GEF Results-Based Management:

Agency Self-Evaluation and the GEF Portal

GEF Results-Based Management:

Agency Self-Evaluation and the GEF Portal

Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

20 ANNIVERSARY

Contents

Annexes

References 77

Boxes

Figures

Tables

Foreword

This report presents the Evaluation of the Agency Self-Evaluation Systems and the Evaluation of the Global Environment Facility (GEF) Portal. These evaluations were conducted as inputs to the Seventh Comprehensive Evaluation of the GEF and were undertaken to examine the key aspects the GEF's results-based management (RBM) system.

The Evaluation of Agency Self-Evaluation Systems addressed the key challenges in supporting learning, incentivizing candor, and harmonizing reporting through the self-evaluations prepared by the GEF Agencies. The Evaluation of the GEF Portal assessed the performance of the Portal in improving data quality and security, supporting monitoring and adaptive management, and meeting the needs of the various users across the GEF partnership. Both evaluations identified several strengths in the GEF RBM system and provide pointers on where it may improve.

This report was presented to the GEF Council in June 2021. Management agreed to the recommendations of the two evaluations and presented action plans to address them. The GEF Council took note of the management action plan and endorsed it.

Through this report, the GEF Independent Evaluation Office intends to share the lessons from these evaluations with a wider audience.

Juha I. Uitto
Director, GEF Independent Evaluation Office

Acknowledgments

Abbreviations

ADB	Asian Development Bank
AfDB	African Development Bank
CAF	Development Bank of Latin America (Corporación Andina de Fomento)
CDM	Clean Development Mechanism
DBSA	Development Bank of Southern Africa
EBRD	European Bank for Reconstruction and Development
FAO	Food and Agriculture Organization of the United Nations
FECO	Foreign Economic Cooperation Office, Ministry of Environmental Protection of China
FUNBIO	Brazilian Biodiversity Fund (Fundo Brasileiro para a Biodiversidade)
FY	fiscal year
GCF	Green Climate Fund
GEF	Global Environment Facility
IDB	Inter-American Development Bank
IEG	Independent Evaluation Group
IEO	Independent Evaluation Office
IFAD	International Fund for Agricultural Development
IRENA	International Renewable Energy Agency
ITS	Information and Technology Solutions

IUCN	International Union for Conservation of Nature
M&E	monitoring and evaluation
OPS	overall performance study (now termed comprehensive evaluation)
PIR	project implementation report
PMIS	Project Management Information System
RBM	results-based management
STAP	Scientific and Technical Advisory Panel
UN	United Nations
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
UNIDO	United Nations Industrial Development Organization
WWF-US	World Wildlife Fund–US

GEF replenishment periods

Pilot phase: 1991–94	**GEF-1:** 1995–98
GEF-2: 1999–2002	**GEF-3**: 2003–06
GEF-4: 2006–10	**GEF-5:** 2010–14
GEF-6: 2014–18	**GEF-7:** 2018–22
	GEF-8: 2022–26

Executive summary

The Independent Evaluation Office (IEO) of the Global Environment Facility (GEF) has undertaken several evaluations and reviews that address different aspects of the GEF results-based management (RBM) system as part of its work for the Seventh Comprehensive Evaluation of the GEF (OPS7). This report brings together evidence from two of these evaluations: (1) the Evaluation of GEF Agency Self-Evaluation Systems and (2) the Evaluation of the GEF Portal.

AGENCY SELF-EVALUATION SYSTEMS

This evaluation aimed at assessing the extent to which GEF Agency self-evaluation systems meet GEF requirements and provide credible, quality, and timely information to support accountability and learning. 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 conducted to analyze challenges related to Agency self-evaluation systems and to identify solutions through a design thinking–based participatory process.

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

- Terminal evaluations are managed by the project staff and validated by the Agency's evaluation unit (internal or independent unit).
- Terminal evaluations are managed by the project staff but not validated by the Agency's evaluation unit, or such a unit is absent.
- Terminal evaluations are managed by the Agency internal evaluation units.

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.

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) notes 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 Environment 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–US (WWF-US), have explicit guidance on midterm reviews. The Agencies that joined the partnership during the second round of expansion are interested in having more guidance on the conduct of midterm reviews.

Arrangements in place within Agencies to harmonize their self-evaluation systems with GEF requirements vary. Some provide a step-by-step guide, and some have mainstreamed these requirements within their standard guidelines. Development finance institutions such as the African Development Bank (AfDB), the Asian Development Bank (ADB), the European Bank for Reconstruction and Development (EBRD), the Inter-American Development Bank (IDB), the International Fund for Agricultural Development (IFAD), and the World Bank 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.

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 ADB, IDB, and the World Bank, these units validate terminal evaluations of only those projects that involve funding beyond a certain threshold. Evaluation units in FAO, UNEP, and UNIDO manage the evaluation process, including providing quality assurance. The GEF coordination units of Conservation International, 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 PIRs and midterm reviews. Only WWF-US and, upon exceptional request, FAO and IUCN validate midterm reviews.

The ratings 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, 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).

There are gaps in submission of PIRs and midterm review reports, particularly in low-performing projects. PIRs 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, PIRs were less likely to be available for projects that were rated unsatisfactory or highly unsatisfactory at implementation completion. Midterm reviews are available for 43 percent of completed full-size projects. Thus, the instrument is not being used as often as it should be.

GEF Agencies generally comply with the GEF IEO's terminal evaluation guidelines for full-size 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, 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.

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.

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 PIRs trigger corrective actions. The focus is on activities and outputs.

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 United Nations entities—some Agencies such as IFAD and UNDP have a strong emphasis on learning on doing the right things; others have not given as much attention to this. Nongovernmental organizations have almost no systemic arrangements for supporting learning on doing the right things.

GEF Agencies generally do not incentivize candor in self-evaluation, although some are making efforts to encourage such an evaluation culture. The key measure of success in Agencies is total resources and number of projects managed. This organizational logic informs project design and management, with the result that evaluation comes to be seen as a necessary requirement rather than an opportunity to learn. Project staff have incentives to move their projects along without issue and in their receiving 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.

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 supports projects from design to post-evaluation.

GEF PORTAL

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. The evaluation assessed the extent to which the GEF Portal is meeting its objectives and is an improvement over the PMIS. It provides lessons from the experience with the Portal's development and rollout.

Information was gathered from multiple sources. These included a review of relevant GEF Council documents, publications by the Secretariat, and evaluations by the GEF IEO; interviews; 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 Portal users on new features. Along with the GEF Portal, the web pages of three peer portals— the Green Climate Fund's 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 from the viewpoint of an external user for comparison.

The Portal has enhanced 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 proposals via the Portal. The Portal creates a clear audit trail of the who, what, and when of a given action to facilitate accountability.

Overall, the Portal has contributed to improvement in data quality, though some errors in data outputs were noted. Improvements were especially noted for the more recent projects; this has been accomplished through increased automation and arrangements to ensure data entry discipline. The Portal has streamlined the process of submitting project proposals, PIRs, 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 in order 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.

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 have a clean layout. The icons used in the web pages are simple, elegant, and consistent. The GEF Portal compares well with the three selected peer portals on criteria such as navigation, visual appeal, and accessibility. User perception on ease of navigation and use of web pages is varied and appears to be linked with frequency of usage—those who use the Portal more often have a more favorable perception of it than those who do not—and user group type. For example, a lower percentage of Agency staff than of other Portal users assess performance of the Portal to be in the satisfactory range on ease in navigation and use of web pages.

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. Several areas need to be developed further. These include development of a comprehensive system of alerts through email, 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 the limited resources budgeted for its development. Of the $922,229 cumulatively allocated for Portal development, $658,650 was available at the start of fiscal year 2017; at the end of fiscal year 2020, $299,000 remained unspent. These figures are a low-end estimate because they do not reflect full costs.

Users perceive a need for a transparent process to collect information on problems, prioritize problems, and report on 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 beginning in 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. 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 team's responsiveness; however, some users are unaware of the process for contacting the 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.

The Portal has contributed to the acquisition of more and better data, and access to data is more regulated than before. More data are being generated through the online processing of project proposals, reviews, uploading of documents, and entry of data on key milestones. However,

the 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.

Connectivity is a major concern for many users. Recurrent issues were cited related to logging in, connection losses, and "silent logouts" caused by the page timing out, leading to wasted effort. Earlier, these problems were common for all users, but this was partially addressed when the Portal team enhanced the available bandwidth. For many users based in the least developed countries and remote areas, it continues to be difficult to use the Portal with ease because Internet connectivity is a constraint.

CONCLUSIONS

GEF Agency self-evaluation systems

How do policy frameworks in the GEF Agencies support their self-evaluation systems? The evaluation found that the policy frameworks of GEF 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.

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. The GEF IEO Annual Performance Report 2020 (GEF IEO 2022b) showed that PIRs generally give overly optimistic performance ratings. This

evaluation found that even though the ratings provided in the PIRs 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. Midterm reviews are generally regarded as credible and useful, but they are not prepared for a majority of full-size 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 the right things.

To what extent do the Agency self-evaluation systems meet GEF requirements according to the relevant GEF policies and guidelines? Agency 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 into 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 making comparisons across the Agencies. The midterm reviews are mandatory for full-size projects but are not prepared for a majority of the others. The submission rate of PIRs is somewhat lower for low-performing projects. The quality of terminal evaluation reports is improving, although reporting on whether feedback from operational focal points has been sought and addressed, and on the application of safeguards, is generally weak.

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.

GEF Portal

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, 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 peer portals on some of the technical parameters on which 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.

Several interview respondents and online survey participants expressed frustration with 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 of the process 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 users so that they are better aware of the process and have a clearer understanding of the road map to further development of the Portal.

The advent of the Portal shifted the burden of data entry 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 with which they have to work. At the same time, there are substantial resources for the special initiative for the Portal that have not been fully utilized. GEF management needs to assess whether additional support is necessary and how speedier development of the Portal can be aided.

RECOMMENDATIONS

GEF Agency self-evaluation systems

- The GEF Secretariat and Agencies should strengthen the use of midterm reviews for learning and adaptive management.

- 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.

GEF Portal

- The GEF Secretariat should strengthen its process to address user feedback on the Portal.

- GEF Management should develop and implement a time-bound plan to speed up the development of the Portal.

Introduction

The Independent Evaluation Office (IEO) of the Global Environment Facility (GEF) undertook several evaluations and reviews addressing different aspects of the GEF results-based management (RBM) system as part of its work for the Seventh Comprehensive Evaluation of the GEF (OPS7).[1] This report brings together evidence from two of these evaluations: (1) the Evaluation of GEF Agency Self-Evaluation Systems and (2) the Evaluation of the GEF Portal.

A related effort, also conducted for OPS7, was a review of the GEF terminal evaluation validation process (GEF IEO 2020b); this found that the GEF IEO's validation process was well established and the data set from the process allows for comparisons across the GEF partnership. It found several areas where the GEF IEO's approach to the validation process might be strengthened: greater sharing of information on the validation process with the Agencies, graduation of Agencies that have a robust validation process in place, and giving more attention to the newer GEF Agencies. It also called for greater attention to strengthening a community of practice on the validation of terminal evaluations.

The two evaluations presented in this report expand on this work and bring other areas relevant to RBM into focus. The Evaluation of GEF Agency Self-Evaluation Systems, presented in chapter 2, assesses the extent to which these systems meet GEF requirements and provide credible, quality, and timely information to support accountability and learning. Using a systems and design thinking process, it identifies bottlenecks and solutions.

The Evaluation of the GEF Portal, presented in chapter 3, assesses the extent to which the GEF 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.

Chapter 4 presents the respective conclusions and recommendations of the two evaluations. A set of annexes supports and supplements the analyses outlined in the main text.

---

[1] The 2021 Annual Performance Report presents an overall analysis of RBM in the GEF, focusing on arrangements managed by the GEF Secretariat (GEF IEO 2023).

GEF Agency self-evaluation systems

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 2.1](#)). These 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.

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 learn from experience. At the corporate level, they facilitate portfolio portfolio and may provide early warnings on trends that warrant attention and contribute to learning through sharing of lessons. Although the GEF IEO covered aspects of Agency self-evaluation systems in 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 was aimed at addressing this information gap.

2.1   Background

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 UN system and across the multilateral development banks consider self-evaluation to be a project/program management responsibility and primarily intended for management's own use (ECG 2018; UN 2018; UNEG 2017). A self-evaluation system takes into account scope, responsibilities, methods, data, and products; and applies to the different stages of the project/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.

The expectations from Agency self-evaluation systems are outlined in several GEF policy

Figure 2.1 Timeline of GEF Agencies joining the GEF partnership

1992	**1st round of expansion 1999–2006**	**2nd round of expansion 2013–15** **Today**
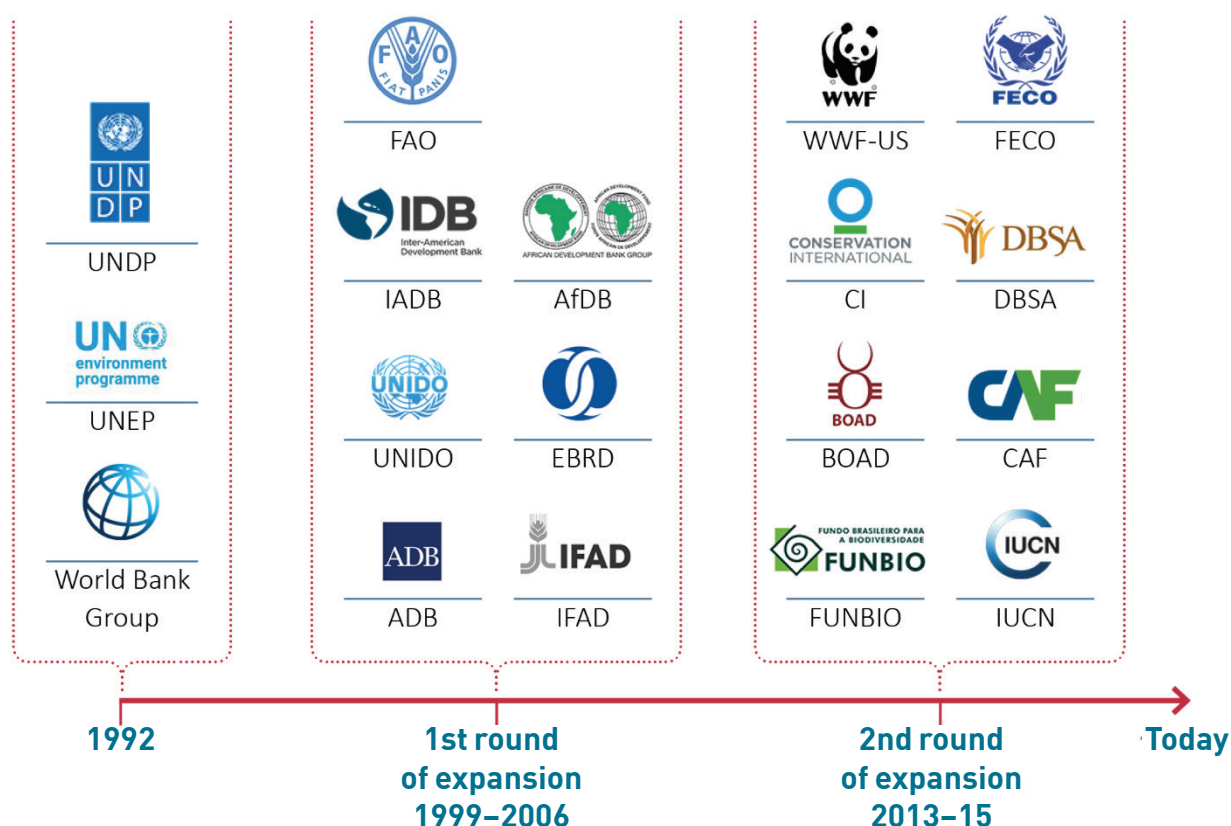
documents and Agency policies. These include the GEF Evaluation Policy (GEF IEO 2019), the GEF Monitoring Policy (GEF 2019b), the Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects (GEF IEO 2017a), and the Minimum Fiduciary Standards for GEF Partner Agencies (GEF 2019a), all of which specify several requirements.

- The **GEF Evaluation Policy** requires the Agencies to prepare terminal evaluations and midterm reviews (where applicable), and to monitor their respective GEF portfolios.

- The **terminal evaluation guidelines for full-size 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 full-size projects (those involving more than $2.0 million in GEF funding), the Agencies are encouraged to use them to develop terms of reference for evaluation of GEF-funded medium-size projects and for programs.

- The **GEF Monitoring Policy** addresses the guiding principles for monitoring along with other requirements, including reporting through project implementation reports (PIRs), midterm reviews, and tracking tools. It requires that the Agencies submit to the GEF 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 PIRs no later than 75 calendar days after the end of the first full fiscal year

(July 1–June 30). At project midterm, the Agencies are required to submit midterm reviews for all full-size projects, and for medium-size projects where available; and tracking tools and information on achieved results across applicable core indicators and subindicators.

The GEF IEO undertook this evaluation to assess the extent to which Agency self-evaluation systems provide quality and timely information, comply with 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 Endeva, a consulting firm with expertise in design thinking.

2.2 Literature and GEF Agency reviews

The literature on self-evaluation systems in international development organizations is not well developed. However, self-evaluation systems are received increasing attention from the evaluation units of 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 (AfDB IDEV 2018). The Independent Evaluation Department 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) also recently undertook an assessment of its self-evaluation system (Kirk 2019); the United Nations Development Programme (UNDP) has a similar initiative under way.[1]

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

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 to convert tacit knowledge into explicit knowledge (Nonaka 1994; Spender 1996). Taut's "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.

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

---

[1] Source: Communication with UNDP Independent Evaluation Office.

commonly used approach to address the credibility challenge.

In the evaluations conducted by the evaluation units of the GEF Agencies, there is a recognition that self-evaluation systems are not being used for learning to the extent they should be. The evaluation of the World Bank's self-evaluation systems concluded that these systems complement independent evaluation systems, but the focus of the former is more on results reporting and meeting accountability needs and less on learning to enhance performance (IEG 2016). The review of ADB's project-level self-evaluation system concluded that the system is robust and credible, and is useful for both learning and accountability (ADB IED 2020). However, it also noted that the 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-evaluation and argue that, without appropriate incentives, it is likely that self-evaluation would become a bureaucratic requirement and less useful for learning (ADB IED 2020; IEG 2016).

2.3   Key questions

This evaluation sought to answer the following questions:

- **How do policy frameworks in the GEF Agencies support their self-evaluation systems?** The evaluation assessed 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-evaluation, provide guidance on how self-evaluations ought to be

conducted, and clarify relationships with independent evaluation.

- **To what extent do the Agency self-evaluation systems provide credible, quality, and timely information to support accountability and learning?** The evaluation considered the arrangements within the Agencies to conduct self-evaluations (or equivalent[2]). It considered arrangements to address quality assurance, harmonization, information management, and knowledge sharing, as they relate to self-evaluation of GEF projects. The evaluation assessed the extent to which these arrangements allow for credible self-evaluation of GEF-supported activities. It also assessed opportunities and solutions for improving self-evaluation systems.

- **To what extent do the Agency self-evaluation systems meet GEF requirements according to the relevant GEF policies and guidelines?** The evaluation assessed the quality of self-evaluations and the extent to which the GEF Agencies are meeting relevant GEF requirements. These include the minimum requirements related to project M&E stipulated by the GEF Evaluation Policy and its predecessor the 2010 GEF M&E Policy (GEF IEO 2010a), the Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects, and the GEF Monitoring Policy.

- **What are the factors that influence the effectiveness of the self-evaluation systems that could have an impact on the quality and timeliness of information provided to the GEF?** The evaluation assessed whether and how different variables—such as policy frameworks, information management arrangements, incentives

---

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

to promote candid reporting, quality assurance arrangements, level of resources provided, and capacities of the local partners—affect the evaluation of self-evaluation systems.

2.4   Methodology

EVALUATION DESIGN

The evaluation incorporated 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 to which all the GEF Agencies were invited to participate.

The evaluation used a multiple-case design (Yin 2017) and covered all the GEF Agencies. The self-evaluation system of a GEF Agency—as it relates to GEF-supported activities—is the primary unit of analysis. Within each Agency, information on the self-evaluation system was gathered through review of relevant Agency policy documents, and interviews with the unit responsible for RBM and with senior management. The products of the self-evaluation system—PIRs, midterm reviews, and terminal evaluations—were reviewed to determine the areas where these reports are strong and where they might be improved. In these analyses, each product is treated as a unit of analysis, and, where possible, performance on a specific dimension is aggregated and reported by Agency. The coverage for products is through a representative sample. Terminal evaluation templates used by the GEF Agencies were reviewed to assess reporting practices.

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 the effectiveness of self-evaluation systems at different levels.

Design thinking was a key element of the evaluation approach. Design thinking considers the user's 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

- **Literature review.** The evaluation drew on the literature relevant to self-evaluation systems. Information gathered through the literature review was used in the development of the evaluation design and approach. Reports prepared by the United Nations Evaluation Group, the Evaluation Cooperation Group, the Multilateral Organisation Performance Assessment Network, and the Joint Inspection Unit of the UN System 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 for this evaluation (GEF IEO 2020a).

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

- **Workshops.** Five workshops were held. 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 Agency level and within the GEF partnership.

- **Desk reviews.** Several desk reviews were conducted as part of the evaluation. These comprised a review to assess Agency policies and guidelines relevant to self-evaluation, a review to assess compliance with GEF IEO guidance on the preparation of terminal evaluations, a review of PIRs, and an assessment of the availability of midterm reviews.

  - 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 for 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 which had submitted terminal evaluations to the GEF IEO in the preceding four years were analyzed.

  - The **review of terminal evaluations** to assess compliance with GEF IEO guidelines covered 270 completed projects. The GEF IEO issued guidelines for the GEF Agencies in conducting terminal evaluations for full-size projects in April 2017. These guidelines specify reporting requirements on topics such as project identification, basic project details, theory of change, feedback from stakeholders, methodology, outcomes, 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

medium-size projects and to programs, the Agencies are encouraged to use the guidelines in preparing terminal evaluations for these as well. Terminal evaluations for both full- and medium-size projects were covered in the review.

- The **review of PIRs** assessed submission gaps and consistency and timeliness in reporting. A stratified random sample of 150 completed projects approved from GEF-3 onward was drawn. This comprised three samples of 50 projects that represented 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.

- The 2019 GEF Monitoring Policy, as well as its predecessor the 2010 GEF Monitoring and Evaluation Policy, require that the Agencies prepare and submit a **midterm review** for full-size projects at the midterm of the project implementation period, and submit a midterm review for medium-size projects where such review has been conducted. The evaluation assessed gaps in submission of these reviews for the completed full-size projects and in the 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 the evaluation to assess PIR-related submission gaps.

- **Data sets.** The terminal evaluation review data set maintained by the GEF IEO was used. This data set contains validated ratings on performance dimensions such as project outcomes, sustainability, M&E, implementation, execution, and quality of terminal evaluation. The data set available as of February 2021 was used

for analysis on trends in quality of terminal evaluations.

CONDUCT OF THE EVALUATION

The evaluation was led by the GEF IEO, and the IEO designed its overall approach. The Office conducted the review of Agency compliance with GEF IEO terminal evaluation guidelines, trends in quality of terminal evaluations, and review of reporting through PIRs. Endeva conducted the systems and design thinking workshops, interviews to assess Agency self-evaluation systems, and review of the terminal evaluation templates.

The approach to the evaluation was developed during fiscal year (FY) 2020. 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, the GEF Secretariat, and the GEF Agencies was incorporated to finalize the approach paper.

The evaluation was conducted between September and 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 the GEF Agencies; their feedback has been addressed in finalizing this report.

2.5 Findings

STRUCTURE OF SELF-EVALUATION SYSTEMS

The structure of the self-evaluation systems across the 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. Some of these Agencies—including ADB, AfDB, EBRD, the Inter-American Development Bank (IDB), the International Fund for Agricultural Development (IFAD), UNDP, and the World Bank—have a structurally independent evaluation unit, whose head reports directly to its governing board and is independent of management. These independent evaluation units validate at least some of the terminal evaluations prepared by management. For example, evaluation units of ADB, IDB, and the World Bank validate terminal evaluations of those projects that involve funding beyond a certain threshold. In IFAD, the evaluation unit provides feedback on the draft of the terminal evaluation reports for a few projects; another IFAD unit, which is not independent of its management, conducts validations. UNDP's evaluation unit conducts validation for terminal evaluations of all GEF projects.

Another set of Agencies, where management conducts terminal evaluations, do not have an evaluation unit. These Agencies include Conservation International, the International Union for Conservation of Nature (IUCN), and the World Wildlife Fund–US (WWF-US), all of which became part of the GEF partnership during the second round of expansion. These three Agencies 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.

The evaluation units of still another set of Agencies are responsible for conducting terminal evaluations. These Agencies are the Development Bank of Latin America (CAF), the Development Bank of Southern Africa (DBSA), the Food and Agriculture Organization of the United Nations (FAO), the Foreign Economic Cooperation Office in the Ministry of Environmental Protection of China

(FECO), the United Nations Environment Programme (UNEP), and the United Nations Industrial Development Organization (UNIDO). 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 UNEP's executive director; and those of UNIDO and FAO report to their respective director generals. Peer reviews of the evaluation function of UNIDO, UNEP, and FAO have assessed their evaluation units to be functionally independent (MoFA Denmark 2010; UNEG 2012a, 2012b). The CAF, DBSA, and FECO evaluation units have not undergone a peer review process.

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 the conduct of terminal evaluations, they are able to dedicate more resources and time to learning-focused activities at the corporate level.

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 PIRs and midterm 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 the World Bank provides technical support to projects that incorporate experimental approaches to assess the impact of development interventions and provide lessons on what works and in what contexts. Similarly, the World Bank Quality Assurance Group, which operated until FY 2010, provided early feedback on the quality of project design and on the quality of supervision

to management. See annex A for Agency good practices in self-evaluation beyond those highlighted in this chapter and annex B for profiles of the self-evaluation systems of the GEF Agencies. Annex C provides more information on relevant self-evaluation characteristics of the GEF Agencies.

FRAMEWORK FOR SELF-EVALUATION IN AGENCIES

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 2.2). The purpose of *learning* is to enhance decision making and project performance within and across projects. But how the twin objectives of accountability and learning are to be met by the self-evaluation system is usually not clearly defined. A survey of the policy documents of the GEF Agencies revealed the following:

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

Figure 2.2 **Purposes and objectives of self-evaluation**

Source: Agency policy documents.

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

All Agencies provide guidance on evaluation criteria and processes as well as on quality assurance (figure 2.3). The majority incorporate at least some elements of the GEF requirements and IEO guidance for terminal evaluations in their respective guidance. Seven Agencies (Conservation International, FAO, FECO, UNDP, UNEP, UNIDO, WWF-US) have explicit guidance on midterm reviews. This difference is because midterm reviews are mandatory only for GEF-supported programs and full-size 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—expressed interest in receiving more guidance from the GEF Secretariat on the conduct of midterm reviews.

Guidance on how to harmonize self-evaluation systems with GEF requirements varies among Agencies. ADB, FAO, FECO, IDB, and UNIDO provide step-by-step guides. The GEF Agency NGOs

and UN entities (except for IFAD) have mainstreamed these requirements within their standard guidelines. The GEF Agency development finance institutions such as ADB, AfDB, EBRD, IDB, IFAD, and the World Bank 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. New Agencies, including Conservation International, FECO, and WWF-US, have established self-evaluation procedures for GEF projects to meet requirements. CAF and DBSA are currently reviewing their policies and guidance for GEF projects and programs. EBRD is currently redesigning its entire self-evaluation system.

Most Agencies provide generic guidance on tools and data capture for evaluation. FAO, IFAD, UNEP, UNIDO, and the World Bank provide guidance on the use of a theory of change in self-evaluation. UNEP has mainstreamed GEF requirements in its terms of reference for terminal

Figure 2.3 Coverage of issues in GEF Agency guidance documents

Issue	Agencies
Guidance on GEF core indicators	4 Agencies
Use of theory of change	5 Agencies
Midterm review preparation	7 Agencies
GEF requirements	11 Agencies
Role of monitoring	13 Agencies
Evaluation criteria and processes	15 Agencies
Questions and answers, including role of evaluation unit	15 Agencies

evaluations and guidance notes on assessment of stakeholder consultations and tool kit on assessment of safeguards.

The GEF core indicators are specified in the guidance provided by CAF, FAO, FECO, and UNDP. 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.

Some Agencies are not clear about GEF requirements and about good practices in implementing them. Although the GEF IEO has issued guidelines for conducting terminal evaluations for full-size projects, these have not been issued for medium-size projects and for programs. Lack of GEF guidance on the conduct of midterm reviews also contributes to a perceived lack of clarity on GEF requirements. The staff of Agencies that joined the GEF partnership during the second round of expansion and have yet to submit a terminal evaluation—that is, the West African Development Bank (BOAD), CAF, DBSA, FECO, and the Brazilian Biodiversity Fund (Fundo Brasileiro para a Biodiversidade, FUNBIO)—are less conversant with the GEF requirements related to terminal evaluations and need more support (GEF IEO 2022b).

Agencies expressed an appetite for learning more on the use of a theory of change in the 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 GEF Scientific and Technical Advisory Panel (STAP) has developed a primer on the use of a theory of change to support the design of GEF interventions (GEF STAP 2020). The primer covers different stages of the 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 theory of change guidance that may be applied in evaluations. It is still too early to empirically assess the extent to which the primer prepared by the 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 levels of rigor because Agencies differ in their interpretation and implementation approaches.

QUALITY ASSURANCE

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 project teams, so that they follow GEF requirements. GEF coordination units also review reports before submitting them to the GEF through the GEF Portal.

Depending on the Agency, the evaluation unit or 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 ADB GEF projects. The evaluation units of IDB and the World Bank only validate terminal evaluations for projects above a certain funding threshold. In IFAD, the evaluation unit provides feedback on the drafts of a few terminal evaluations; validations are conducted by another unit that is not independent of management. UNDP's evaluation unit validates terminal evaluations for all GEF projects.

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

The respective GEF coordination units of Conservation International, IUCN, and WWF-US manage the conduct of terminal evaluations. The coordination units provide feedback and are responsible for quality assurance. To date, these Agencies have submitted five terminal evaluations in total.

Many Agencies have internal or external peer reviews of terminal evaluation reports. AfDB and the World Bank undertake internal reviews and clearance processes for terminal evaluations that involve several review loops among different units and the evaluation unit (box 2.1). EBRD's independent evaluation unit pre-reviews terminal evaluations before their final submission and provides feedback to the team to ensure quality (box 2.2). The independent evaluation units at ADB and IDB prepare a draft validation report that is reviewed internally by their independent evaluation unit staff and shared for comment with the project teams. Similarly, at the 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 comment to the regional practice group director and to the recipient country. Some GEF Agencies (FAO, FECO, UNDP, UNEP, UNIDO) 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 (box 2.3). UNDP requires the commissioning unit to present preliminary findings to the GEF operational focal point in the recipient country.

Some evaluation units conduct postcompletion project evaluations and/or impact evaluations for a set of projects. The ADB and World Bank evaluation units undertake field-based verifications for a sample of projects postproject closure (box 2.4). Four development finance institutions (ADB, AfDB, IDB, World Bank) and four UN entities (FAO, IFAD, UNEP, UNIDO) conduct impact and thematic

Box 2.1 African Development Bank'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.

1. A first draft of the report is submitted to the sector manager for clearance.
2. The draft is sent to peer reviewers who should either be based in a field office or have previous field experience.
3. 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 IDEV for validation and quality assessment. The IDEV prepares its evaluation note based on all available documents and discussions from the review process, as noted above. Additionally, the evaluation note itself is subject to peer review to ensure consistency in applying all guidelines and is distributed to the operational department for response; the department then can either agree or disagree with the findings. If it disagrees, the department has to provide additional evidence to support its response.

This review process allows for extensive dialogue throughout the preparation and validation of the project completion report. While disagreements do occur, especially regarding ratings, the overall process is seen as highly useful in that it helps AfDB end up with an independent, consensual, and reasonable assessment of a project's performance and reduces any bias stemming from self-assessment of projects.

Sources: Interviews with Guirane NDiaye, Ayanleh Daher Aden, Hedi Manai, and Lafeta Camera, AfDB; AfDB 2018.

Box 2.2 EBRD's independent evaluation unit pre-reviews terminal evaluations before submission

To improve the quality of terminal evaluations, EBRD's independent Evaluation Department (EvD) has introduced a feedback loop before validation. Project teams hand in a first draft. The EvD reviews the draft by checking for quality and completeness—for example, whether outputs, outcomes, impacts, and lessons are reported on appropriately. Based on the feedback from the EvD, the operational teams make adjustments before final submission. An evaluator has five days to review the draft, and the operational teams have another three to five days to incorporate the feedback before submission. After the final submission, the report undergoes validation by the 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 the quality of the terminal evaluation ratings.

Sources: Interview with Barry Kodolkin, EBRD; Kirk 2019.

evaluations that may cover a sample of projects at implementation completion and/or postcompletion.

PIRs and midterm reviews are generally not reviewed by evaluation units. These reports are

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

The GEF Agencies are expected to complete terminal evaluations within six months of implementation completion and to submit the terminal evaluations to the GEF IEO within a year of

implementation completion. In addition to their usual program management systems, the GEF coordination units in some Agencies also trigger timely preparation of terminal evaluations. Most Agencies do not feel constrained by the time available for preparation and submission of terminal evaluations. One exception is ADB, which has an internal timeline for terminal evaluations that exceeds one year, and hence finds it difficult to meet the GEF deadline.

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 for UNDP and the World Bank, where the GEF IEO accepts terminal evaluation validations prepared by their respective independent evaluation units. The GEF IEO reports on submission gaps only when a submission is delayed for several years.

APPROACHES TO RATING PROJECT PERFORMANCE

The 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 Agency 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).

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 GEF Agencies address relevance to the local and national contexts, with some differences:

- **Development finance institutions do not assess relevance of GEF projects using a GEF perspective.** For example, EBRD only briefly addresses a project's relevance to the local context but does not rate the relevance. ADB and EBRD do not address the project's relevance to the GEF.

- **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.

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. There are some differences in the subcomponents assessed as part of effectiveness between Agencies:

- UNEP explicitly uses a theory of change as part of the assessment of effectiveness. Other Agencies use a theory of change implicitly and apply it inconsistently (GEF IEO 2022b).

- 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 indicators of project impact and not project effectiveness.

- FAO and the World Bank include the achievement of global environmental objectives as part of their effectiveness assessment.

- UNIDO includes project management as a component of effectiveness.

Most Agencies rate the overall effectiveness of a project; EBRD rates the achievement of each project component separately but does not indicate an overall effectiveness rating.

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 across the Agencies regarding efficiency assessment:

- UNEP includes a section on the use of preexisting resources.
- FAO includes a section on the M&E of the project.
- IFAD includes project management.
- Most Agencies rate the overall efficiency of the project; WWF-US does not do so.

Sustainability

The GEF defines sustainability as 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, 13)

Agencies broadly follow this guidance, with some differences:

- UNEP does not consider environmental risks; and Conservation International, EBRD, FAO, and IUCN do not explicitly address the four aspects.
- The World Bank does not rate or assess sustainability because, in its judgment, there is not enough evidence at project closure to report meaningfully on this criterion. Instead, it assesses risk to development outcomes, which is the approach used by the GEF IEO to assess likelihood of sustainability.

Overall outcome rating

There are minor differences in how the GEF Agencies assign an overall outcome rating. According to GEF requirements, the overall outcome rating should include the three criteria of effectiveness, efficiency, and relevance. Sustainability should not be part of the overall outcome rating (GEF IEO 2017a). 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. If ratings on these two criteria are not 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 those followed by the GEF Agencies:

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

Most Agencies use the rating scales following the GEF requirements. The 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 scale ranging from likely to unlikely (GEF IEO 2017a). Although most Agencies are aligned with the GEF IEO in using a four-point scale ranging from likely to unlikely, some exceptions are notable:

- EBRD and UNEP use 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

PIRs document challenges faced during project implementation and identify corrective actions; however, there are gaps in the submission of

these reports, particularly in low-performing projects. PIRs 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 PIRs generally document challenges during implementation in a timely manner, but in some instances may not adequately reflect the level of urgency required to meet the challenge. While overall this suggests that the PIRs are providing information that may be useful for corrective action, PIRs are less likely to be available for projects that were rated unsatisfactory or highly unsatisfactory at implementation completion than for projects that were rated satisfactory or highly satisfactory (table 2.1).

Midterm reviews are being prepared for only a limited number of full-size projects. The percentage is much lower for medium-size projects (table 2.1), because Agencies are encouraged, but not required, to prepare midterm reviews for these projects. Overall, the midterm review—though noted as a useful instrument for learning by the Agencies—is not being used as often as it should.

The GEF Agencies generally comply with the GEF IEO's terminal evaluation guidelines for full-size projects. Compliance with the guidelines was

assessed for terminal evaluations that had been prepared in or after October 2017, six months after guideline issuance. A total of 270 terminal evaluations—185 for full-size projects and 85 for medium-size 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 the preparation of the terminal evaluation is weak (table 2.2). Similar patterns of compliance were observed for terminal evaluations of medium-size projects.

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-size 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 generally do not provide the GEF project ID number, but it is easy to ascertain this based on the World

Table 2.1 Percentage of completed projects with PIRs and midterm reviews, by overall outcome rating

Availability of PIR/midterm review	Project rating		
	Highly unsatisfactory/ unsatisfactory	Moderately unsatisfactory/ moderately satisfactory	Satisfactory/ highly satisfactory
At least one PIR submitted for project	88 (50)	90 (50)	98 (50)
PIR submitted for each year project was under implementation[a]	62 (50)	68 (50)	80 (50)
Projects with a midterm review	26 (50)	36 (50)	38 (50)
Full-size projects	36 (28)	44 (32)	49 (35)
Medium-size projects	14 (22)	22 (18)	13 (15)

Source: Review of PIRs carried out for this evaluation and for GEF IEO 2022b.

Note: Number of observations in parentheses.

a. Gaps in the first year of implementation are not considered, as projects may not complete a PIR for this year.

Table 2.2 **Level of compliance on reporting requirements for terminal evaluations of full-size projects**

Compliance dimension	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
No. of terminal evaluations reviewed	19	73	25	15	33	20	185
Timeliness of conduct	partially	fully	partially	mostly	mostly	partially	mostly
General information	fully	fully	fully	fully	mostly	fully	fully
OFP feedback (national projects)	noncompliant	partially	partially	noncompliant	partially	noncompliant	partially
Feedback of other stakeholders	mostly	mostly	mostly	mostly	mostly	mostly	mostly
Theory of change	mostly	partially	fully	mostly	mostly	mostly	mostly
Sources of information (methodology)	fully	mostly	mostly	mostly	partially	mostly	mostly
Reporting on outcomes	mostly	mostly	fully	mostly	fully	mostly	mostly
Reporting on sustainability	mostly	fully	mostly	mostly	partially	partially	mostly
Reporting on project M&E	mostly	fully	mostly	mostly	mostly	partially	mostly
Reporting on Agency performance	partially	fully	partially	mostly	mostly	mostly	mostly
Cofinancing	mostly	mostly	mostly	mostly	mostly	mostly	mostly
Social and environmental safeguards	partially	mostly	mostly	noncompliant	fully	partially	partially
Lessons and recommendations	mostly	mostly	mostly	mostly	mostly	mostly	mostly
Provide performance ratings	fully	mostly	mostly	fully	partially	partially	mostly
Consistency of ratings	mostly	mostly	fully	mostly	fully	mostly	mostly

Source: review carried out as part of this evaluation and APR 2020. OFP= GEF Operational Focal Point.

Note: ■ = fully compliant; ■ = mostly compliant; ■ = partially compliant; ■ = noncompliant. OFP = operational focal point.

Bank's project ID number which is linked with the GEF's. Other gaps are more substantial. For example, most terminal evaluations prepared by Agencies do not report on whether the evaluation team sought and incorporated input from the relevant GEF operational focal point during preparation of the terminal evaluation. As a result, it is difficult to ascertain the extent to which the GEF Agencies are complying with Minimum Requirement 4 (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 FAO, UNDP, UNEP, 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.

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 2.4). 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. Among the GEF

Figure 2.4 Trend in overall quality of terminal evaluations
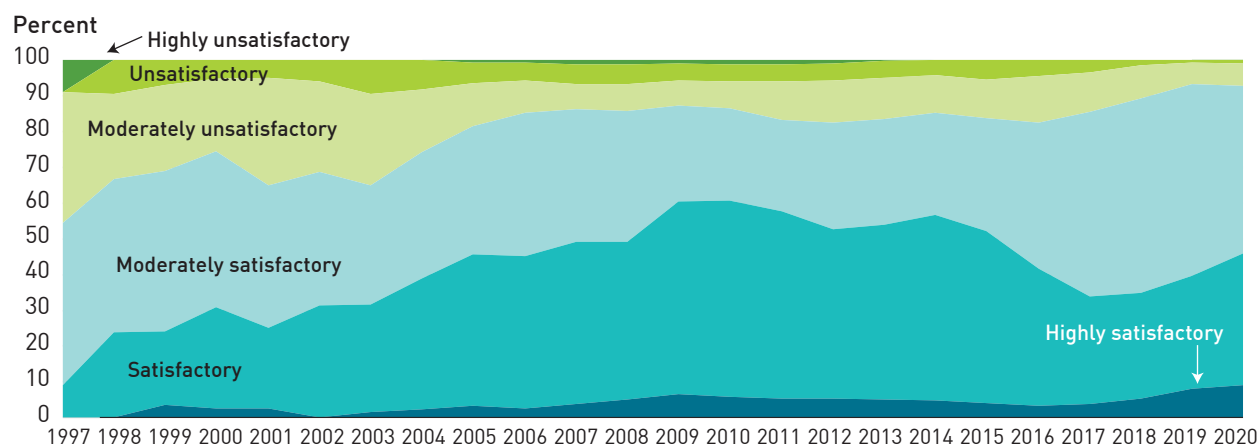
Source: GEF IEO terminal evaluation review data set 2021.

Note: *n* = 1,774. Data presented are three-year rolling average of terminal evaluation quality ratings, based on year of terminal evalution completion.

The latter have the advantage of receiving feedback from its evaluation unit.

SELF-EVALUATION SYSTEMS AND LEARNING

The types of learning taking place in Agencies may be broadly classified as learning about doing things right and learning about doing the right

Agencies, compared to the period before 2010, terminal evaluations prepared by UNEP have improved in quality (table 2.3).

Overall, terminal evaluations for full- and medium-size projects are of similar quality. The quality of terminal evaluations of World Bank–implemented medium-size projects was assessed to be somewhat lower than for its full-size projects.

Table 2.3 Terminal evaluations rated in the satisfactory range for quality, by Agency and project modality

GEF Agency	Full-size projects						Medium-size projects					
	Through 2009		2010 on		All years		Through 2009		2010 on		All years	
	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.
ADB	0	0	93	15	93	15	0	3	0	0	0	3
FAO	0	0	96	25	96	25	0	0	82	11	82	11
IDB	0	0	70	17	70	17	0	0	60	5	60	5
IFAD	0	0	80	20	80	20	0	0	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	0	0	85	33	85	33
World Bank	88	136	89	265	88	401	76	79	66	44	72	123
Joint/other	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 set 2021.

Note: Satisfactory range includes highly satisfactory, satisfactory, and moderately satisfactory ratings.

things. Doing things right relates to whether projects are being implemented as planned; doing the right things relates to the effectiveness of interventions and understanding what interventions work, how, and why.

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 PIRs and midterm 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 PIRs trigger corrective actions; the focus is on activities and outputs. Midterm reviews also consider outcomes (to the extent possible) and review the project logic to provide recommendations. Midterm 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 by year to review 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. GEF Agency NGOs do not have these arrangements for systematic reporting at the portfolio level.

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

Terminal evaluations are not used adequately to support learning on doing the right things. Although terminal evaluations are intended to fulfill this purpose, the format is not adequate for this type of learning. Information provided in the lessons learned 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 midterm reviews have some potential to facilitate learning on doing the right things. However, they are not conducted for the majority of projects. PIRs 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 the right track.

Agency evaluation units 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 midterm reviews are also used. These reviews are an important element of the independent evaluation units' mandate in development finance institutions and UN entities. Impact evaluations and technical

reviews also help with learning about what works, because they include outcome and impact data collected after project end.

Databases of lessons learned are available, but are reported by the Agencies to be hard to search. The 2020 Evaluation of Knowledge Management in the GEF (GEF IEO 2022a) and the Evaluation of the GEF Portal presented in chapter 3 of this report identified gaps in the search-related abilities of the Portal and make a case for upgrading its capabilities for knowledge sharing and learning.

Learning practices play an important role in learning on doing the right things, because they create experiences for staff to update their assumptions and mental models on what works. Stakeholder workshops and face-to-face debriefings between project staff, evaluation staff, and national counterparts are common practice at FAO, FECO, UNDP, and UNIDO. WWF-US runs regular learning workshops to update the project's theory of change (box 2.5). IDB organizes an end-of-year open house at which all evaluations that have been produced throughout the year are presented. ADB hosts events to discuss evaluations based on what works and what does not." At EBRD, the independent evaluation unit produces a continuous stream of webinars to share best practices. Some Agencies such as ADB, IDB, and UNIDO have brown bags/green bags to share information on good practices and failures from projects.

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 the GEF IEO and the GEF STAP. These reviews are valued highly by the GEF Agencies. Additionally, IW:Learn is seen as highly effective in enabling learning across the partnership.

Box 2.5 WWF-US uses theory of change for learning

The 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. Situation and results chain diagrams are typically developed using Miradi software. Miradi helps project managers produce key outputs such as situation analysis diagrams, results chains (theory of change), action plans, and monitoring plans in line with WWF standards and procedures.

In the context of GEF projects, two RBM specialists, serving in 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. They also support GEF project teams in updating the theory of change during annual reflection and learning workshops. The development of situation diagrams and results chains has helped WWF in good project design, the selection of appropriate indicators for the results framework, and learning practices within projects.

Source: Interview with Amelia Kissick, WWF-US.

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. 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 could be of great value for mutual learning.

Candor is an important factor that facilitates learning, but it is rarely incentivized. The key measure of success in Agencies is total resources and number of projects managed. This organizational logic informs project design and management, with

the result that evaluation comes to be seen as a necessary requirement rather than an opportunity to learn. Project staff have incentives to move their projects along without issue and in their receiving 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.

Agencies are exploring ways to promote candor in self-evaluations. 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 aims to apply an impartial perspective to assess whether available evidence justifies the ratings. The re-rating of project performance 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. As another example, IDB has created a Development Effectiveness Unit, which supports projects from design to post-evaluation. 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 culture of learning from evaluation.

Other Agencies are investing in training project management staff. The UNDP evaluation unit has developed an evaluation certification training course that is mandatory for all UNDP staff whose work involves M&E. Investment in capacity building and skills development in quality evaluation has helped build an evaluation culture among project and program staff—creating an environment in which inquiry, evidence, and

learning are valued as essential to good management, but not limited to use in reporting. ADB, UNDP, and the World Bank have instituted evaluation awards to recognize teams and departments that have excelled in producing high-quality self-evaluation reports (box 2.6).

Box 2.6 Asian Development Bank's excellence in evaluation award

The ADB's Independent Evaluation Department (IED) introduced annual IED awards in 2014. These recognize ADB teams and departments that have excelled in producing high-quality self-evaluation reports. Candidates are drawn from the terminal evaluations validated by the IED. The assessed criteria include adherence to evaluation guidelines, adequacy of evidence and analysis, and quality of lessons and recommendations. Awards are given out in five categories including best reports for sovereign operations, best reports for nonsovereign operations, and best operation department on self-evaluation. All ADB staff, board, management, and consultants are invited to the award ceremony, and the award is made by IED management, after reading out the justification for nomination of the reports. Winners receive a trophy, plaque, and certificate. IED extensively promotes the winning reports on ADB's internal communications platform as well as on social media. Additionally, an awards ceremony video is produced and disseminated both internally and externally.

The awards have increasingly gained traction within ADB, as evidenced by the growing number of attendees at the ceremony over the past years, reaching more than 100 staff and consultants at the 2020 virtual ceremony. Among the participants were the vice presidents and director generals of most operational departments.

The awards have had a positive effect on 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.

Source: Interview with Garrett Kilroy.

The example of collaborative preparation of PIRs at UNDP illustrates that the right incentives offered by management can promote an evaluation culture (box 2.7).

Box 2.7 Collaborative PIR reporting in UNDP improves candor

UNDP puts a strong emphasis on stakeholder-led approaches in evaluation. One mechanism to achieve this is its collaborative reporting and rating of PIRs. The reporting for the PIR in UNDP is conducted in collaboration by the GEF technical advisor, the program manager, the country office focal point, the GEF operational focal point, and—sometimes—the regional manager. Collaborative reporting and rating of the PIR helps UNDP ensure coherence and coordination within its 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 the collaborative process is greater candor in reporting. This is substantiated by the GEF IEO 2020 Annual Performance Report, which notes: "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 IEO 2022b, 30).

Sources: Interview with Pascale Bonzom; GEF IEO 2022b.

Evaluation of the GEF Portal

The 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 2018, 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).

Concerns related to the PMIS had been noted in several evaluations. The Second Overall Performance Study (OPS2) cited concerns related to data on cofinancing (Christoffersen et al. 2002). OPS3 noted that the PMIS did not capture information systematically and did not facilitate monitoring of the GEF portfolio, and that the GEF Secretariat and M&E staff did not find its data to be fully reliable (GEF IEO 2005). OPS4 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 2010b). It observed that the pace of improvement in the PMIS had been slow because of cost-cutting measures adopted by the GEF Secretariat and that the choice of delayed development was not optimal. OPS6 noted that the availability

and quality of information provided by the PMIS remained an area of concern (GEF IEO 2018). It further noted that the primary reason for the low quality of data was that data were being entered manually and not updated regularly. OPS6 highlighted the need for improving systems for project management data, monitoring, and knowledge sharing.

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

In an effort to address the shortcomings of the PMIS, and following up on a recommendation from the independent review, the GEF elected to participate in the World Bank's Financial Intermediary Funds Information Technology Project, an initiative aimed at consolidating the information

systems of similar secretariats housed at the Bank. After two years, the GEF dropped out of the effort, and engaged with the World Bank's Information and Technology Solutions (ITS) unit to develop an entirely new system to replace the PMIS. Work began in 2017, and the GEF Portal was launched at the start of GEF-7 in July 2018. The design of the Portal, and the procedures and rules regarding its use, are aimed at delivering the following benefits (GEF 2018):

- Enhanced project review and processing abilities to improve operational efficiency;

- Improved accessibility of data through capturing 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;

- 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;

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

At its launch, the GEF Portal had limited capabilities. Several features have been rolled out subsequently, and several other 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, and incorporates feedback and comments from across the GEF partnership.

3.1 Background

The PMIS was the predecessor of the GEF Portal. It started operation in 1998, when the GEF Council requested a list of projects the GEF had funded. From that data set, it became a web interface–based system in 2009. During GEF-4, $700,000 was earmarked for the special initiative of developing 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. Notably, the PMIS was not housed on the 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.

The independent review of GEF systems had recommended that the PMIS be integrated into the World Bank's Financial Intermediary Funds Information Technology Project (GEF 2011). The project management information systems of four funds—the Global Partnership for Education, the Global Agriculture and Food Security Program, the Climate Investment Funds, and the Adaptation Fund—were eventually integrated into the project. However, the GEF dropped out of the initiative after two years because development progress was slower than expected and did not account for the complexity of the GEF project cycle modalities and workflow.

Following this decision, the GEF continued to explore other options. Over time, there was greater support for developing a new system rather than revamping the PMIS. The GEF Secretariat entered into discussions with the ITS unit of the World Bank to develop the new system. Subsequently, the Secretariat signed a contract with

the ITS unit to develop the GEF Portal. At the start of FY 2017, $658,650—the remaining balance of the amount approved in 2005 plus $222,229 repurposed from the remaining balance of the allocation for the Performance-Based Allocation System—was available for development of the Portal (GEF 2017). This does not include the full costs of staff time involved in the development.

The GEF Portal task team within the GEF Secretariat led the process of Portal conceptualization. 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 the World Bank ITS unit, held consultations with Secretariat staff from the program and policy and operations units, the GEF Agencies, the GEF IEO, and the GEF STAP. The team that managed the PMIS also provided inputs to this conceptualization process.

ITS focused on developing a minimally viable product—within the resource envelope provided to it—as a basis for 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 offered limited functionality. Subsequently, several additional features have been introduced. These rollouts have been accompanied with online workshops to introduce the new features and to obtain feedback on areas of further Portal improvement.

3.2   Key questions

The key questions of the evaluation were as follows:

- To what extent has the Portal facilitated enhanced project review and processing abilities and contributed to improvement in operational efficiency?

- To what extent has the Portal improved access to data for key stakeholders in the GEF partnership?
- How well are the key GEF programming strategies and policy requirements integrated in the Portal?
- To what extent has the Portal enhanced transparency through increased real-time availability of data?
- To what extent are the users of the Portal satisfied with its performance?

3.3   Methodology

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 users on new features.

SURVEY OF THE PORTAL

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, running different queries and data downloads, and assessing the capabilities of the Portal. The survey was conducted during October 2020.

STAKEHOLDER INTERVIEWS

During the interview with the GEF Portal team, key stakeholder and user categories were identified. For each category, key informants were identified who were then contacted and interviewed (annex D). In all, 33 key informants representing different categories of GEF Portal stakeholders were interviewed (table 3.1). Each key informant was asked questions to capture his or her unique perspective on design, implementation, and use of the Portal and/or issues relevant to assessing its performance.

ONLINE SURVEY

The online survey focused on gathering information on the experience of 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 questionnaire consisted of both open-ended questions and questions that used a rating scale. It covered topics such as website

Table 3.1 **Stakeholder coverage through interviews**

Stakeholder category	Number interviewed
GEF Secretariat	15
Program manager	7
Portal team	3
RBM team	3
Management	1
Country relations	1
World Bank ITS	2
GEF IEO	4
GEF STAP Secretariat	1
GEF Agency staff	8
Peer portal (Green Climate Fund)	3
Total	33

navigation, completeness of web pages, information accessibility, process of proposal submission and appraisal, uploading and downloading of documents, and transparency.

The online survey was administered to active users of the Portal. A list of the email IDs of users who have access to the GEF Portal was acquired from the ITS unit responsible for managing access. The list was cleaned by excluding repeated and inactive email addresses.[1] After these exclusions, the number of unique recipients of the invitation to participate in the online survey was estimated to be 845. However, all those who have access are not active users; the number of active users who use the GEF Portal at least once a month is estimated to be 667. Overall, 158 active users (24 percent) responded to the survey, and 133 (20 percent) answered all the survey questions (table 3.2).

For the questions that involved the use of a rating scale, six-point scales were used without a neutral option, but with a seventh option generically labeled "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 made little sense. For example, when assessing the importance of an attribute, possible responses ranged from "not important at all" to "extremely important." The survey results

---

[1] The prevalence of inactive email addresses was assessed for addresses in the GEF and World Bank domain, and these data were extrapolated to estimate inactive addresses in the list acquired from the ITS unit. This information was then used to assess the response rate of the active users.

Table 3.2 Participation figures for survey

Total email addresses listed as users having access to the Portal	1,001
Estimated email addresses listed after excluding repeats and inactive addresses[a]	845
Estimated active users who use the Portal at least once a month	667
Total respondents to online survey (% of estimated survey recipients)	205 (24%)
Active users who responded to online survey (% of estimated active users)	158 (24%)
Active users who answered all survey questions (% of estimated active users)	133 (20%)
Active users who answered all survey questions as % of those who started the survey	84%

a. This is an estimate because inactivity could be ascertained only for GEF and World Bank domain–based email addresses. For other Agencies, the rate of inactivity found for the GEF and the World Bank was extrapolated.

presented in this report exclude "do not know/ unable to assess" responses.

COMPARATIVE ANALYSIS

A comparative analysis was conducted to assess the quality and usefulness of the Portal vis-à-vis other peer portals. The evaluation team first searched for evaluations of other, similar portals that covered their performance, including evaluations that looked at these portals as a main or additional topic.[2] This search did not yield enough data for analysis; therefore, an alternative approach was adopted. The evaluation team examined three portals—the Green Climate Fund (GCF) 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—as external users.

These portals are used to manage information on projects 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 team attempted to interview those involved in the management of these portals to obtain their perspective, and was able to interview three key informants related to the GCF portal. No response was received 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.

LIMITATIONS

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 a willingness to be interviewed, the informant eventually dropped out. As a result, the perspective of the PMIS team is not reflected in this evaluation.

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 team members had to be canceled because of no-shows and nonresponses despite follow-up. As a result, the evaluation team is not able to fully address issues related to the ITS role and performance and to fully reflect the ITS perspective in this evaluation.

---

[2] The search covered organizations such as EBRD, the UN, IDB, the International Finance Corporation, the European Investment Bank, the Nordic Development Fund, International Renewable Energy Agency, and the Green Climate Fund.

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

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

3.4 Findings

PROJECT PROPOSAL SUBMISSION AND REVIEW

The process for submission and review of project proposals has shifted to the Portal. The PMIS offered little support for project proposal submission and review activities (GEF IEO 2017b). Rather, proposals were submitted via email, 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 Agencies submit proposals, and 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 them, and when. This audit trail and chain of custody facilitates audit of actions taken—or not taken—at the Portal and establishes accountability. During the COVID-19 pandemic, the ability to submit project proposals and review submitted proposals online was 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.

DATA QUALITY

The Portal has streamlined the process of submitting PIRs, midterm reviews, and terminal evaluations, which may help in reducing submission gaps. Before the GEF shifted to the Portal, Agencies submitted these documents via email, with a correspondingly greater risk of documents being misplaced or uploaded incorrectly to the PMIS. Further, it was difficult to obtain an overview of submission gaps in real time; it would take considerable time and effort to ensure compliance. Submissions are now made by the Agencies directly at the Portal, and it is possible to track document submission more effectively.

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. Data entry tasks are linked to further movement of a given project along the project cycle. That is, project submitters (Agencies) and project reviewers (GEF Secretariat) are unable to move the project to the next step without entering complete information for the preceding step. This guides those responsible for entering information—who also have an interest in moving the process forward—to ensure all fields are filled in. Auto-validation is inbuilt for many fields so those entering information receive real-time feedback on potential errors 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 core indicator data has been filled and allows the Agencies to reach the next-stage web page only after the relevant information has been entered. Validation checks incorporated in the Portal design cover financial envelopes, the GEF System for Transparent Allocation of Resources (STAR) limits, and impact program set-asides. Use of a menu of response options limits the scope for errors in data entry by restricting 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 the 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. For such projects, the data may be entered only by the lead Agency. The respective Agencies thus need some offline back and forth on the information that the lead Agency is to enter on behalf of the other Agencies.

The availability of real-time data on project cycle status has improved, although 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 (excluding "do not know/unable to assess" responses) agreed that real-time availability of data on GEF activities has improved as a result of the GEF Portal (see table E.3 in annex E). Many actions related to the project life cycle—such as project proposal submission; project appraisal; and submission of PIRs, midterm 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 improves the quality of data on project cycle milestones, project financials, and the status of project proposals, and provides it in real time.

The quality of historical data remains a concern, despite efforts at improvement. The Portal team filled some gaps in the PMIS project data through collaboration with the Agencies. However, several challenges remain. Some problems have arisen because of data migration from the PMIS to the Portal. For example, original document upload dates were automatically revised by the system: the date of the document is the date of migration, and every time the system is refreshed, a new date based on that latest system update is listed. Documents that used to span several years suddenly have the same date, making document tracking by date difficult.

Several users who were interviewed reported errors in data outputs. These are fixed when they are brought to the attention of the Portal team, but the turnaround time is perceived to be long, which discourages some from reporting on these issues.

Portal users have the option of keeping a document as a draft version until they deem it final. Once final, Agencies need to go through the GEF Secretariat and the ITS unit to be able to edit any data entered. This ensures data integrity and avoids changes being made to already final submissions. At the same time, it disincentivizes corrections to uploaded information. Moreover, sometimes, when the wrong report has been uploaded, Agencies will subsequently upload the correct version without notifying the GEF Secretariat and ITS so the incorrect version can be removed. The presence of multiple versions of a document can be confusing for a user who may need to spend additional effort in determining which of the versions is correct.

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 requests for new data products could be addressed only by the PMIS manager. Dependence on any one person to such an extent entails risk. The PMIS manager and another member of the PMIS team were empowered to make changes to the data and related information. While this ability had tactical utility in addressing 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 regulates who enters data, about what, and when. It also has an auditable trail of who did what, when, and why.

NAVIGATION, VISUAL APPEAL, AND ACCESSIBILITY

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 (see table E.2 in annex E). It is easy for the user to identify an HTML link; web pages are well composed and have 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 to the online survey were satisfied with the visualizations, displays, completeness, and clarity of the Portal web pages (figure 3.1).
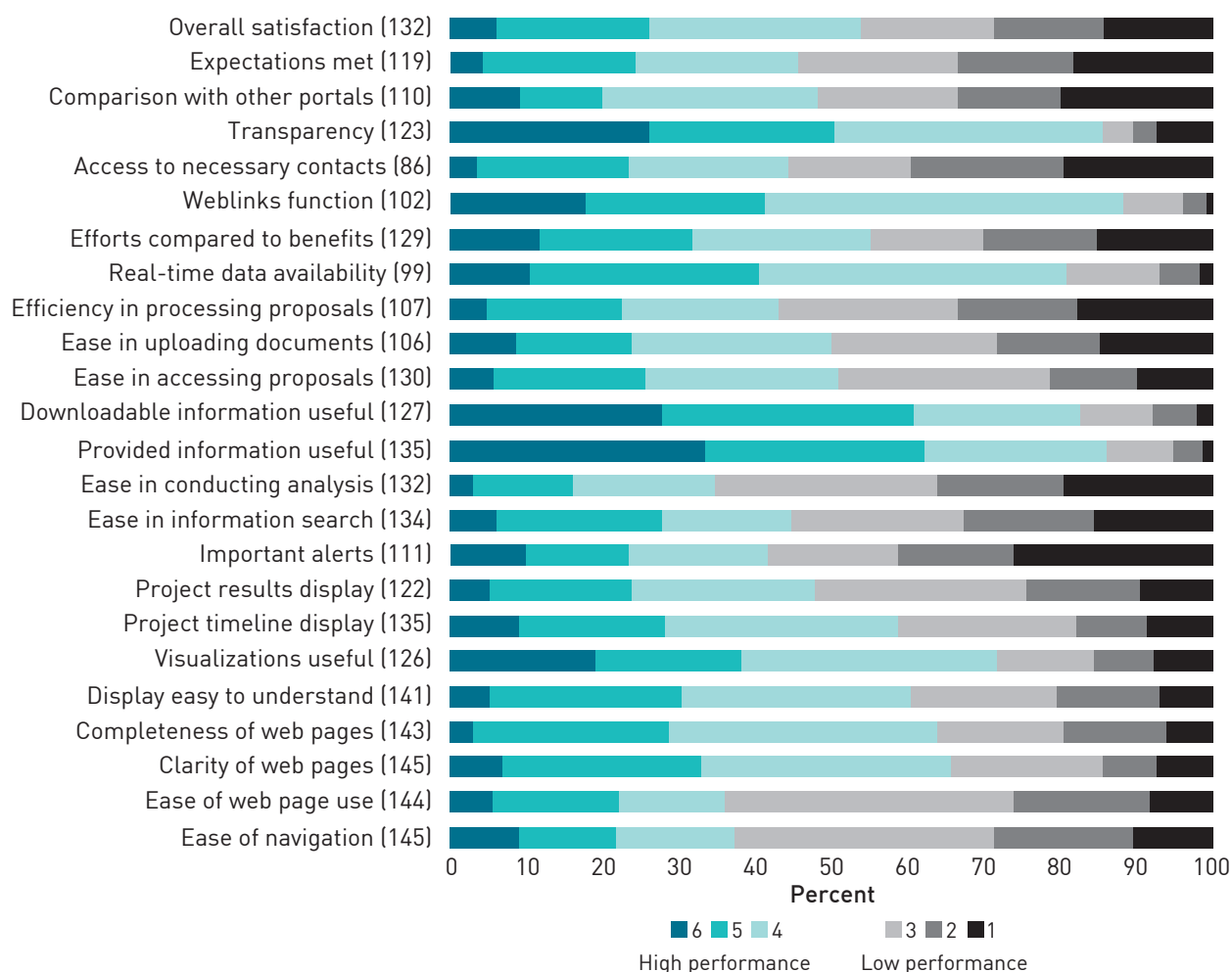
User perception of ease of navigation is varied. Of the online survey respondents, only a minority felt that it was easy to navigate the website and that the web pages of the Portal were easy to use (see figure 3.1 and table E.3). Thirty-seven percent of the respondents assessed navigation of the Portal 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 assessed navigation to be somewhat to very easy (n = 38), 30 percent of those who use it at a lower frequency found 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; part of the difficulty may be due to the arrangement of pages.

The GEF Portal compares well with peer portals on criteria such as navigation, visual appeal, and accessibility. The evaluation assessed the design and performance of the GEF Portal with that of peer portals on several criteria (see table E.2 in annex E). 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.

User assessment of the quality of the design of the GEF Portal is mixed. Fifty-two percent of the online survey respondents—again, setting aside "do not know/unable to assess" responses—felt that the quality of the GEF Portal was slightly to much worse than that of other portals with which they were familiar (see figure 3.1 and table E.2). 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 the other groups described the quality of the GEF Portal to be slightly to much better than other portals; the difference is statistically significant at a 90 percent significance level. Based on the

Figure 3.1 User perceptions on portal performance, on a scale of 6 to 1

Overall satisfaction (132)
Expectations met (119)
Comparison with other portals (110)
Transparency (123)
Access to necessary contacts (86)
Weblinks function (102)
Efforts compared to benefits (129)
Real-time data availability (99)
Efficiency in processing proposals (107)
Ease in uploading documents (106)
Ease in accessing proposals (130)
Downloadable information useful (127)
Provided information useful (135)
Ease in conducting analysis (132)
Ease in information search (134)
Important alerts (111)
Project results display (122)
Project timeline display (135)
Visualizations useful (126)
Display easy to understand (141)
Completeness of web pages (143)
Clarity of web pages (145)
Ease of web page use (144)
Ease of navigation (145)

Percent

■ 6 ■ 5 ■ 4   ■ 3 ■ 2 ■ 1

High performance   Low performance

Source: Online survey administered November 2020.

Note: Figures in parentheses are numbers of responses. 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.

supporting evidence gathered through interviews, it appears that the difference in perspective is driven by inadequate prioritization of the needs of the GEF Agencies in Portal design.

DEVELOPMENT PROCESS

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 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 see how it worked, but this experience was insufficient and there was little time to address any 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 more work for users. This issue was subsequently resolved by 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 obtain the 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 reported that they still use the PMIS for several functions because the Portal is not fully functional and the PMIS seems better suited for certain tasks. A shared frustration among several respondents was that there is no clear indication as to when the Portal will achieve full functionality.

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. In the years since the Portal's launch, several additional features have been rolled out, accompanied by workshops and training on the use of the new features. The Portal team notes that it is working on an improved reporting platform and on 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, they regard the Portal as a work in progress and have identified several areas where it ought to be developed and glitches need to be fixed. These include the following:

- Development of a comprehensive system of alerts on status changes and approaching deadlines/events;
- The ability to batch download documents;
- Enhancing the capabilities of the search function;
- Ensuring that calculations presented in 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 ITS attribute the slower development to the limited resources with which they had to work. During GEF-4, the GEF Council allocated $700,000 for the Portal through a special initiative (GEF 2017).[3] Subsequently, $222,229 was repurposed from the special initiative for the Performance-Based Allocation System to the budget for the Portal. Of the $922,229 allocated, $658,650 was available at the start of FY 2017 for development of the Portal; at the end of FY 2020, $299,000 remained unspent. These cost figures are a low bound and do not reflect the full costs incurred in development of the Portal.

The present search function is inadequate to facilitate users in accessing information from the Portal. The current search engine offers the possibility of filtering and aggregating (sum, average, etc.) variables by country, focal area, etc. The material available for searching 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 in the Portal, and 65 percent found it slightly to very difficult to analyze content within the Portal. Users would like a feature to facilitate sorting and filtering of documents within the Portal. For example, they want to 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 follow:

- PIRs, midterm reviews, and terminal evaluations are difficult to find.

---

[3] Referred to in the budget document as "GEF Management Information System."

- Documents that are saved as annexes are not well integrated with the main document and are difficult to find.
- The Portal is expected to provide the ability to use georeferenced data; so far, it allows only for upload of maps, text files, and pictures but does not upload georeferenced shape files that digitally demarcate geographical areas covered by the GEF project.
- 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 receive a list or tabulation.
- Search by indicators and project components is not enabled. For example, a user cannot know which projects within a portfolio were aimed at capacity building.

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 conventions followed for the Portal, documents cannot, in most instances, be identified without opening them. It would be useful if, when searching for documents, their names were more obvious so a user could easily infer the content. Developing stricter naming conventions would help alleviate this problem. Dates are also a challenge because original document dates are automatically revised by the system. The date of the documents is the date of migration; and every time the system is refreshed, documents receive a new date. Documents that used to span several years suddenly all have the same date. This is an issue, particularly when tracking published documents.

The Portal does not allow bulk download of documents, which is a major barrier for users who analyze content. The currently available function enables the download of data in multiple formats as well as of core indicator reports. Users are generally appreciative of this ability. However, bulk download of documents—which is an important function for users conducting a portfolio-level analysis—is an area for improvement. In the PMIS, bulk downloading was enabled. In the new Portal, it is not: users are required to download files one at a time. 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 took more than five workdays to download the full document set.

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 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, the system needs to provide alerts that are targeted to other action takers within the partnership as well. It could also provide alerts further upstream of approaching deadlines and could incorporate a system of alerts and reminders for activities that need to be completed within a time frame instead of an exact date (e.g., terminal evaluations). Improvements of this nature in the alert system would provide information on status update and would have a positive affect on users' perceived quality of the workflow within the Portal.

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—for example, progress over time, region,

theme, or focal areas. Along with an easy view of the current portfolio based on real-time data, it will facilitate user-friendly project data access.

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 migrating the data, the team conducted verification checks to assess whether the data had been migrated free of errors. Through this process, several issues were found and corrected. But several challenges related to data quality remain. For example, in the absence of guidelines on the categorization of data, the categories to classify data are often used inconsistently. In several instances, data given in the Chief Executive Office (CEO) letter for endorsement/approval do not match the corresponding data provided in the Portal. The quality of historical data, despite efforts at improvement, remains poor and data are 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, small island developing states were reported as least developed countries, and vice versa. These problems are indeed fixed once they are brought to the attention of the Portal team, but it means that data users have to make an extra effort in checking whether downloaded data are correct and can be used for further analysis.

UTILITY AND TRANSPARENCY

User perceptions of their overall Portal experience have improved since launch. Several interviewees noted a marked increase in positive experiences with the Portal since mid-2019, when users were able to gain full access to the Portal (submit, draft proposals); they note that since then the Portal has broadly served their needs. However, there is dissatisfaction among users—including from the Agencies, the GEF IEO, and the GEF Secretariat—with what they perceive as numerous low-level glitches and challenges. Several interviewees noted that the new Portal should have been a dramatic improvement in operational efficiency over the PMIS, and its performance closely aligned with industry standards for such a portal, but that their expectations had not been met.

Although users are generally satisfied with the responsiveness of the Portal team, several are dissatisfied with what they perceive as a lack of clarity in the approach used to identify and prioritize problems. A system that is fully responsive to user feedback should meet the following criteria:

- Clear process for recording complaints;
- Tracking of progress toward resolution;
- 100 percent or nearly 100 percent response rate;
- Commitment to a time frame for providing a response.

While there is a process in place for users to report issues they encounter in using the Portal, the arrangements are not as systematic as they could be. Therefore, user experience of the Portal team's performance in addressing 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, although most respondents found the team to be responsive. The feedback to the Portal team, including the ITS staff designated for the purpose, is usually through emails. Users from the GEF 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 this need 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 very 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.

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 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 expressed the need for more proactive communication by the Portal team. Several respondents stated that there should be at least a 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 function as an open data source, several interviewees mentioned the need for additional efforts to make data transparent and easily distributed—for example, by giving users more detailed guidance on the data being presented. In part, this could be accomplished through a codebook for spreadsheets describing what can be found in the different columns.

The Portal has shifted some of the work burden to the Agencies. Analysis of online survey data shows that GEF Agencies account for 65 percent of the active users and 54 percent of the frequent (at least a few times a week) users of the Portal. Agencies are responsible for submitting project proposals,

PIRs, midterm reviews, and terminal evaluations to the GEF. Previously, they submitted these documents via email, which were then uploaded to the PMIS by GEF Secretariat and GEF IEO (for terminal evaluations) staff. While the Portal has eliminated the need for Secretariat staff to upload project documents, Agency staff now spend more time than they did before uploading documents. Although the new approach reduces errors in submission, the Agencies feel that the additional benefits have not been commensurate with the additional effort at their end.

DATA ACQUISITION AND ACCESS

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. The Portal feeds some of the information to the GEF website, which is accessible to the public. However, the availability of data to users is more restricted. During interviews, users noted that sometimes data are simply not available 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 differences in the disclosure policies of the GEF and the GEF Agencies, resulting in different levels of access to data for various stakeholders.

Users are generally happy with the level of transparency the Portal provides. Eighty-five percent of the active users of the Portal among the survey respondents 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.

Connectivity is a major concern for many users. Earlier, most users had difficulty in working with 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 became more acute when there was heavy use of the Portal, such as around deadlines. Much of this problem was due to a relatively narrow bandwidth assigned to the Portal; this was subsequently addressed by ITS and sufficient bandwidth is now provided to the Portal. At the user end, challenges remain in countries with poor Internet connectivity, which disproportionately affects users in least developed countries and remote areas. Connectivity becomes a major issue with the decentralization of data entry at the Agency level, especially for project managers who are posted in recipient countries. The Portal needs to accommodate the technical constraints of most, if not all, of its users.

Conclusions and recommendations

4.1 Conclusions

GEF AGENCY SELF-EVALUATION SYSTEMS

Conclusion 1: How do policy frameworks in the GEF Agencies support their self-evaluation systems? The evaluation found that the policy frameworks of GEF 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.

Conclusion 2: 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. The GEF IEO Annual Performance Report 2020 (GEF IEO 2022b) showed that PIRs generally give overly optimistic performance ratings. This evaluation found that even though the ratings

provided in the PIRs 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. Midterm reviews are generally regarded as credible and useful, but they are not prepared for a majority of full-size 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 the right things.

Conclusion 3: To what extent do the Agency self-evaluation systems meet GEF requirements according to the relevant GEF policies and guidelines? Agency 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 into 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 making comparisons across the Agencies. The

midterm reviews are mandatory for full-size projects but are not prepared for a majority of the others. The submission rate of PIRs is somewhat lower for low-performing projects. The quality of terminal evaluation reports is improving, although reporting on whether feedback from operational focal points has been sought and addressed, and on the application of safeguards, is generally weak.

Even though midterm reviews are required for full-size 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 on impact programs—to deliver GEF support makes the gaps in guidance for conduct of midterm reviews even more salient. These programs are more likely to benefit from midterm 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 the midterm review 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, the GEF IEO will need to work more proactively with the Agencies in providing them with feedback on the quality of terminal evaluation reports and on compliance gaps.

Conclusion 4: **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.

Some Agencies, such as ADB, UNDP, the World Bank, and WWF-US, 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 inter-Agency workshops and other instruments for knowledge sharing.

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 the 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.

GEF PORTAL

Conclusion 1: **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, 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 peer portals on some of the technical parameters on which 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.

Conclusion 2: Several interview respondents and online survey participants expressed frustration with 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 of the process 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 users so that they are better aware of the process and have a clearer understanding of the road map to further development of the Portal.

Conclusion 3: The advent of the Portal shifted the burden of data entry 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 with which they have to work. At the same time, there are substantial resources for the special initiative for the Portal that have not been fully utilized. GEF management needs to assess whether additional support is necessary and how speedier development of the Portal can be aided.

4.2   Recommendations

GEF AGENCY SELF-EVALUATION SYSTEMS

Recommendation 1: The GEF Secretariat and the GEF Agencies should strengthen use of midterm reviews for learning and adaptive management. The evaluation shows that, despite their potential, midterm reviews are conducted in a limited number of instances and the guidance on midterm reviews is inadequate. The Secretariat should

provide more guidance to the Agencies on the conduct of midterm reviews, should share good practice examples, and should track timely conduct and submission of midterm reviews. The Agencies should conduct midterm reviews for GEF-supported projects, as mandated by the 2019 GEF Monitoring Policy.

Recommendation 2: 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 a theory of change, and design and implementation of M&E plans. Similarly, enhancement of the search and analysis capabilities of the GEF Portal may be useful in strengthening learning across the partnership. The Secretariat may also need to bring in Agencies, the GEF IEO, and/or the GEF 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 a theory of change, and the GEF IEO on guidance on midterm reviews.

GEF PORTAL

Recommendation 1: 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 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 toward resolution.

Recommendation 2: GEF management should develop and implement a time-bound plan to speed up the development of the Portal. Users of the GEF Portal perceive that, despite significant progress, the development of Portal has continued for a long time. 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. GEF management should assess how best it may speed up the process and implement its plan.

GEF Agency good practices

Box A.1 IDB enhances development effectiveness through guidance and support infrastructure

Context. As a development institution, the Inter-American Development Bank (IDB) is constantly faced with the challenge of accurately assessing the Bank's performance and ensuring its development effectiveness. While already following the international standards and good practices for evaluation of the Evaluation Cooperation Group, IDB acknowledged that more work could be done internally to fully overcome deficiencies and improve its development effectiveness. In 2008, IDB introduced the Development Effectiveness Framework (DEF), which is embedded in the project cycle of all IDB projects. With the purpose of setting clear standards for the evaluation of IDB interventions, providing guidance for project teams, and monitoring IDB progress on key development effectiveness indicators, the DEF encompasses various tools, offering a comprehensive guide for project teams.

Good practice. IDB established the Office of Strategic Planning and Development Effectiveness (SPD) to provide essential support to teams throughout the project cycle, including on the DEF and evaluation guidelines and their application. Core products featured in the DEF and supported by the SPD are the development effectiveness matrix at the design stage of a project and the project completion report (PCR) to assess quality at intervention exit. The matrix is a checklist comprising best practice standards to design and monitor evidence-based projects. Its dimensions are evaluability, strategic alignment with the Corporate Results Framework, project risks, and additionality of the project to ensure and control the quality of IDB operations. The SPD offers on-demand trainings on the preparation of high-quality PCRs. These sessions are organized around presentations on the core and noncore criteria of the PCRs, followed by breakout work groups where project teams can work on individual PCRs with the support of an SPD facilitator. The SPD also offers support in conducting impact evaluations; to date, IDB has carried out more than 600 impact evaluations, which are a valuable source for determining institutional effectiveness.

Effects. "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," IDB's Leonardo Corral observed. As a result, the SPD regularly receives requests for training sessions from all sectoral divisions within IDB, with project teams appreciating the support. Furthermore, the SPD has become the "go to" department for project teams for all questions related to the DEF and its tools, including questions on impact evaluations, evaluability at entry, monitoring during execution, and PCRs at closure.

Sources: Interview with Leonardo Corral, IDB; IDB 2008.

Box A.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. The Foreign Economic Cooperation Office, Ministry of Environmental Protection of China (FECO) also joined the GEF partnership only a few years ago and has not yet submitted any terminal evaluations. Therefore, both institutions are particularly eager to exchange information 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," explained René Gómez-García Palao. As a new GEF Agency, CAF highly values GEF standards and policies on self-evaluation processes, as they help the Agency identify potential flaws in its own system and foster internal procedures to continuously improve. Similarly, FECO has systematically leveraged its pre-accession harmonization to review and compare GEF requirements and international good practice with its own approaches. In this way, FECO refined

its own evaluation system for GEF projects with an extensive review function by its GEF coordination unit and learning meetings with various stakeholders. Within the GEF ecosystem, CAF particularly appreciates platforms such as the inter-Agency retreats 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 fluid communications among the entire GEF partnership. The retreats' presentations and discussions allow for learning from past experiences and, at the same time, help the new Agencies become fit for the future. This includes discussions of upcoming trends and coordination of future actions that are aligned with the GEF agenda. Lastly, various Agencies have noted 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 Gómez-García Palao, CAF; and Liu Lei, FECO.

Box A.3 **Two Agencies innovate their self-evaluation systems with the application of theory of change approach**

Context. Despite shifts from a logical framework to a theory of change model in the development field, 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) emphasizes the need to increasingly look at projects and programs from a systems perspective, focusing on the interdependent configuration of factors that lead to outcomes rather than a simple cause-effect model. The World Bank and the United Nations Environment Programme (UNEP) are among the GEF Agencies that are innovating their self-evaluation systems by applying a theory of change methodology as both systems-oriented process and to construct systems models.

World Bank

Good practice. The theory of change approach was made mandatory in 2017 as part of the World Bank's reform of project terminal self-evaluations. The theory of change 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's operational policy vice presidency, with support from the independent evaluation function, encouraged the use of theory of change among operations staff. Enforcement of the mandatory requirement scaled up the uptake of theory of change as a tool.

Effects. Application of the theory of change was seen by the Bank's Independent Evaluation Group (IEG) to be generally successful, although there has been variation in practice across sectors/regions—meaning that the quality of theory of change can vary greatly. The IEG, as a result, frequently offers training clinics and gives support to teams in creating their theory of change. The use of a theory of change helps teams in various project phases. At project design, it helps improve clarity and communication around project logic as well as internal quality review processes to debate design logic; it is an especially valuable tool for engaging clients through a participatory

process. In evaluation, it helps teams debate what worked and did not work from a results (outcomes and impacts) rather than a delivery perspective (outputs). Over the long term, the World Bank would like to encourage teams to use a theory of change to articulate longer-term goals/higher-level outcomes, beyond the accountability function. It would like to see its application in understanding impacts beyond the project level (cumulative effects, replication/scale-up effects, etc.) as well as in identifying the type of indicator data to collect to track higher-level country outcomes.

UNEP

Good practice. UNEP evaluations are structured around a project or program's theory of change to assess the causal logic of the intervention and to determine whether all external factors affecting outcomes, impact, sustainability, and up-scaling have been carefully considered. This means that UNEP not only reconstructs a theory of change for the terminal evaluation, but the evaluation method uses the reconstructed theory of change to assess the effectiveness of the intervention as a guiding framework. It can also be used to inform other evaluation criteria. Of interest, the reconstructed theory of change in the terminal evaluation may add intermediate outcomes that were missing in the logic to better assess the validity of the causal pathways.

Effects. Standardized use of a theory of change, together with characterization of standard project performance ratings, has improved the consistency of evaluative judgment and the quality of evaluation reports. Specifically, use of the theory of change 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 UNEP's independent Evaluation Office with use of the theory of change has been fed back into formal Agency project and program design requirements—leading to improvements in design quality.

Sources: Interviews with Stephen Hutton, IEG; Michael Spilsbury, UNEP; and Amelia Kissick, WWF-US; Patton 2010.

Box A.4 IW:Learn, an innovative learning infrastructure

Context. IW:Learn was created when the GEF made a virtue out of necessity with regard to the international waters focal area. Unlike the other GEF focal areas, international waters did not on its establishment in 1997 have an overarching convention or conference of the parties. In the absence of these, a vision of a global knowledge network for the focal area 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 also by 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 interproject 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 has hosted 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 has grown 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 interproject 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 the IW:Learn 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 the formerly separated domains of freshwater and marine ecosystems. Through means of knowledge production and sharing around projects, results and best practices are made explicit and accessible, hence contributing to overall portfolio performance and learning in the GEF partnership.

Source: Interview with Andrew Hudson, UNDP.

Box A.5 **UNIDO undertakes systematic meta-reviews for cross-project learning from terminal evaluations**

Context. GEF requirements for Agencies mainly focus on project evaluations. However, there is much value for future project design and other learning needs from a comparison of project lessons learned; such reviews are called systematic or meta-reviews. These reviews are likely to be done on an ad hoc basis, to focus on a comparative perspective and to draw insights around existing lessons learned.

Good practice. The United Nations Industrial Development Organization (UNIDO) conducted such a meta-review of its renewable energy–based minigrid projects in rural areas with the twofold aim of reviewing its experience as an institution and evaluating the lessons learned from each project to refine the design of future projects. Several projects and their evaluations were reviewed according to five key criteria determining long-term sustainability of minigrid projects. The results were captured in a working paper (UNIDO 2017), collaboratively written by staff from various UNIDO departments—further contributing to breaking down silos for future project design.

Effects. Various GEF Agencies have noted that the potential of project evaluations is not fully realized for cross-project learning. Generating cross-cutting lessons learned is an important step toward feeding learning 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 fosters learning between the independent evaluation unit and project managers involved in synthesizing lessons learned from projects and their evaluations. To scale this good practice from an ad hoc to a regular exercise, dedicated resource allocations will be needed.

Sources: Interviews with Rana Ghoneim and Jossy Thomas, UNIDO; UNIDO 2017.

GEF Agency profiles

Table B.1 **African Development Bank (AfDB)**

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with GEF as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by GEF unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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

(continued)

Good practice example	• Extensive peer review opportunities as part of Quality Assurance Framework
Areas for improvement	• More use of self-evaluation for learning between its independent evaluation unit, project staff, stakeholders, and beneficiaries • Shifting current focus on ratings and disconnect of ratings between project teams and independent evaluation unit more toward learning

Table B.2 **Asian Development Bank (ADB)**

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Good quality assurance mechanisms • Useful system for accountability and generating information for learning
Areas for improvement	• Project monitoring and evaluation systems during implementation • Improve the sharing of lessons learned

Table B.3 **Conservation International (CI)**

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Being a GEF Agency is helping CI develop a self-evaluation system, not only for GEF projects, but possibly for all projects
Areas for improvement	• More use of self-evaluation for learning between its independent evaluation unit, project staff, stakeholders, and beneficiaries • More peer learning from other Agencies with more advanced self-evaluation systems

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with GEF as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by GEF unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Continuous reviewing of policies and guidelines to align with GEF partnership standards
Areas for improvement	• Need for more developed guidelines, quality assurance processes, information management and learning practices

Table B.5 **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 terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with GEF as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by GEF unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Continuous improvement through Development Results Working Group
Areas for improvement	• More alignment and cooperation between independent evaluation unit and GEF coordination unit

Table B.6 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 terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with GEF as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by GEF unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Prevalidation 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

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with GEF as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by GEF unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Inclusion of reference groups with preparation and dissemination of terminal evaluations
Areas for improvement	• Better aggregation of learnings from terminal evaluations to feed into high-level decision making

Table B.8 Foreign Economic Cooperation Office, Ministry of Environmental Protection of China (FECO)

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• 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 independent evaluation unit, project staff, stakeholders, and beneficiaries

Table B.9 **Inter-American Development Bank (IDB)**

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with GEF as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by GEF unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Good quality assurance mechanisms • Useful system for accountability and generating information for learning • Office of Strategic Planning and Development Effectiveness that supports project teams with monitoring and evaluation throughout the project cycle • IDB frequently conducts impact evaluations for projects
Areas for improvement	• More use of self-evaluation for learning between independent evaluation unit, project staff, stakeholders, and beneficiaries

Table B.10 **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 terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with GEF as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by GEF unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Advanced data capture as part of self-evaluation system, with a range of indicators for rural poverty on all its projects • A Development Effectiveness Unit dedicated to applying self-evaluation results in informing project design
Areas for improvement	• More use of self-evaluation for learning between independent evaluation unit, project staff, stakeholders, and beneficiaries

Table B.11 International Union for Conservation of Nature (IUCN)

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Only GEF Agency nongovernmental organization 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

Table B.12 **United Nations Development Programme (UNDP)**

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Excellent collaborative project implementation report reporting, which adds value to collaboration, evidence, and oversight
Areas for improvement	• Lack of validation by independent evaluation unit before GEF submission may compromise quality and consistency of reports • Lack of involvement of independent evaluation unit in managing self-evaluation also influences impartiality of self-evaluation • More capacity building and investment in the use of theory of change for understanding project's causal pathways • More use of self-evaluation for learning between independent evaluation unit, project staff, stakeholders, and beneficiaries

Table B.13 United Nations Environment Programme (UNEP)

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Very robust guidance for evaluation • Very robust peer reviews for evaluation
Areas for improvement	• More use of self-evaluation for learning between independent evaluation unit, project staff, stakeholders, and beneficiaries

GEF Results-Based Management: Agency Self-Evaluation and the GEF Portal

Table B.14 **United Nations Industrial Development Organization (UNIDO)**

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Meta reviews comparing and synthesizing lessons from terminal evaluations • Debriefs on project evaluations with project managers, country offices, and evaluators
Areas for improvement	• Better feedback loops between project delays and optimal timing of midterm reviews for learning and compliance

Table B.15 World Bank Group (WBG)

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• 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

Table B.16 World Wildlife Fund–US (WWF-US)

Purpose of self-evaluation	• Accountability • Learning • Management support • Capacity building
Relationship with independent evaluation	• Independent evaluation unit • Semi-independent evaluation unit • Manages terminal evaluations • Validates terminal evaluations
Specific guidance provided on	• Evaluation criteria and processes • Quality assurance including role of independent evaluation office • Preparation of midterm review • How to comply with gef as step-by-step guide • Role of monitoring in evaluation • Role of GEF core indicators in evaluation • Role of theory of change in evaluation

Arrangements ensuring	• Credibility and quality
	  • Quality assurance by gef unit or program managers on project implementation report and midterm review
	  • Internal peer review on terminal evaluation
	  • External consultation on terminal evaluation
	  • Field-based/technical reviews for a sample of terminal evaluations
	  • Updating theory of change at regular intervals
	  • Use of theory of change 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	• Long tradition of using a results chain framework as methodology for theory of change
Areas for improvement	• More use of self-evaluation for learning between independent evaluation unit, project staff, stakeholders, and beneficiaries • More peer learning from other Agencies with more advanced self-evaluation systems

Summary tables of Agency self-evaluation system characteristics

The tables in this annex use the following abbreviations:

- ADB = Asian Development Bank
- AfDB = African Development Bank
- BOAD = West African Development Bank
- CAF = Development Bank of Latin America (Corporación Andina de Fomento)
- CI = Conservation International
- DBSA = Development Bank of Southern Africa
- EBRD = European Bank for Reconstruction and Development
- FAO = Food and Agriculture Organization of the United Nations
- FECO = Foreign Economic Cooperation Office, Ministry of Environmental Protection of China
- FUNBIO = Brazilian Biodiversity Fund (Fundo Brasileiro para a Biodiversidade)

- IDB = Inter-American Development Bank
- IFAD = International Fund for Agricultural Development
- IUCN = International Union for Conservation of Nature
- UNDP = United Nations Development Programme
- UNEP = United Nations Environment Programme
- UNIDO = United Nations Industrial Development Organization
- WBG = World Bank Group
- WWF-US = World Wildlife Fund–US.

Table C.1 **Summary of monitoring and evaluation guidance provided by GEF Agencies**

Agency	Evaluation criteria and processes	Quality assurance including role of IEU	Guidance for midterm review	Guidance on complying with GEF IEO	Role of monitoring in evaluation	Role of GEF IEO core indicators in evaluation	Role of theory of change in evaluation
ADB	Yes	Yes	—	Yes[a]	Yes	—	—
AfDB	Yes	Yes	—	—	Yes	—	—
BOAD	—	—	—	—	—	—	—
CAF	Yes[b]	Yes	—	Yes[c]	Yes	Yes[d]	—
DBSA	—	—	—	—	—	—	—
EBRD	Yes	Yes	—	—	Yes	—	—
IDB	Yes	Yes	—	Yes[a]	Yes	—	—
WBG	Yes	Yes	—	—	Yes	—	Yes
FAO	Yes	Yes	Yes	Yes[a, e]	Yes	Yes[f]	Yes
IFAD	Yes	Yes	—	—	Yes[g]	—	Yes
UNDP	Yes	Yes	Yes	Yes[e]	Yes	Yes	—
UNEP	Yes	Yes	Yes	Yes[e]	—	—	Yes
UNIDO	Yes	Yes	Yes	Yes[a, e]	Yes	—	Yes
CI	Yes	Yes	Yes	Yes[e]	—	—	—
FECO	Yes	Yes	Yes	Yes[a]	Yes	Yes[f]	—
FUNBIO	—	—	—	—	—	—	—
IUCN	Yes	Yes		Yes[e]	Yes	—	—
WWF-US	Yes	Yes	Yes	Yes[e]	Yes	—	—

Note: — = no data provided by Agency or not applicable; IEU = independent evaluation unit. ▢ = Development finance bank, ▢ = United Nations entity, ▢ = nongovernmental organization.

a. Step-by-step guide (explicit for GEF IEO).

b. Processes are outlined, but evaluation criteria for independent evaluations are not fully in line with the standards of the Organisation for Economic Co-operation and Development's Development Assistance Committee (OECD DAC, four of whose criteria are used by the GEF IEO).

c. Only some process details for project implementation report (PIR) as part of project cycle guidance; none for terminal evaluation and midterm review, as they are managed under independent evaluation function and/or external evaluators.

d. CAF references a table of core indicators and a methodology for the selection of subindicators.

e. Some GEF IEO requirements internalized within Agency guidelines for terminal evaluation, midterm review, and PIR.

f. Some basic guidance referring to core indicators and where to find GEF IEO guidance on this.

g. 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).

Table C.2 Summary of quality assurance procedures used by GEF Agencies

Agency	Project/program management		Self-evaluation system									
	Quality control by GEF unit or Agency program/portfolio managers	Update ToC at regular intervals	Validation by independent/semi-independent unit			Internal peer review process			External consultation on TE	Use ToC for evaluation/learning	Field-based/technical reviews for sample	Impact evaluations
			TE	MTR	PIR	TE	MTR	PIR				
ADB	Yes	—	Yes	—	—	Yes[a]	—	—	—	Yes	Yes	Yes
AfDB	Yes	—	Yes	—	—	Yes[b]	—	—	—	Yes	—	Yes
BOAD	—	—	—	—	—	—	—	—	—	—	—	—
CAF	Yes	—	—	—	—	—	—	—	—	—	—	—
DBSA	—	—	Yes	—	—	—	—	—	—	—	—	—
EBRD	Yes	—	Yes	—	—	Yes[c]	—	—	—	Yes	—	—
IDB	Yes	—	Yes	—	—	Yes[a,d]	—	—	—	No	—	Yes
WBG	Yes	Yes	Yes	—	—	Yes[a,b]	—	—	—	Yes	Yes	—
FAO	Yes	Yes	Yes	Yes[e]	No	Yes	Yes[f]	No	Yes[g]	Yes	—	Yes
IFAD	—	—	Yes	—	—	Yes	—	—	—	—	—	Yes
UNDP	Yes	Yes	No[h]	No	—	Yes	Yes	Yes	Yes[i]	—	—	—
UNEP	Yes	—	Yes	—	—	Yes[a]	—	—	Yes[g]	Yes	—	Yes
UNIDO	Yes	—	Yes	—	—	Yes	—	—	Yes	Yes	—	Yes
CI	Yes	—	—	—	—	—	—	—	—	—	—	—
FECO	Yes	—	Yes	—	—	—	—	—	Yes[g]	—	—	—
FUNBIO	—	—	—	—	—	—	—	—	—	—	—	—
IUCN	Yes	—	—	—	—	—	—	—	—	Yes	—	—
WWF-US	Yes	Yes	Yes	Yes	Yes	—	—	—	—	Yes	—	—

Note: — = no data provided by Agency or not applicable; ToC = theory of change; TE = terminal evaluation; MTR = midterm review; PIR = project implementation report. ▢ = Development finance bank, ▢ = United Nations entity, ▢ = nongovernmental organization.

a. Peer reviews within independent evaluation unit during validation of TEs.

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

c. Prevalidation of reports through EBRD's independent evaluation unit before submission.

d. Office of Strategic Planning and Development Effectiveness.

e. Only upon request by the GEF Coordination Unit.

f. Peer review is only highly recommended but is under budget holder's responsibility; independent evaluation office provides feedback.

g. Optional consultations with external stakeholders such as operational focal points, national counterparts, and civil society organizations, e.g. on preliminary findings.

h. UNDP independent evaluation unit only validates completed and cleared TEs retrospectively. Its independent review processes as well as coordination of the TE process are achieved through monitoring and evaluation specialists at its commissioning units.

i. Discussions on preliminary findings.

Table C.3 **Summary of GEF Agency information and knowledge management and learning practices on self-evaluation**

Agency	Practices in place by independent evaluation office		Learning practices in place for learning from self-evaluation
	Aggregation and synthesis	Database	
ADB	Yes	Yes	Yes
AfDB	Yes	Yes	Yes
BOAD	—	—	—
CAF	—	Yes	Yes
DBSA	No	—	No[a]
EBRD	Yes	Yes	Yes
IDB	Yes	Yes	Yes
WBG	Yes	Yes	Yes
FAO	Yes	Yes	Yes
IFAD	Yes	—	Yes
UNDP	Yes	Yes	Yes
UNEP	Yes	—	—
UNIDO	Yes	—	Yes
CI	—	—	—
FECO	—	Yes	Yes
FUNBIO	—	—	—
IUCN	Yes	—	—
WWF-US	—	—	Yes

Note: — = no data provided by Agency or not applicable. ☐ = Development finance bank, ☐ = United Nations entity, ☐ = nongovernmental organization.

a. None apart from submitting evaluation reports to the DBSA Board, Strategy Department, and Communication Department.

Table C.4 Summary of GEF Agency terminal evaluation process by rating criteria and alignment with GEF IEO

Agency	Effectiveness		Efficiency		Relevance		Sustainability		Overall outcome rating in line with GEF IEO
	Assessed and rated	Aligned with GEF IEO rating scale	Assessed and rated	Aligned with GEF IEO rating scale	Assessed and rated	Aligned with GEF IEO rating scale	Assessed and rated	Aligned with GEF IEO rating scale	
ADB	Yes	No	Yes	No	Yes	No	Yes	No	No[a,b]
AfDB	—	—	—	—	—	—	—	—	—
BOAD	—	—	—	—	—	—	—	—	—
CAF	—	—	—	—	—	—	—	—	—
DBSA	—	—	—	—	—	—	—	—	—
EBRD	No[c]	Yes	No[c]	Yes	No	—	Yes	No[d]	No
IDB	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WBG	Yes[e]	No	Yes	No	Yes	No	No[f]	—	Yes
FAO	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IFAD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No[a,b]
UNDP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UNEP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No[d]	No[a]
UNIDO	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FECO	—	—	—	—	—	—	—	—	No
FUNBIO	—	—	—	—	—	—	—	—	No
IUCN	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No[g]
WWF-US	No[h]	—	No[h]	—	No[h]	—	Yes	Yes	No[h]

Note: — = no data provided by Agency or not applicable. ☐ = Development finance bank, ☐ = United Nations entity, ☐ = nongovernmental organization.

a. Overall outcome rating includes more criteria than effectiveness, efficiency, and relevance.

b. Outcome rating is not in line with the GEF IEO's guidance on rating scale (provided in terminal evaluation guidelines).

c. Rating of subcomponents/components but no indication of an overall rating for the criterion.

d. Uses GEF IEO six-point rating scale for sustainability.

e. Assesses and rates efficacy instead of effectiveness.

f. Assesses and rates risk to development outcome instead of sustainability.

g. Does not indicate an overall outcome rating

h. Effectiveness, efficiency, and relevance are not rated separately.

Interviewees

GEF SECRETARIAT

Ulrich Apel, Senior Environmental Specialist

Cyril Blet, Senior Specialist Results-Based Management

William Ernest Ehlers, Senior Country Officer

Claude Gascon, Manager, GEF Programs

Fareeha Iqbal, Senior Climate Change Specialist

Leah Bunce Karrer, Senior Environmental Specialist

Peter Lallas, Advisor

Omid Parhizkar, Operations Officer

Naying Peng, Operations Analyst

Quynh Xuan Thi Phan, Senior Financial Officer

Henry Salazar, Senior Operations Officer

Christian Holde Severin, Senior Environmental Specialist

Jean Marc Sinnassamy, Senior Environmental Specialist

Sarah Wyatt, Biodiversity Specialist

Mark Zimsky, Senior Biodiversity Specialist

GEF SCIENTIFIC AND TECHNICAL ADVISORY PANEL

Guadalupe Duron, Program Officer

GEF IEO

Anupam Anand, Evaluation Officer

Sara Choufi, Evaluation Analyst

Molly Watts Sohn, Evaluation Analyst

Peixuan Zhou, Evaluation Analyst

GEF AGENCIES

Arun Abraham, Consultant, Asian Development Bank

Ayanleh Daher Aden, GEF Coordinator, African Development Bank

Jenin Assaf, Evaluation Officer, Office of Evaluation, Food and Agriculture Organization of the United Nations

Rachel Bedouin, Team Lead, Office of Evaluation, Food and Agriculture Organization of the United Nations

Pascale Bonzom, Good Growth Partnership Global Project Manager, United Nations Development Programme

Lafeta Camera, Evaluation Officer, African Development Bank

Leonardo Corral, Principal Specialist, Inter-American Development Bank

Rana Ghoneim, Division Chief, UNIDO Energy Department, United Nations Industrial Development Organization

Olga Gordiievska, GEF Coordination Office, United Nations Industrial Development Organization

Andrew Hudson, Head, Water & Ocean Governance Programme, United Nations Development Programme

Stephen Hutton, Senior Evaluation Officer, Independent Evaluation Group, World Bank

Priyanka Chandrahas Kamat, Information Technology Officer, World Bank

Garrett Kilroy. Principal Evaluation Specialist, Asian Development Bank

Amelia Kissick, Results-Based Management Specialist, World Wildlife Fund–US

Barry Kodolkin, Director, Deputy Chief Evaluator, European Bank for Reconstruction and Development

Luzmila Lambrano, Program Associate, United Nations Development Programme

Liu Lei, Project Officer, Foreign Economic Cooperation Office, Ministry of Environmental Protection of China

Hedi Manai, Division Manager, African Development Bank

Lavinia Monforte, Evaluation Manager, Office of Evaluation, Food and Agriculture Organization of the United Nations

Rosario Narciso, GEF Portfolio Management Officer (Consultant), Asian Development Bank

Guirane NDiaye, GEF Portfolio Supervision and Results Specialist, African Development Bank

Jaturong Padungsapya, Global Environmental Finance, United Nations Development Programme

René Gómez-García Palao, Senior Executive and Head of Green Business Unit, Development Bank of Latin America

Viven Prasad Sade, Senior Information Technology Officer, World Bank

Orissa Samaro, Senior Director, Conservation International

Estefania Samper, Results-Based Management Specialist, United Nations Development Programme

Michael Spilsbury, Director, Evaluation Office, United Nations Environment Programme

Jossy Thomas, Industrial Development Officer, United Nations Industrial Development Organization

Stefanie Valcheva, GEF Coordination Team, United Nations Industrial Development Organization

GREEN CLIMATE FUND

Nebi Bekiri, Head of Enterprise ICT Programme Management Office

Johann Elysee, Senior Portfolio Specialist

Lilian Macharia, Head of Office of Portfolio Management

GEF Portal performance

Table E.1 GEF Portal dimensions evaluated by operational activity

Dimension	Design	Management	Utilization
Portal navigation/ organization	Arrangement/sequence of pages	• Inclusion of all necessary pages • Functionality of links	• Ease of navigation • Ability to understand page layout and ease of search (by main portal categories)
Portal information	Effective visualization (maps, lists, summaries)	• Content analysis • Completeness of information	• Ability to find information needed • Ability to understand information provided, including balance between detail and visualization
Tracking	Efficient internal display of project timeline and results; deadline alerts	Efficient internal display of project timeline and results; deadline alerts	Efficient external display of project timeline and results; deadline alerts
Document systems		• Ease of accessing and processing submissions • Efficient workflow	• Ease of upload • Efficient workflow
Interactivity		• Contact provided • Correct channeling of queries	• Ability to find contact • Question answered quickly/ accurately
Alignment with overall mission	Efficient use of internal users' time	Accomplishment of portal mission to the GEF	• User satisfaction • Main positive/negative features • Main missing features • Efficient use of external users' time • Comparison with similar portals
Users		User growth rate	Likelihood to influence decision to engage with the GEF in future

Best practice	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 icons match the functions; others are misleading or unclear	Topic cluster used, a good resource for this portal	Clearly understandable page layout and topic classification; within each page are well-defined topic clusters	Website is divided into two well-differentiated pages; within each are well-defined topic clusters
Clear location on sitemap through use of "breadcrumbs"	No	Yes, limited to two layers	Yes	No, nor are "go back" buttons offered
Ease of understanding site components/ features	• Menu icons are used, which is a good resource • Buttons with different functions have the same color, screen location, shape, and size; a better way to visually indicate intent should be used	Lot of information, which makes using the sitemap necessary in order to understand the portal; sitemap is clearly presented, however	Images and icons help guide the user, giving hints of what the related feature is meant to do	Easy to understand where each button leads; no icons are used, which would likely be a better, less text-heavy way to guide users
Ease of finding the information needed	• Filter menu could benefit from hierarchical organization • An "x" button allows filter selections to be reversed; there may be other ways to help users navigate filtering as well • Icons leading to pop-ups are not intuitive	• Project search offers a simple, useful filter tool • Table presenting search results lacks dynamism, but is still useful • There is a quick search option for entire portal	Clean, very intuitive, page layout, which makes finding sections easy; however, search function does not work	• Filter menu is well distributed and offers a useful way to filter different aspects by enabling multiple selection via checkboxes; this also provides more flexibility in deselecting filtering criteria • Filter menu could benefit from a "submit" or "go" button, as every time a checkbox is selected/ deselected, both the filter menu and the results dashboard reload, which could be undesirable

(continued)

Table E.2 Comparison of the GEF Portal with peer portals by dimension and practice *(continued)*

Best practice	GEF Portal	UNFCCC CDM	IRENA Project Navigator	Green Climate Fund project portfolio
Avoid "analysis paralysis" by reducing number of choices	Number of menus offered is minimal	There is too much information available (mainly as text), and it is not adequately prioritized	Number of choices is limited on main pages; once inside subpages, the number of choices increases but is still user friendly, since content scope is framed within a topic	Website is divided into two well-differentiated pages Gallery presentation of projects can be visually overwhelming at first, because user attention is drawn to it first rather than to the filter menu
Visual appeal				
Consistent branding	Yes	Yes	Yes	Yes
Appealing color scheme/useful color contrast	• Green and blue are highly saturated, possibly competing for visitor attention • Clean title bar, but logo and green text are subsequently lost from the navigation; there may be ways to better leverage color and design to guide users back to top level • Buttons with different functions have same color; this should be changed to better guide the user	• Neutral aesthetic lacking dynamism, which makes it difficult to find topics of interest • Color contrast should be optimized	Simple, clean design, featuring brand color scheme, contrast between fonts, negative space, and homogeneity between elements of the same page	• Clean design, which adheres to brand color scheme; homogeneity between elements of the same page • Color contrast is well used for parts with white background; for parts with images used in background, contrast with text should be optimized
Content easy to read by being concise, minimizing text, and/or using short paragraphs	Minimal text	Text divided in short paragraphs; however, there is too much information	Yes	Yes
Content easy to read through proper text layout	Pop-ups with some distracting text wrap	Some links are of disproportionate size, being relatively larger for no reason	Yes, text size helps guide and focus the user	Yes, but in some places, interactive text pop-ups and counter simulators may be undesirable
Content easy to read through block reading	Not applicable, as the amount of text presented is minimal	Yes	Yes, portal makes use of text boxes combined within interactive diagrams enabling block reading	Block reading enabled, but at some points, block boundaries are not clearly distinguishable

(continued)

Best practice	GEF Portal	UNFCCC CDM	IRENA Project Navigator	Green Climate Fund project portfolio
Content easy to read through use of bullet points	Yes, alternating gray bars are used for project lists; it would be interesting to see if this style could be applied to other types of listings	Yes, bulleted list items are indicated with a "+" character	Yes	Rarely used, could be optimized
Use of negative space to avoid crowded pages	Clean design with minimal text and negative space well distributed	Used, but not optimized	Clean design with negative space well distributed	Used, but could be optimized
Use of quality images/photos and balance between words and images	No, only icons are used; more images could make the site more visually appealing	No, lack of balance between images and text; text predominates	Yes, uses good-quality diagrams and images to assist in understanding ideas and information presented	Yes, uses appropriate diagrams and high-quality images, which contribute to a better user experience
Use of human face aimed at "putting a face to the name"	No	Yes, but images are small	No	Yes
Accessibility				
Accessibility for the public	• Account needed: only entities recognized as legitimate through World Bank Group can access • Pertinent result from web search leads to a blank page; this can be corrected by deleting "/App" at end of URL; this should be fixed	Account needed; anyone can create an account	Account needed; anyone can create an account	Open to the public; this is not a portal to log into, but only a website
Functional links	• 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 font types in a small space	Yes	Yes

(continued)

Best practice	GEF Portal	UNFCCC CDM	IRENA Project Navigator	Green Climate Fund project portfolio
Optimization for mobile	Unable to assess	No; mobile version is same as desktop version	No; homepage seems to be optimized for mobile, but portal itself is not	Yes; visitors are able to easily navigate website on mobile devices
Optimization for speed[a]	• Load time: 982 ms • Overall performance grade A (99/100)	• Load time: 645 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)
Prioritization of search engine optimization (the lower the optimization potential, the better it is)[b]	90% optimization potential • GEF Portal URL redirects to Microsoft Online Login page for the World Bank • 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[c]	Medium; server does not support Forward Secrecy with the reference browsers	Strong	Medium; server accepts RC4 cipher, but only with older protocols	Strong
Interactivity				
Strong calls to action	Unable to assess	• Tab for contact with FAQs for different topics, with questions addressed through other websites or complex documents • "Contact us" section is well organized, and institutional contact emails are categorized by topic or issue addressed	Not applicable; portal not conceived as a platform for interactivity with the institution	Each project fact sheet has a contacts section, which provides contact details of officers from both the GCF and corresponding accredited entity
Email marketing enabled	Unable to assess	No	Yes	Yes
Social media enabled	Unable to assess	Yes, both at homepage before logging in as well as once logged into the portal	Yes, at the homepage before logging in	Yes
Chatbots used	Unable to assess	No	No	No

Note: ms = millisecond; s = second. UNFCCC CDM = United Nations Framework Convention on Climate Change Clean Development Mechanism. Performance indicators selected are from Morgeson (2011). Data are based on observations made as an external user.

a. Website speed test performed using the Pingdom Website Speed Test.

b. Website search engine optimization check performed using the SEO Checker from IONOS.

c. Website security check performed using the SSL Server Test from Qualys SSL Labs.

Table E.3 User satisfaction with GEF Portal performance

Performance dimension	GEF Agencies		Secretariat		OFPs		Other		All users	
	No.	%	No.	%	No.	%	No.	%	No.	%
Ease of navigation	95	28	28	46	10	90	13	46	145	37
Ease of web page use	95	28	28	39	9	100	13	46	144	36
Clarity of web pages	95	62	28	57	10	100	13	85	145	66
Completeness of web pages	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 documents	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
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 satisfaction 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 administered November 2020.

Note: No. = number of observations (responses received); % = percentage of respondents that assessed performance on a given dimension to be in the top half of a six-point scale. The specific response options for the scale differed based on the dimension, but the top half broadly indicates an acceptable performance range, and the bottom an unacceptable performance range. Variation in user totals reflects overlap in respondent categories.

References

All URLs were checked before publication.

AfDB IDEV (African Development Bank Independent Development Evaluation). 2018. "Evaluation of Quality Assurance across the Project Cycle of the African Development Bank Group (2012–2017). Synthesis Report." Abidjan, African Development Bank Group.

ADB IED (Asian Development Bank Independent Evaluation Department). 2020. "Annual Evaluation Review 2020: ADB's Project Level Self-Evaluation System." Asian Development Bank.

Arrow, Kenneth J. 1984. "The Economics of Agency." No. TR-451. Stanford University California Institute for Mathematical Studies in the Social Sciences.

Bester, Angela. 2016. "Results-Based Management in the United Nations Development System." A report prepared for the United Nations Department of Economic and Social Affairs for the 2016 Quadrennial Comprehensive Policy Review.

Christoffersen, Leif, et al. 2002. "The First Decade of the GEF: Second Overall Performance Study." GEF, Washington, DC.

ECG (Evaluation Cooperation Group). 2018. "Self-Evaluation in ECG Member Institutions." ECG Practice Note 3.

FAO (Food and Agriculture Organization of the United Nations). 2015. "OED Evaluation Manual."

GEF (Global Environment Facility). 2011. "Independent Review of GEF Systems." GEF/C.41/Inf.12. GEF, Washington, DC.

———. 2012. "Options for Strengthening GEF Systems: Addressing the Findings and Recommendations of the Independent Review of GEF Systems." GEF/42/Inf.12. GEF, Washington, DC.

———. 2017. "GEF Business Plan and Corporate Budget for FY18." GEF/C.52/06. GEF, Washington, DC.

———. 2018. "Progress Report on the Development of the New GEF Portal." GEF/C.54/Inf.06. GEF, Washington, DC.

———. 2019a. "Minimum Fiduciary Standards for GEF Partner Agencies." GA/PL/02. GEF, Washington, DC.

———. 2019b. "Policy on Monitoring." ME/PL/03. GEF, Washington, DC.

GEF IEO (Global Environment Facility Independent Evaluation Office, formerly Office of Monitoring and Evaluation; Evaluation Office). 2005. *OPS3: Progressing toward Environmental Results. Third Overall Performance Study of the GEF*. Washington, DC: Office of Monitoring and Evaluation.

———. 2010a. "The GEF Monitoring and Evaluation Policy 2010." GEF Evaluation Office, Washington, DC.

———. 2010b. *OPS4: Progress toward Impact. Fourth Overall Performance Study of the GEF*. Washington, DC: GEF Evaluation Office.

———. 2014. *GEF Annual Performance Report 2013*. Evaluation Report No. 92. Washington, DC: GEF IEO.

———. 2017a. "Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects." GEF IEO, Washington, DC.

———. 2017b. *Review of Results-Based Management in the GEF*. Evaluation Report No. 112. Washington, DC: GEF IEO.

———. 2018. *OPS: The GEF in the Changing Environmental Landscape. Sixth Comprehensive Evaluation of the GEF*. Evaluation Report No. 110. Washington, DC: GEF IEO.

———. 2019. "The GEF Evaluation Policy." GEF IEO, Washington, DC.

———. 2020a. "Evaluation of the Agency Self Evaluation Systems: Revised Approach Paper." GEF IEO, Washington, DC.

———. 2020b. "Report on the Review of the GEF Terminal Evaluation Validation Process." GEF IEO, Washington, DC.

———. 2022a. *Evaluation of Knowledge Management in the GEF*. Evaluation Report No. 152. Washington, DC: GEF IEO.

———. 2022b. *GEF Annual Performance Report 2020: Quality of Reporting*. Evaluation Report No. 143. Washington, DC: GEF IEO.

———. 2023. *GEF Annual Performance Report 2021*. Evaluation Report No. 155. Washington, DC: GEF IEO.

GEF STAP (Global Environment Facility Scientific and Technical Advisory Panel). 2020. "Theory of Change Primer: A STAP Advisory Document." GEF STAP, Washington, DC.

Grossman, Sanford J., and Oliver D. Hart. 1992. "An Analysis of the Principal-Agent Problem." In *Foundations of Insurance Economics: Readings in Economics and Finance*, edited by Georges Dionne and Scott E. Harrington, 302–40. Dordrecht: Springer.

IDB (Inter-American Development Bank). 2008. "Development Effectiveness Framework."

IEG (Independent Evaluation Office). 2014. "Learning and Results in World Bank Operations: How the Bank Learns." World Bank, Washington, DC.

———. 2016. *Behind the Mirror: A Report on the Self-Evaluation Systems of the World Bank Group*. Washington, DC: World Bank.

Kirk, Colin. 2019. "Independent External Evaluation of EBRD's Evaluation System. Final Report: Main Report." European Bank for Reconstruction and Development.

Kusek, Jody Zall, and Ray Rist. 2004. "Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners." World Bank, Washington, DC.

MoFA (Ministry of Foreign Affairs) Denmark. 2010. "Peer Review of the Evaluation Function of UNIDO." Evaluation Department, MoFA.

Morgeson, Forrest V. 2011. "E-Government Performance Measurement: A Citizen-Centric Approach in Theory and Practice." In *Electronic Governance and Cross-Boundary Collaboration: Innovations and Advancing Tools*, edited by Yu-Che Chen and Pin-Yu Chu, 150–65. Hershey, PA: IGI Global.

Nonaka, Ikujiro. 1994. "A Dynamic Theory of Organizational Knowledge Creation." *Organization Science* 5 (1): 14–37.

OECD DAC (Organisation for Economic Co-operation and Development Development Assistance Committee). 2002. "Glossary of Key Terms in Evaluation and Results Based Management."

Patton, M.Q. 2010. "Future Trends in Evaluation: From Policies to Results." Developing Capacities for Country Monitoring and Evaluation Systems webinar.

Picciotto, Robert. 1999. "Towards an Economics of Evaluation." *Evaluation* 5 (1): 7–22.

Ross, Stephen A. 1973. "The Economic Theory of Agency: The Principal's Problem." *The American Economic Review* 63 (2): 134–39.

Scriven, Michael. 1975. "Evaluation Bias and Its Control." Evaluation Center, Western Michigan University, Kalamazoo.

Spender, J.-C. 1996. "Making Knowledge the Basis of a Dynamic Theory of the Firm." *Strategic Management Journal* 17 (S2): 45–62.

Taut, Sandy. 2007. "Studying Self-Evaluation Capacity Building in a Large International Development Organization." *American Journal of Evaluation* 28 (1): 45–59.

UN (United Nations). 2018. "Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation." Secretary-General's Bulletin. ST/SBG/2018/3.

UNEG (United Nations Evaluation Group). 2012a. "Professional Peer Review of the Evaluation Function of United Nations Environment Programme."

———. 2012b. "Peer Review: The Evaluation Function of Food and Agriculture Organization (FAO)." Global Environment Facility Evaluation Office, Washington, DC.

———. 2017. "Norms and Standards for Evaluation." UNEG, New York.

UNIDO (United Nations Industrial Development Organization). 2017. "Renewable Energy-Based Mini-Grids: The UNIDO Experience." Working Paper 1/2017. UNIDO, Vienna.

Yin, Robert K. 2017. *Case Study Research and Applications: Design and Methods*. 6th ed. New York: Sage Publications.

The Independent Evaluation Office (IEO) of the Global Environment Facility (GEF) was established by the GEF Council in July 2003. The Office is independent from GEF policy making and its delivery and management of assistance.

The Office undertakes independent evaluations at the strategic level. These evaluations typically focus on cross-cutting themes, such as focal area–wide topics or integrated approaches to delivering global environmental benefits. The IEO presents a GEF-wide annual performance report and also undertakes institutional evaluations, such as assessing GEF governance, policies, and strategies. The Office's work culminates in a quadrennial comprehensive evaluation of the GEF.

The Office cooperates with professional evaluation networks on developing evaluation approaches, setting standards, and delivering training—particularly with regard to environmental evaluation and evaluation at the interface of environment and socioeconomic development. We also collaborate with the broader global environmental community to ensure that we stay on the cutting edge of emerging and innovative methodologies.

To date, the Office has produced over 150 evaluation reports; explore these on our website: www.gefieo.org/evaluations.

Independent Evaluation Office, Global Environment Facility
1818 H Street, NW Washington, DC 20433, USA
www.gefieo.org   X @gefieo   ▶ @gefieo