> • aggregation and synthesis in annual reports 		
		<ul style="list-style-type: none"> • evaluation and learning database 		
		<ul style="list-style-type: none"> • learning practices 		
		<ul style="list-style-type: none"> • thematic reviews 		
		<ul style="list-style-type: none"> • meta reviews 		
		<ul style="list-style-type: none"> • impact evaluation 		

good practice example

Inclusion of reference groups with preparation and dissemination of TEs

areas for improvement

better aggregation of learnings from TEs to feed into high-level decision-making

10. Agency profile United Nations Industrial Development Organization (UNIDO)

purpose of self-evaluation				
• accountability				
• learning				
management support				
capacity building				
relationship with independent evaluation			arrangements ensuring:	credibility and quality
• independent evaluation unit			• QA by GEF unit or program managers on PIR and MTR	
semi-independent evaluation unit			• internal peer review on TE	
• manages TEs			• external consultation on TE	
validates TEs			field based/ technical reviews for a sample of TEs	
specific guidance provided on		updating ToC at regular intervals		
• evaluation criteria and processes		• use of ToC for evaluation and learning		
• quality assurance including role of independent evaluation office		information management		
• preparation of MTR		• aggregation and synthesis in annual reports		
• how to comply with GEF as step-by-step guide		evaluation and learning database		
• role of monitoring in evaluation		learning		
role of GEF core indicators in evaluation		• learning practices		
• role of ToC in evaluation		• thematic reviews		
		• meta reviews		
		• impact evaluation		

good practice example

Meta reviews comparing and synthesizing lessons from TEs
 Debriefs on project evaluations with project managers, country offices and evaluators

areas for improvement

Better feedback loops between project delay and optimal timing of MTRs for learning and compliance

11. Agency profile Foreign Economic Cooperation Office, Ministry of Environmental Protection of China (FECO)

<p>purpose of self-evaluation</p> <ul style="list-style-type: none"> • accountability • learning <p>management support</p> <p>capacity building</p>		
<p>relationship with independent evaluation</p> <ul style="list-style-type: none"> • independent evaluation unit <p>semi-independent evaluation unit</p> <ul style="list-style-type: none"> • manages TEs <p>validates TEs</p>		
<p>specific guidance provided on</p> <ul style="list-style-type: none"> • evaluation criteria and processes • quality assurance including role of independent evaluation office • preparation of MTR • how to comply with GEF as step-by-step guide • role of monitoring in evaluation • role of GEF core indicators in evaluation <p>role of ToC in evaluation</p>	<p>arrangements ensuring:</p>	<p>credibility and quality</p> <ul style="list-style-type: none"> • QA by GEF unit or program managers on PIR and MTR • internal peer review on TE • external consultation on TE <p>field based/ technical reviews for a sample of TEs</p> <p>updating ToC at regular intervals</p> <ul style="list-style-type: none"> • use of ToC for evaluation and learning
<p>information management</p> <ul style="list-style-type: none"> • aggregation and synthesis in annual reports • evaluation and learning database 		
<p>learning</p> <ul style="list-style-type: none"> • learning practices <p>thematic reviews</p> <p>meta reviews</p> <p>impact evaluation</p>		

good practice example

Annual supervision missions to projects fosters implementation and data collection
 Extensive guidance following GEF requirements

areas for improvement

More use of self-evaluation for learning between its IEU, project staff, stakeholders, and beneficiaries.

12. Agency profile Development Bank of Latin America (CAF)

purpose of self-evaluation		
•	accountability	
•	learning	
	management support	
	capacity building	
relationship with independent evaluation		
•	independent evaluation unit	
	semi-independent evaluation unit	
•	manages TEs	
	validates TEs	
specific guidance provided on		
•	evaluation criteria and processes	
	quality assurance including role of independent evaluation office	
	preparation of MTR	
	how to comply with GEF as step-by-step guide	
	role of monitoring in evaluation	
	role of GEF core indicators in evaluation	
	role of ToC in evaluation	
arrangements ensuring:		
		• QA by GEF unit or program managers on PIR and MTR
		internal peer review on TE
		external consultation on TE
		field based/ technical reviews for a sample of TEs
		updating ToC at regular intervals
		use of ToC for evaluation and learning
		information management
		aggregation and synthesis in annual reports
		• evaluation and learning database
		learning
		learning practices
		• thematic reviews
		meta reviews
		impact evaluation

good practice example

Continuous reviewing of policies and guidelines to align with GEF partnership standards

areas for improvement

Need for more developed guidelines, QA processes, information management and learning practices.

13. Agency profile Development Bank of Southern Africa (DBSA)

purpose of self-evaluation			
	accountability		
	learning		
	management support		
	capacity building		
relationship with independent evaluation			
•	independent evaluation unit		
	semi-independent evaluation unit		
•	manages TEs		
	validates TEs		
specific guidance provided on		arrangements ensuring:	
	evaluation criteria and processes		
	quality assurance including role of independent evaluation office		
	preparation of MTR		
	how to comply with GEF as step-by-step guide		
	role of monitoring in evaluation		
	role of GEF core indicators in evaluation		
	role of ToC in evaluation		
			credibility and quality
•	QA by GEF unit or program managers on PIR and MTR		
•	internal peer review on TE		
	external consultation on TE		
	field based/ technical reviews for a sample of TEs		
	updating ToC at regular intervals		
	use of ToC for evaluation and learning		
		information management	
	aggregation and synthesis in annual reports		
	evaluation and learning database		
		learning	
	learning practices		
	thematic reviews		
	meta reviews		
	impact evaluation		

good practice example

Continuous improvement through Development Result Working Group

areas for improvement

More alignment and cooperation between independent evaluation unit and GEF coordination unit

14. Agency profile International Union for Conservation of Nature (IUCN)

purpose of self-evaluation	arrangements ensuring:		
• accountability			
• learning			
• management support			
capacity building			
relationship with independent evaluation			credibility and quality
independent evaluation unit			• QA by GEF unit or program managers on PIR and MTR
• semi-independent evaluation unit			Internal peer review on TE
• manages TEs			External consultation on TE
validates TEs			Field based/ technical reviews for a sample of TEs
specific guidance provided on	Updating ToC at regular intervals		
• evaluation criteria and processes	• Use of ToC for evaluation and learning		
• quality assurance including role of independent evaluation office			
preparation of MTR	information management		
how to comply with GEF as step-by-step guide	• aggregation and synthesis in annual reports		
• role of monitoring in evaluation	evaluation and learning database		
role of GEF core indicators in evaluation	learning		
role of ToC in evaluation	learning practices		
	thematic reviews		
	meta reviews		
	impact evaluation		

good practice example

Only NGO that aggregates self-evaluation results for management learning

areas for improvement

More internal and external peer reviews
 More use of self-evaluation for learning between its semi-independent unit, project staff, stakeholders, and beneficiaries.

15. Agency profile World Wildlife Fund - US (WWF-US)

purpose of self-evaluation	arrangements ensuring:		
• accountability			
• learning			
• management support			
capacity building			
relationship with independent evaluation			credibility and quality
independent evaluation unit			• QA by GEF unit or program managers on PIR and MTR
• semi-independent evaluation unit			internal peer review on TE
• manages TEs			external consultation on TE
validates TEs			field based/ technical reviews for a sample of TEs
specific guidance provided on	• updating ToC at regular intervals		
• evaluation criteria and processes	use of ToC for evaluation and learning		
• quality assurance including role of independent evaluation office			
preparation of MTR	information management		
how to comply with GEF as step-by-step guide	aggregation and synthesis in annual reports		
• role of monitoring in evaluation	evaluation and learning database		
role of GEF core indicators in evaluation	learning		
role of ToC in evaluation	learning practices		
	thematic reviews		
	meta reviews		
	impact evaluation		

good practice example

WWF-US has a long tradition of using a result chain framework as methodology for ToC

areas for improvement

More use of self-evaluation for learning between its IEU, project staff, stakeholders, and beneficiaries.
More peer learning from other Agencies with more advanced SES

16. Agency profile Conservation International (CI)

purpose of self-evaluation	 <p>CONSERVATION INTERNATIONAL</p>
<ul style="list-style-type: none"> • accountability 	
<ul style="list-style-type: none"> • learning 	arrangements ensuring:
management support	
capacity building	
relationship with independent evaluation	
independent evaluation unit	
semi-independent evaluation unit	
manages TEs	
validates TEs	
specific guidance provided on	
<ul style="list-style-type: none"> • evaluation criteria and processes 	
<ul style="list-style-type: none"> • quality assurance including role of independent evaluation office 	
preparation of MTR	
how to comply with GEF as step-by-step guide	
role of monitoring in evaluation	
role of GEF core indicators in evaluation	
role of ToC in evaluation	
credibility and quality	
<ul style="list-style-type: none"> • QA by GEF unit or program managers on PIR and MTR 	
internal peer review on TE	
external consultation on TE	
field based/ technical reviews for a sample of TEs	
updating ToC at regular intervals	
use of ToC for evaluation and learning	
information management	
aggregation and synthesis in annual reports	
evaluation and learning database	
learning	
learning practices	
thematic reviews	
meta reviews	
impact evaluation	

good practice example

Being a GEF agency is helping CI develop a SES, not only for GEF projects, but possibly for all projects.

areas for improvement

More use of self-evaluation for learning between its IEU, project staff, stakeholders, and beneficiaries.
More peer learning from other Agencies with more advanced SES

Annex A.2 Summary tables of Agency Self Evaluation System Characteristics

TABLE 1: SUMMARIZED DATASET GUIDANCE PROVIDED BY AGENCIES

<i>Guidance given on</i>							
<i>Agencies</i>	Evaluation				Monitoring		
	Evaluation criteria and processes	Quality assurance including role of IEU	Specific guidance for MTR	Guidance how to comply with GEF	Role of monitoring in evaluation	Role of GEF core indicators in evaluation	Role of ToC in evaluation
<i>WBG</i>	x	x			x		x
<i>ADB</i>	x	x		x ¹	x		
<i>IADB</i>	x	x		x ¹	x		
<i>AfDB</i>	x	x			x		
<i>EBRD</i>	x	x			x		
<i>CAF</i>	x ⁶	x		x ⁴	x	x ³	
<i>DBSA</i>	NA	NA	NA	NA	NA	NA	NA
<i>BOAD</i>	NA	NA	NA	NA	NA	NA	NA
<i>FAO</i>	x	x	x	x ^{1, 2}	x ⁶	x ⁷	x
<i>IFAD</i>	x	x			x		x
<i>UNDP</i>	x	x	x	x ²	x	x	
<i>UNEP</i>	x	x	x	x ²			x
<i>UNIDO</i>	x	x	x	x ^{1,2}	x		x
<i>IUCN</i>	x	x		x ²	x		
<i>CI</i>	x	x	x	x ²			
<i>WWF-US</i>	x	x	x	x ²	x		
<i>FECO</i>	x	x	x	x ¹	x	x ⁷	
<i>FUNBIO</i>	NA	NA	NA	NA	NA	NA	NA

NA No data collected

1 Step by step guide (explicit for GEF)

2 Some GEF requirements internalised within Agencies guidelines for TE, MTR, PIR2

3 CAF references a table of core indicators & a methodology for the selection of sub-indicators

4 Only some process details for PIR as part of project cycle guidance; none for TE and MTR as they are managed under independent evaluation function and/or external evaluators

5 Processes are outlined, but evaluation criteria for independent evaluations are not fully in line with OECD-DAC standards (4 criteria used by GEF)

6 IFAD measures several indicators for rural poverty impact (household income and assets, human and social capital and empowerment, food security and agricultural productivity, natural resources and the environment, institutions and policies)

7 Some basic guidance referring to core indicators and where to find GEF guidance on it

TABLE 2: SUMMARISED DATASET QUALITY ASSURANCE PROCEDURES IN AGENCIES

Agencies	Part of Project and Program Management		Part of self-evaluation system									
	Quality control by GEF unit or Agencies' program/ portfolio managers	Updating ToC at regular intervals	Validation by independent or semi-independent unit			Internal peer review process			External Consultation on TE	Use of ToC for evaluation/ learning	Field based/ technical reviews for a sample	Impact Evaluations
			TE	MTR	PIR	TE	MTR	PIR				
WBG	x	x	x			x ^{2,6}				x	x	
ADB	x		x			x ⁶				x	x	x
IADB	x		x			x ^{1,6}				No		x
AfDB	x		x			x ²				x		x
EBRD	x		x			x ³				Yes		
CAF	x	NA	NA	NA	NA	NA		NA		NA	NA	
DBSA	NA		x	NA	NA	NA	NA	NA		NA	NA	
BOAD	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
FAO	x	x	x	x ⁴	No	x	x ⁵	No	x ⁷	Yes		x
IFAD			x			x						x
UNDP	x	x	No ⁹	No		x	x	x	x ⁸			
UNEP	x		x			x ⁶			x ⁷	x		x
UNIDO	x		x			x			x	x		x
IUCN	x									x		
CI	x											
WWF-US	x	x	x	x	x					x		
FECO	x	NA	x						x ⁷	NA	NA	
FUNBIO	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	

NA No data collected

1 Office of Strategic Planning and Development Effectiveness (SPD)

2 Peer reviews including other units within the organization (e.g., sectoral manager, country/ regional teams and other peer reviewers)

3 Pre-validation of reports through IEO before submission

4 Only upon request by GEF CU

5 Peer review is 'only' highly recommended but is under budget holder's responsibility; Independent evaluation office provides comments on top

6 Peer reviews within IEU during validation of TEs

7 optional consultations with external stakeholders such as Operational Focal Points, national counterparts, and civil society organizations, e.g. on preliminary findings

8 discussions on preliminary findings

9 UNDP IEU only validates in retrospect to the completed and cleared TEs. Its independent review processes as well as coordination of the TE process is achieved through M&E specialists at its Commissioning Units

TABLE 3: SUMMARISED DATASET INFORMATION AND KNOWLEDGE MANAGEMENT AND LEARNING PRACTICES ON SELF-EVALUATION.

<i>Agencies</i>	Aggregating, synthesis, databasing of self-evaluation results by independent evaluation office		Learning practices in place for learning from self-evaluation
	Aggregation and synthesis	Databasing	
<i>WBG</i>	x	x	x
<i>ADB</i>	x	x	x
<i>IADB</i>	x	x	x
<i>AfDB</i>	x	x	x
<i>EBRD</i>	x	x	x
<i>CAF</i>	NA	x	x
<i>DBSA</i>	No		No ¹
<i>BOAD</i>	NA	NA	NA
<i>FAO</i>	x	x	x
<i>IFAD</i>	x		x
<i>UNDP</i>	x	x	x
<i>UNEP</i>	x		
<i>UNIDO</i>	x		x
<i>IUCN</i>	x		
<i>CI</i>			
<i>WWF-US</i>			X
<i>FECO</i>	NA	x	x
<i>FUNBIO</i>	NA	NA	NA

NA No data collected

¹ none apart from submitting evaluation reports to the Board, Strategy Dep and Communication Dep of DBSA

TABLE 4: SUMMARIZED DATASET TE COMPARISON

TE Comparison

<i>Agencies</i>	<i>Effectiveness</i>		<i>Efficiency</i>		<i>Relevance</i>		<i>Sustainability</i>		<i>Overall</i>
	Assessed & Rated	Aligned with GEF rating scale	Assessed & Rated	Aligned with GEF rating scale	Assessed & Rated	Aligned with GEF rating scale	Assessed & Rated	Aligned with GEF rating scale	Overall outcome rating in line with GEF
<i>WBG</i>	x ¹	No	x	No	x	No	No ⁴	-	x
<i>ADB</i>	x	No	x	No	x	No	x	No	No ^{6,7}
<i>IADB</i>	x	x	x	x	x	x	x	x	x
<i>AfDB</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>EBRD</i>	No ²	x	No ²	x	No	-	x	No ⁵	No
<i>CAF</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>DBSA</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>BOAD</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>FAO</i>	x	x	x	x	x	x	x	x	x
<i>IFAD</i>	x	x	x	x	x	x	x	x	No ^{6,7}
<i>UNDP</i>	x	x	x	x	x	x	x	x	x
<i>UNEP</i>	x	x	x	x	x	x	x	No ⁵	No ⁶
<i>UNIDO</i>	x	x	x	x	x	x	x	x	x
<i>IUCN</i>	x	x	x	x	x	x	x	x	No ⁸
<i>CI</i>	x	x	x	x	x	x	x	x	x
<i>WWF-US</i>	No ³	-	No ³	-	No ³	-	x	x	No ⁶
<i>FECO</i>	NA	NA	NA	NA	NA	NA	NA	NA	No
<i>FUNBIO</i>	NA	NA	NA	NA	NA	NA	NA	NA	No

NA No data collected

1 Assesses and rates efficacy instead of effectiveness

2 Rating of subchapters/components but no indication of an overall rating for the criterion

3 Effectiveness, efficiency and relevance are not rated separately

4 Assesses and rates risk to development outcome instead of sustainability

5 Uses GEF 6-point rating scale for sustainability

6 Overall outcome rating includes more criteria than effectiveness, efficiency, and relevance. Is not true only based on those three criteria.

7 Outcome rating is not in line with the proposed GEF rating scale

8 Does not indicate an overall outcome rating

Annexes - Part B. Evaluation of the GEF Portal

Annex B.1 List of interviews conducted

S. No.	Name	Title/Affiliation
GEF Secretariat		
1.	Peter Lallas	Advisor
2.	Henry Salazar	Senior Operations Officer
3.	Quynh Xuan Thi Phan	Senior Financial Officer
4.	Claude Gascon	Manager GEF Programs
5.	Ulrich Apel	Senior Environmental Specialist
6.	Leah Bunce Karrer	Senior Environmental Specialist
7.	Jean Marc Sinnassamy	Senior Environmental Specialist
8.	Sarah Wyatt	Biodiversity Specialist
9.	Fareeha Iqbal	Senior Climate Change Specialist
10.	Mark Zimsky	Senior Biodiversity Specialist
11.	Christian Holde Severin	Senior Environmental Specialist
12.	Cyril Blet	Senior Specialist RBM
13.	Omid Parhizkar	Operations Officer
14.	Naying Peng	Operations Analyst
15.	William Ernest Ehlers	Senior Country Officer
GEF IEO		
16.	Peixuan Zhou	Evaluation Analyst
17.	Molly Watts Sohn	Evaluation Analyst
18.	Sara Choufi	Evaluation Analyst
19.	Anupam Anand	Evaluation Officer
World Bank ITS		
20.	Viven Prasad Sade	Senior IT Officer
21.	Priyanka Chandrahas Kamat	IT Officer
STAP Secretariat		
22.	Guadalupe Duron	Program Officer, GEF Scientific and Technical Advisory Panel
GEF Agencies		
23.	Olga Gordiiievskia	GEF Coordination Office, United Nations Industrial Development Organization (UNIDO)
24.	Stefanie Valcheva	GEF Coordination Team UNIDO
25.	Estefania Samper	RBM Specialist, UNDP
26.	Jaturong Padungsapya	Global Environmental Finance, United Nations Development Programme (UNDP)
27.	Luzmila Lambrano	Program Associate UNDP
28.	Orissa Samaro	Senior Director, Conservation International
29.	Rosario Narciso	GEF Portfolio Management Officer (Consultant), Asian Development Bank
30.	Arun Abraham	Consultant- Asian Development Bank
Green Climate Fund – peer portal		
31.	Lilian Macharia	Head of Office of Portfolio Management
32.	Johann Elysee	Senior Portfolio Specialist
33.	Nebi Bekiri	Head of Enterprise ICT Programme Management Office

Annex B.2 Portal evaluation aspects

		Operational activities		
		Design	Management	Utilization
Aspects	Portal Navigation / Organization	<ul style="list-style-type: none"> ▪ Arrangement / sequence of pages 	<ul style="list-style-type: none"> ▪ Inclusion of all necessary pages. ▪ Functionality of links 	<ul style="list-style-type: none"> ▪ Ease of navigation. ▪ Ability to understand page layout and ease of search (by main portal categories)
	Portal Information	<ul style="list-style-type: none"> ▪ Effective visualization (maps, lists, summaries) 	<ul style="list-style-type: none"> ▪ Content analysis. ▪ Completeness of information. 	<ul style="list-style-type: none"> ▪ Ability to find information needed. ▪ Ability to understand information provided, including balance between detail and visualization
	Tracking	<ul style="list-style-type: none"> ▪ Efficient internal display of project timeline and results; deadline alerts 	<ul style="list-style-type: none"> ▪ Efficient internal display of project timeline and results; deadline alerts 	<ul style="list-style-type: none"> ▪ Efficient external display of project timeline and results; deadline alerts
	Document Systems		<ul style="list-style-type: none"> ▪ Ease of accessing and processing submissions. ▪ Efficient workflow. 	<ul style="list-style-type: none"> ▪ Ease of upload. ▪ Efficient workflow.
	Interactivity		<ul style="list-style-type: none"> ▪ Contact provided. ▪ Correct channeling of queries. 	<ul style="list-style-type: none"> ▪ Ability to find contact. ▪ Question answered quickly / accurately
	Alignment with overall mission	<ul style="list-style-type: none"> ▪ Efficient use of internal users' time 	<ul style="list-style-type: none"> ▪ Accomplishment of portal mission to GEF 	<ul style="list-style-type: none"> ▪ User satisfaction ▪ Main positive / negative features. ▪ Main missing features. ▪ Efficient use of external users' time. ▪ Comparison with similar portals.
	User percentage		<ul style="list-style-type: none"> ▪ User growth rate 	<ul style="list-style-type: none"> ▪ Likelihood to influence decision to engage with GEF in future.

Annex B.3: Comparison of the GEF Portal with peer portals – based on observation as an external user

Aspect	Best practices	GEF Portal	UNFCCC CDM	IRENA Project Navigator	Green Climate Fund Project Portfolio
Easy navigation	Clear page layout through use of proper headings and topic clusters	Menu icons are used, which is a good resource. Some of the ones used at the GEF Portal match their function with the icon. Others mislead the user into expecting other type of information.	Topic cluster used, a good resource for this portal.	The portal has a clearly understandable page layout and topic classification. Within each page there are well-defined topic clusters.	The website is divided into two well differentiated pages. Within each page there are well-defined topic clusters.
	Clear location on the sitemap through use of breadcrumbs	No.	Yes, limited to two layers.	Yes.	No. Nor are 'go back' buttons offered.
	Ease of understanding the components/features of the site	Menu icons are used, which is a good resource. Buttons with different functions share color, screen location, shape, and size. A better way to visually guide the reader of the intent should be used.	There is a lot of information, what makes using the sitemap as definitely needed in order to understand the portal. The sitemap presented is clear though.	Images and icons help guiding the user and giving hints of what the feature is meant to do.	Easy to understand where each button is leading to, but no icons used. Would likely do better with use of icons as a better way to guide visitors, lessening text.

	<p>Ease of finding the information needed</p>	<p>The filter menu may benefit from some hierarchical organization.</p> <p>Having a way to reverse filter selections is good – currently it is the “x” button on filters applied. There may be other ways to help users navigate filtering as well.</p> <p>Icons leading to pop-ups are not intuitive.</p>	<p>The project search offers a simple and useful filter tool.</p> <p>The table presenting results from the search lacks dynamism, but it is still useful.</p> <p>There is a quick search option for the whole portal.</p>	<p>The clean page layout is very intuitive, what make finding a section easy.</p> <p>Although through the menus and section it is easy to find information, the search function offered does not work.</p>	<p>The filter menu is well distributed.</p> <p>Given the nature for the information to filter, the menu offers a useful way to filter through different aspects by enabling multiple selection via check boxes. This also enables a more versatile way to deselect filtering criteria.</p> <p>The filter menu may benefit from a submit or go button, because every time a checkbox is selected or deselected, both the filter menu and the results dashboard reloads, what could be undesirable.</p>
	<p>Avoid analysis paralysis by reducing the number of choices</p>	<p>Number of menus offered is minimal.</p>	<p>There is too much information available (mainly as text) and is not adequately prioritized.</p>	<p>On the main pages the number of choices is limited. Once inside subpages, the number increases; however, it is still comfortable for the user, since the content scope is already framed within a topic.</p>	<p>The website is divided into two well differentiated pages.</p> <p>The gallery presentation of projects can be visually overwhelming at first, because it calls the attention first rather than the filter menu.</p>

Visual appeal	Consistent branding	Yes.	Yes.	Yes.	Yes.
	Appealing color scheme/Useful color contrast	<p>Clean title bar.</p> <p>The green color and blue are highly saturated, and it may be that they compete here for our attention.</p> <p>The logo and green text disappear from the rest of the navigation experience. There may be ways to leverage this color and design to guide users back to the top level.</p> <p>Buttons with different functions share color. This should be changed to guide better the user.</p>	<p>It has a neutral esthetic and lacks dynamism, which makes it difficult to find the topics of interest.</p> <p>Color contrast is not useful. It should be optimized.</p>	<p>Simple and clean design, which take advantages of the brand color scheme, contrast between fonts, negative spaces, and homogeneity between the elements of the same page.</p>	<p>Clean design, which stick to the brand color scheme and homogeneity between the elements of the same page.</p> <p>Color contrast is well used for parts with white background. Nevertheless, contrast between background and text for parts with images as background should be optimized.</p>
	Content easy to read by being concise, minimizing text and/or using short paragraphs	Minimal text.	Text are divided in short paragraphs. However, there is too much information available.	Yes.	Yes.
Content easy to read through proper text layout	Pop-ups with some distracting text-wrap.	Some links have disproportionated size, being relatively bigger without need.	Yes. Text sizes help to guide and focus the user.	Yes. However, in some parts the interactivity of the text popping up or simulating a counter can be undesirable.	

	Content easy to read through block reading	Not applicable since the amount of text presented is minimal.	Yes.	Yes. This portal makes use of text boxes combined within interactive diagrams enabling block reading.	Block reading enabled. However, at some points the blocks 'boundaries' are not clearly distinguishable.
	Content easy to read through use of bullet points	Yes. Use of alternating gray bars used for project lists. It would be interesting to check if this style can be pushed in other areas where things are listed.	Yes. Everywhere where a list is need bullet points with the character '+' are used.	Yes.	Rarely used, could be optimized.
	Use of negative spaces to avoid crowded pages	Clean design with minimal text and negative spaces well distributed.	Used, but not optimized.	Clean design with negatives spaces well distributed.	Used, but could be optimized.
	Use of quality images/photos and balance between words and images	No. Only icons used. More images could make the site more visually appealing.	No. There is a lack of balance between images and text. There is predominantly text.	Yes. The portal uses qualitatively good diagrams and images as a way to assist in the understanding of the ideas and information presented.	Yes. The website uses appropriately diagrams and high-quality images, contribute to a better user experience.
	Use of human face aiming at "putting a face to the name"	No.	Yes, but the images are small.	No.	Yes.
Accessibility	Accessibility for the public	Account needed: Only persons/organizations legitimate through the World Bank Group can access.	Account needed: All public can create one.	Account needed: All public can create one.	For all public. This is not a portal to log in, but only a website.

		The pertinent result from the web search leads to a blank page. When correcting the URL deleting the “/App” at its end, it works. This should be fixed.			
	Functional links	The home button leads to a currently empty page. Too many pop-ups. Not only does that disable the back buttons, but it is likely to run into problems with secure browser setups that disable pop-ups.	Yes, but there are too many of them, which makes it difficult to read given the different fonts types in a small space.	Yes.	Yes.
	Optimization for mobile	<i>Unable to assess</i>	No. The mobile version is simply the same desktop version.	No. The homepage seems to be optimized for mobile. However, the portal itself no.	Yes. Visitors are able to easily navigate the website on mobile devices.
	Optimization for speed ¹⁰	Load time: 982 ms Overall performance grade A (99/100)	Load time: 6.45 s Overall performance grade C (79/100)	Load time: 930 ms Overall performance grade C (75/100)	Load time: 828 ms Overall performance grade C (80/100)

¹⁰ Website speed test performed using the [Pingdom Website Speed Test](#).

	Prioritization of SEO ¹¹ (lower the optimization potential, the better it is)	90% optimization potential. The GEF Portal URL (gefportal.worldbank.org) redirects to the Microsoft Online Login page for the World Bank. The first results from web search are documents or news about the portal rather than the portal itself.	29% optimization potential.	35% optimization potential.	30% optimization potential.
	Strong security ¹²	Medium The server does not support Forward Secrecy with the reference browsers.	Strong	Medium This server accepts RC4 cipher, but only with older protocols.	Strong
Interactivity	Strong CTAs (calls-to-action)	<i>Unable to assess</i>	There is a tab for contact with FAQs for different topics, where the questions are addressed through other websites or	This portal is not conceived as a platform for interactivity with the institution; therefore, it does not count with features enabling	Each project factsheet has a 'contacts' section, where contact details of officers from both the GCF and the corresponding accredited entity are shown.

¹¹ Website SEO check performed using the [SEO Checker from IONOS](#).

¹² Website security check performed using the [SSL Server Test from Qualys SSL Labs](#).

			<p>complex documents.</p> <p>The 'contact us'-section is well organized and the institutional contact emails are categorized regarding the topic or issue addressed.</p>	<p>communication channels with IRENA.</p>	
	Email marketing enabled	Unable to assess	No.	Yes.	Yes.
	Social media enabled	Unable to assess	Yes, both at the homepage before logging in as well as once logged in in the portal.	Yes, at the homepage, before logging in.	Yes.
	Chatbots used.	Unable to assess	No.	No.	No.

Note: SEO = search engine optimization.

Annex B.4: Summary of key results of the online survey

The percentage of respondents that assess performance on a given dimension to be in top half of a six-point scale. The specific response options for the scale differed based on the dimension, but broadly the top half indicates acceptable performance range whereas the bottom half indicates unacceptable performance range. There is some overlap in the respondent categories; therefore, the horizontal total may not add up. (Obs.=number of observations)

Performance Dimension	Agencies		Secretariat		Operational focal points		Others		All users	
	Obs. *	HS to MS	Obs.	HS to MS	Obs.	HS to MS	Obs.	HS to MS	Obs.	HS to MS
Ease of Navigation	95	28%	28	46%	10	90%	13	46%	145	37%
Ease of webpage use	95	28%	28	39%	9	100%	13	46%	144	36%
Clarity of webpages	95	62%	28	57%	10	100%	13	85%	145	66%
Completeness of webpages	94	66%	28	46%	9	78%	13	77%	143	64%
Display easy to understand	94	62%	27	37%	9	89%	12	83%	141	60%
Visualizations useful	86	72%	21	62%	9	100%	11	64%	126	71%
Project timeline display	88	55%	26	62%	10	90%	12	58%	135	59%
Project results display	81	41%	20	45%	10	100%	12	58%	122	48%
Important alerts	73	37%	23	48%	8	63%	7	43%	111	41%
Ease in information search	87	45%	28	32%	8	75%	12	50%	134	45%
Ease in conducting analysis	85	31%	28	32%	8	88%	12	42%	132	35%
Provided information useful	87	84%	28	86%	9	89%	12	100%	135	86%
Downloadable information useful	82	83%	26	73%	9	100%	11	91%	127	83%
Ease in accessing proposals	86	41%	26	65%	8	75%	10	80%	130	51%
Ease in uploading docs	83	41%	13	85%	3	100%	7	71%	106	50%
Efficiency in processing proposals	78	35%	20	55%	4	100%	5	80%	107	43%
Real-time data availability	62	79%	22	77%	7	100%	9	89%	99	81%
Efforts compared to benefits	85	52%	27	56%	7	86%	11	64%	129	55%
Weblinks function	67	85%	20	95%	6	100%	10	90%	102	88%

Performance Dimension	Agencies		Secretariat		Operational focal points		Others		All users	
	Obs. *	HS to MS	Obs.	HS to MS	Obs.	HS to MS	Obs	HS to MS	Obs	HS to MS
Access to necessary contacts	57	39%	15	33%	8	88%	7	71%	86	44%
Expectations with Portal met	78	42%	26	35%	6	83%	10	80%	119	45%
Overall satisfactory with Portal	87	49%	27	48%	8	88%	11	82%	132	54%
Comparison with other portals	74	42%	22	55%	7	86%	8	63%	110	48%
Effect on transparency	81	88%	24	79%	8	88%	11	82%	123	85%

Source: Online survey.