

58th GEF Council
June 02 – 03, 2020
Virtual Meeting

ANNUAL PERFORMANCE REPORT 2020

(Prepared by the Independent Evaluation Office of the GEF)

TABLE OF CONTENTS

List of Abbreviations	v
Executive Summary.....	vi
Overview	vi
Findings	vi
Performance of Completed Projects.....	vi
Management Action Record	vii
Special Thematic Focus: Quality of Reporting	viii
Chapter 1. Overview	1
Chapter 2. Performance of completed projects	4
Methodology.....	4
Findings	5
Outcome and Sustainability	5
Implementation and Execution.....	7
Project Monitoring.....	8
Cofinancing	9
Chapter 3. The Management Action Record 2020	10
Rating Approach.....	11
The GEF IEO Keeps Track of the Reasons for Removing a Decision from the MAR.....	12
Findings	12
GEF Council Decisions with a Substantial Level of Adoption	12
Decision with a Medium Level of Adoption	14
Retired Decisions	14
Chapter 4. Quality of Reporting.....	16
Key Questions	16
Methodology.....	17
Limitations	21
Findings	21
General Information	27
Stakeholder Feedback.....	27
Theory of Change	28

Sources of Information	29
Reporting on Outcomes and Sustainability	29
Monitoring, Implementation, and Cofinancing	31
Social and Environmental Safeguards.....	32
Lessons and Recommendations.....	33
Performance Ratings.....	34
Candor in Reporting through PIRs	35
References	40
Annex A: APR 2020 Project Cohort	41
Annex B: Pilot Phase and GEF-1 Projects added to GEF IEO TER Database.....	48
Annex C: Terminal Evaluation Report Review Guidelines	49
Annex D: Instrument for review of TE quality	56
Annex E: Instrument for Review of Canceled Projects	68
Annex D: Results of Review of Terminal Evaluation Quality.....	73

TABLES AND FIGURES

Table 1: Distribution of completed GEF projects and funding by ratings.....	5
Table 2: Materialization of cofinancing	9
Table 3: Unsatisfactory or lower outcome rating—by GEF Agency.....	20
Table 4: Percentage of terminal evaluation reports rated in the satisfactory range	24
Table 5: Compliance on reporting requirements for terminal evaluations of full-size projects	25
Table 6: Percentage of terminal evaluations for full-size projects that report how stakeholders were involved in conduct of evaluations	28
Table 7: Percentage of terminal evaluations for full-size projects that address theory of change.....	28
Table 8: Percentage of terminal evaluations that discuss information sources—for full-size projects	29
Table 9: Percentage of terminal evaluations that report on given aspect of outcome performance.....	30
Table 10: Reporting on M&E, implementation, and cofinancing—for full-size projects.....	32
Table 11: Percentage of terminal evaluations that report on social and environmental safeguards.....	33
Table 12: Percentage that present lessons and recommendations—for full-size projects	33
Table 13: Percentage of terminal evaluations that provide performance ratings and their consistency ..	34
Table 14: Percentage of projects in the satisfactory range for outcome (or equivalent)- by rating source	35

Figure 1: Projects with outcomes rated in the satisfactory range by GEF replenishment period.....	6
Figure 2: Projects with sustainability of outcomes rated in the likely range by replenishment period.....	6
Figure 3: Projects with implementation/execution quality rated in the satisfactory range by GEF replenishment period	7
Figure 4: Projects with M&E design/implementation rated in the satisfactory range by GEF replenishment period	8
Figure 5: Three-year rolling average of terminal evaluation quality ratings presented—based on year of terminal evaluation completion (N=1,677)	22
Figure 6: Three-year rolling average on terminal evaluation quality dimensions—by year of terminal evaluation completion	23
Figure 7: Percentage of terminal evaluations that comply with GEF IEO guidance—by project size	26
Figure 8: Terminal evaluations for full-size projects that provide information on basic project characteristics	27
Figure 9: Incidence of discussion of concerns that affect project performance—by information source (n=41 projects)	36

LIST OF ABBREVIATIONS

ADB	Asian Development Bank
APR	annual performance report
FAO	Food and Agriculture Organization of the United Nations
GEF	Global Environment Facility
IEO	Independent Evaluation Office
IEG	Independent Evaluation Group
IFAD	International Fund for Agricultural Development
LDCF	Least Developed Countries Fund
MAR	management action record
M&E	monitoring and evaluation
OFP	operational focal points
PIR	project implementation report
PMIS	Project Management Information System
SCCF	Special Climate Change Fund
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UNIDO	United Nations Industrial Development Organization

EXECUTIVE SUMMARY

Overview

1 The Independent Evaluation Office (IEO) of the Global Environment Facility (GEF) prepares the annual performance report (APR) to report on the performance of completed projects, the efficiency of GEF processes, and the responsiveness of GEF management to the recommendations of previous evaluations. APR 2020 presents an evaluation of the outcomes, sustainability, quality of implementation, and monitoring and evaluation (M&E) of 1,706 completed GEF projects. It also includes a summary of the management action record (MAR). The thematic focus of this APR is on the quality of reporting on GEF-supported activities by the Agencies.

2 The analysis in this report is mainly based on data from 1,706 completed GEF projects that account for \$7.5 billion in GEF grants and \$36.2 billion in promised cofinancing. The terminal evaluations are validated by the GEF IEO or the respective Agency evaluation offices. This dataset includes 134 completed projects, from now on referred to as the APR 2020 cohort, for which terminal evaluations were submitted during the calendar year 2019 and are being reported on in APR for the first time. The projects in the APR 2020 cohort account for \$618.8 million in GEF grants, including \$13.8 million for project preparation. The review to assess compliance with the GEF IEO's "Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects" (GEF IEO 2017) examines terminal evaluations of the APR 2020 cohort that were prepared after issuance of the guidelines in 2017. Reporting on canceled GEF projects is based on a review of the documentation for 32 canceled projects that were approved from GEF-3 onwards. The MAR 2020 is based on an assessment of adoption of the recommendations of the eight evaluations; six recommendations pertain to the GEF Council and two are related to the decisions by the Least Developed Countries Fund (LDCF) and the Special Climate Change Fund (SCCF) Council.

3 The conclusions and recommendations based on the findings of this report will be incorporated in the Seventh Comprehensive Evaluation of the GEF.

FINDINGS

Performance of Completed Projects

4 **Cumulatively, eighty percent of GEF projects have satisfactory outcome ratings.** Cumulatively, 80 percent of all the rated projects, which account for 78 percent of the GEF grant, are rated in the satisfactory range for outcome. In comparison, 83 percent of APR 2020 cohort, which account for 83 of the funding, are rated in the satisfactory range.

- (a) Seventy-one percent of the projects in the APR 2020 cohort are in the likely range for sustainability compared with 62 percent for the cumulative portfolio—this difference is statistically significant at 90 percent confidence level. This is consistent with the long-term trend of a modest improvement in sustainability ratings for the projects that have been designed from GEF-4 onwards.

- (b) Eighty-one percent of rated projects are in the satisfactory range for quality of implementation, which assesses the performance of GEF Agencies. In comparison, 85 percent of the projects of the 2020 cohort are rated in the satisfactory range. Although the ratings for the APR 2020 cohort are higher, the difference is not statistically significant.
- (c) To date, 1,448 completed projects have been rated on the quality of project execution, which assesses the performance of agencies which execute the GEF-supported activities on the ground under the supervision of the GEF Agency. Eighty-one percent have been rated in the satisfactory range. Of the 107 projects in the APR 2020 cohort that were rated for quality of project execution, 89 percent were rated in the satisfactory range.
- (d) Cumulatively, 66 percent of projects were rated as satisfactory on M&E design, compared with 71 percent of the projects in the APR 2020 cohort. Sixty-five percent of all completed projects are rated in the satisfactory range for M&E implementation; in comparison 64 percent of the projects in the APR 2020 cohort are rated satisfactory.
- (e) On average, cofinancing for the APR 2020 cohort was higher than the promised amount: \$7.98 materialized compared with \$7.58 promised per dollar of GEF grant. Within the APR 2020 cohort, for 61 percent of projects, at least 90 percent of the promised cofinancing materialized; for 26 percent materialization was less than 50 percent.

Management Action Record

5 The GEF MAR tracks the level of adoption of GEF Council and LDCF/SCCF Council decisions based on the recommendations of the evaluations conducted by the GEF IEO. MAR 2020 reports on the level of adoption for eight different evaluations, which includes two that pertain to the LDCF/SCCF Council:

- (a) Joint GEF–United Nations Development Programme (UNDP) Small Grants Programme Evaluation (GEF/ME/C.48/02)
- (b) Evaluation of the GEF Civil Society Organization Network (GEF/ME/C.50/02)
- (c) Review of the GEF Policy on Minimum Standards on Environmental and Social Safeguards (GEF/ME/C.52/Inf.08)
- (d) Review of GEF Support for Transformational Change (GEF/ME/C.52/Inf.06)
- (e) Review of GEF’s Engagement with Indigenous Peoples (GEF/ME/C.53/Inf.07)
- (f) Biodiversity Focal Area Study (GEF/ME/C.53/Inf.03)

(g) Program Evaluation of the Least Developed Countries Fund (GEF/LDCF.SCCF.20/ME/02)

(h) Program Evaluation of the Special Climate Change Fund (GEF/LDCF.SCCF.22/ME/02)

6 The GEF IEO and GEF Secretariat agree on ratings of adoption of council decisions based on all eight evaluations tracked, with adoption levels rated as substantial in seven evaluations, and as medium in one case. For the Joint GEF-UNDP Small Grants Programme Evaluation, for which adoption was rated medium on one recommendation, the Council's decision had called for reconsideration of the criteria for upgradation of the participating countries. The management reports that it has reconsidered the criteria for upgradation but is using it without any changes for the GEF-7 period. For all other evaluations, the progress has been substantial.

Special Thematic Focus: Quality of Reporting

7 The special thematic focus of the Annual Performance Report 2020 is on the quality of Agency reporting on GEF activities. The focus on quality of reporting through terminal evaluations is timely because more than three years have passed since the issuance of the GEF IEO's "Guidelines for GEF Agencies in Conducting Terminal Evaluations for Full-sized Projects" (GEF IEO 2017). The review on quality of Agency reporting also covers some aspects of reporting through project implementation reports (PIRs) and midterm reviews.

8 The overall quality of terminal evaluation reports improved substantially from 1997 to 2019 but the progress has slowed thereafter. The review analyzed data from 1,677 terminal evaluations for assessment of overall quality, and from 1,135 to 1,211 terminal evaluations to assess trends on various quality dimensions, depending on availability of ratings for each dimension.¹ The review finds that reporting on various terminal evaluation quality dimensions has improved. This is especially in the case of reporting on project financing and M&E, for which the quality of reporting has improved substantially and has converged with other dimensions such as reporting on outcomes and sustainability, in which performance was higher at the baseline. Terminal evaluations prepared by the United Nations Environment Programme (UNEP) show improvements in quality: the proportion of terminal evaluations rated in the satisfactory range for quality increased from 81 percent for those prepared through 2009, to 97 percent for those prepared from 2010 onwards. The improvement in the quality of the percentage of terminal evaluations prepared by UNDP and the World Bank that are in the satisfactory range is marginal.

9 GEF Agencies comply to a greater extent with some of the requirements of the terminal evaluation guidelines and less so with others. Based on examination of the terminal evaluations for 80 full-size and 38 medium-size projects of the APR 2020 cohort, the review found high compliance in areas such as general information on project; reporting on outcomes;

¹ This range is caused by differences in the availability of ratings on the different dimensions. For some dimensions more observations are available than others, although the number of observations remain in the same ballpark.

reporting on project M&E, and consistency in performance ratings. However, only 64 percent of terminal evaluations report on the implementation of social and environmental safeguards.

10 There are gaps in compliance with terminal evaluation guidelines across all agencies.

Terminal evaluations prepared by UNEP and United Nations Industrial Development Organization (UNIDO) comply with most requirements in general; a significant percentage of the terminal evaluations prepared by UNEP do not report on Agency performance, and those by UNIDO do not report on application of environmental and social safeguards. Among other Agencies, terminal evaluations prepared by the World Bank are strong in reporting on project M&E and on environmental and social safeguards but are weak in reporting on their information sources. Terminal evaluations prepared by UNDP are strong in reporting on sustainability and in providing ratings on all performance dimensions but are weak in reporting on project's theory of change and on the application of social and environmental safeguards.

11 Terminal evaluations of medium-size projects meet the guidelines to the same extent as full-size projects. Although the GEF IEO's terminal evaluation guidelines are applicable only to full-size projects, the evidence shows that terminal evaluations for medium-size projects meet them equally well.

12 GEF IEO validations of outcome ratings show a disconnect between project implementation report ratings, terminal evaluation ratings, and IEO ratings. This analysis was based on 596 completed projects for which ratings from project implementation reports, terminal evaluations, and validated terminal evaluations are available. A higher percentage of projects are rated in the satisfactory range for outcome by PIRs (96 percent) and terminal evaluations (92 percent) compared with the terminal evaluation validations conducted by the GEF IEO (84 percent). The disconnect between the GEF IEO validations and PIRs is 10 percent, and between GEF IEO validations and terminal evaluations is 6 percent.

13 PIRs typically flag emerging concerns and challenges in a timely manner for projects with unsatisfactory outcomes. Of the 50 projects that were approved from GEF-3 onwards and had an unsatisfactory or lower rating for outcomes, the PIRs for the vast majority note issues and risk factors that eventually jeopardized the achievement of planned outcomes. However, compared with terminal evaluations, PIRs were less likely to report concerns that affect project performance based on project design or management. Furthermore, the review identified a few instances in which the PIRs either do not adequately raise concerns that affect performance or raise it too late.

14 A high percentage of canceled projects included a discussion on risks and challenges. This analysis is based on 32 projects that were approved from GEF-3 onwards and were canceled after CEO endorsement/approval. Of these, 14 were canceled before start of implementation and 18 after project start. For projects that were canceled before the start of implementation, a cancellation memorandum was available in 12 out of 14 cases. For those that were canceled after start (18), this information is available for all—in 16 instances through PIRs, and in two instances through cancellation memo. The reasons that lead to cancellation vary from civil unrest, issues related to slow startup, delays in fund disbursement, and lack of political willingness to move forward with implementation after a change in political leadership.

For the most part, PIRs raised these challenges and concerns in a timely manner although in some instances reporting was delayed.

CHAPTER 1. OVERVIEW

1 The Independent Evaluation Office (IEO) of the Global Environment Facility (GEF) prepares the annual performance report (APR) to provide an account of the performance of completed projects, the efficiency of the GEF processes, and the responsiveness of GEF management to the recommendations of the past evaluations. APR 2020 is the 16th report of this series. It presents an assessment of the outcomes, sustainability, implementation, and monitoring and evaluation (M&E) of completed GEF projects. It also includes a summary of the management action record (MAR). The thematic focus of APR changes from year to year; APR 2020 is focused on the quality of the Agency reporting on GEF-supported activities.

2 APR 2020 will be an information document for the June 2020 GEF Council meeting. The conclusions and recommendations emerging from the findings of APR 2020 will be incorporated in the Seventh Comprehensive Evaluation of the GEF. The draft report of APR 2020 is being shared across the GEF Partnership for feedback.

3 In addition to this chapter, the report includes the following:

Section A

- **Chapter 2: Performance of completed projects.** The chapter presents an account of the outcomes and sustainability ratings of completed GEF projects. The reporting also covers trends in project implementation, execution, and M&E ratings; and aggregated data on materialization of cofinancing. The data presented are based on a review of the terminal evaluations submitted by GEF Agencies.
- **Chapter 3: Management action record.** A summary of MAR 2020 is presented in this chapter. MAR is shared publicly at the GEF IEO website.²

Section B

- **Chapter 4: Quality of reporting.** The chapter presents an analysis of the trends in the quality of terminal evaluations and the extent to which Agencies comply with the GEF IEO's "Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects" (GEF IEO 2017). It also presents an assessment of candor in reporting of the concerns related to implementation challenges and unsatisfactory performance in project implementation reports (PIRs).

4 The second chapter discusses performance of 1,706 completed GEF projects that together account for \$7.5 billion in GEF grants, including support for project preparation. During the calendar year 2019, the Agencies submitted terminal evaluations for 134 completed GEF projects. These 134 projects account for \$618.8 million in GEF grants, including \$13.8 million for project preparation. Among the GEF Agencies, the United Nations Development Programme (UNDP), the World Bank, the United Nations Environment Programme (UNEP), the United Nations Industrial Development Organization (UNIDO), and the Food and Agriculture

² <https://www.gefio.org/documents/management-action-record-mar-2020>

Organization of the United Nations (FAO), each account for at least 10 terminal evaluations of the APR 2020 cohort.³ The GEF IEO's ongoing effort to fill data gaps for projects that were completed 15 or more years back led to the retrieval of terminal evaluation reports for six projects. These projects are not included in the APR 2020 cohort, but their terminal evaluations have been reviewed and their performance ratings are included in the cumulative terminal evaluation review dataset.. A detailed list of projects included in the APR 2020 cohort is provided in annex A. The dataset on performance ratings of the completed projects will be made available at the GEF IEO website.⁴ The terminal evaluation review reports prepared by the GEF IEO will be shared with the respective GEF Agencies and made publicly available.

5 Chapter 3 of the report presents a summary of the MAR. The MAR assesses the degree to which GEF IEO recommendations endorsed by the Council have been adopted by GEF management. To assess progress on adoption, relevant actions undertaken by the management are considered. The GEF management first provides its self-assessment of the level of adoption of the tracked decisions. The GEF IEO then provides its independent assessment of the level of adoption. MAR 2020 covers eight different evaluations.

6 Reporting on overall quality of terminal evaluations presented in the fourth chapter is based on GEF IEO's terminal evaluation review dataset. Of the 1,706 projects covered in the dataset, data on overall quality of terminal evaluations is available for 1,679. Reporting on compliance with the GEF IEO's terminal evaluation guidelines is based on a review of 118 terminal evaluation reports from within the APR 2020 cohort. In addition to being part of the APR 2020 cohort, these terminal evaluations met the following conditions: they were published in or after October 2017 (six months after issuance of the GEF IEO guidelines) and were prepared in English.

7 The assessment of candor of reporting presented in the second chapter is based on assessment of whether PIRs provide information on challenges experienced during implementation in a timely manner and whether reporting accurately reflects the risks associated with the challenges. It covers projects that were approved in GEF-3 or later, and were either canceled after CEO's endorsement/approval (32 projects) or were rated unsatisfactory or lower on outcomes (50 projects). The documentation available for these projects in the GEF Project Management Information System (PMIS)/Portal was examined. The assessment of reporting on canceled projects involved a review of communications from Agencies pertaining to project progress where available, project implementation reports (PIRs), and the cancellation memorandum, to assess the extent to which these reports raised emerging challenges in a timely manner and without minimizing the potential risks. The assessment of reporting on projects that were rated unsatisfactory considers reporting through PIRs, midterm reviews, and terminal evaluations.

³ Of the two terminal evaluations submitted by AfDB, one is included in the APR 2020 cohort. The other was the excluded from the cohort after it was assessed to lack basic information that would connect the project activities to the GEF grant. The IEO is in conversations with the Agency to find a way to address these information gaps. This project will be included in APR2021 cohort.

⁴ The cumulative data used for APRs is available at: <https://www.gefio.org/data-ratings>

Section A

CHAPTER 2. PERFORMANCE OF COMPLETED PROJECTS

8 This chapter provides updates on the performance of completed projects, including ratings on project outcome, sustainability, quality of implementation and execution, and quality of project M&E. It also presents an update on trends in materialized cofinancing. These topics were analyzed in detail in the Sixth Comprehensive Evaluation of the GEF. APR 2020 gives a brief account on these topics because the underlying causal relationships are not expected to change in a short period. A detailed treatment of each of these topics will be presented in APR 2021, which will be prepared as an input to the Seventh Comprehensive Evaluation of the GEF. Performance of the projects in the APR 2020 cohort on various dimensions is consistent with the long-term average. However, the percentage of projects in the APR 2020 cohort rated in the likely range for sustainability, and satisfactory range for quality of execution, is slightly higher than the long-term average.

Methodology

9 The reporting is based on 1,706 completed GEF projects that together account for \$7.5 billion in GEF grants, including support for project preparation, and \$36.2 billion in promised cofinancing. These include 134 projects in the APR 2020 cohort, for which the terminal evaluations were submitted in the calendar year 2019, which are being covered in the APR reporting for the first time. The APR 2020 cohort accounts for \$618.8 million in GEF grants and \$4.4 billion in promised cofinancing. UNDP (29 percent), World Bank (27 percent), UNEP (19 percent), UNIDO (8 percent), and FAO (7 percent) account for a substantial share of the projects included in the APR 2020 cohort.

10 The GEF IEO's ongoing effort to fill data gaps for projects that were completed 15 or more years ago led to the retrieval of terminal evaluation reports for six projects. These projects are not included in the APR 2020 cohort, but their terminal evaluations have been reviewed and their performance ratings included in the cumulative terminal evaluation review dataset that has been used for analysis. A detailed list of projects included in the 2020 cohort is provided in annex A. The dataset on performance ratings of the completed projects will be made available on the GEF IEO website by the end of FY2020. The terminal evaluation review reports prepared by the GEF IEO will also be made publicly available.

11 Either the GEF IEO or the Agency evaluation offices, or both, assess performance of the completed GEF projects. The GEF IEO reviews terminal evaluation reports, PIRs, midterm reviews, and project documents. The IEO rates the performance of the completed projects based on evidence available in these documents. Project outcomes, project implementation and execution, and M&E design and implementation are rated on a six-point scale ranging from highly satisfactory to highly unsatisfactory, with the top three ratings constituting the satisfactory range and the bottom three the unsatisfactory range. The sustainability of project outcomes is measured on a four-point scale, ranging from likely to unlikely. The methodology used in rating project performance is detailed in annex C.

12 There are two types of sources of the ratings presented in APR 2020. The GEF IEO rates 49 percent of projects (831 projects) and the Agency evaluation offices rate fifty one percent

(875 projects). From 2009 onwards, GEF IEO has been accepting the performance ratings from evaluation offices based on whether these ratings are available when the terminal evaluations are submitted and whether there is prior evidence of consistency between the Agency and IEO ratings. Thus far, the GEF IEO has accepted the project performance ratings from the World Bank, UNDP, and UNEP, which together account for 88 percent of the completed projects (1,496 of 1,706 projects) and 96 percent of the projects for which project performance ratings by the Agency evaluation office were available (1,006 out of 1,053 projects).

13 To address the possible inconsistencies in ratings across Agencies, GEF IEO conducts a terminal evaluation review for a random sample of the terminal evaluations prepared by the World Bank, UNDP, and UNEP. Usually, 10 to 15 percent of the terminal evaluations prepared by these three agencies are selected. So far, 196 projects have been rated by the GEF IEO for quality assurance purposes. On a six-point scale, the average difference in the outcome rating from the two sources is 0.05, which is not statistically significant. Binary-scale on the net ratings match exactly for 94 percent of the projects (184 projects). Thus, there is no substantial difference in the validation of ratings by the IEO or by the Agency.

Findings

Outcome and Sustainability

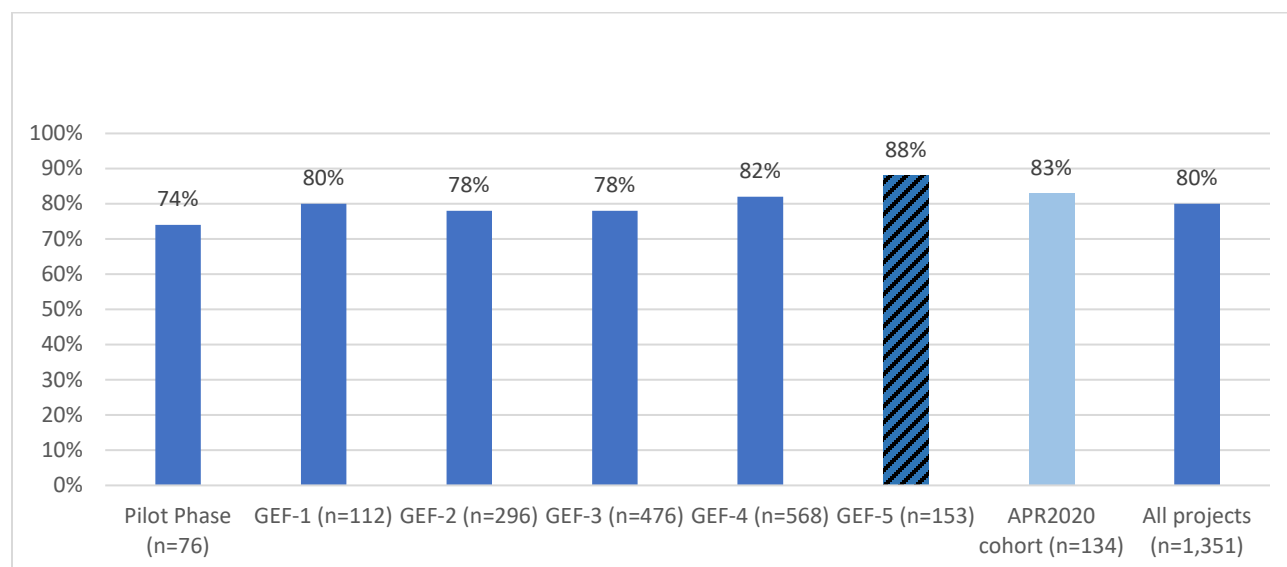
14 **Eighty percent of GEF projects have satisfactory outcomes.** The outcome rating assesses the extent to which a completed project achieved the outcomes expected at implementation completion. Cumulatively, 80 percent of all the rated projects, which account for 78 percent of the GEF grant, are rated in the satisfactory range for outcomes. In comparison, in the APR 2020 cohort, 83 percent of the rated projects that account for 83 percent of the funding are rated in the satisfactory range. The difference between the outcome ratings of the APR 2020 cohort and other completed projects is not statistically significant. The distribution of GEF funding by outcome rating closely tracks the distribution of projects (table 1). The percentage of projects rated in the satisfactory range has consistently been around 80 percent from GEF-1 onwards. Although 88 percent of the completed projects from the GEF-5 period are rated in the satisfactory range, the number of completed projects from the period are still low, and, therefore, likely to change (**Error! Reference source not found.**).

Table 1: Distribution of completed GEF projects and funding by ratings

Outcome Rating	Percentage of Projects		Percentage of GEF Funding	
	APR 2020 Cohort (n=134)	All Other Projects (n=1471)	APR 2020 Cohort	All Other Projects
Highly satisfactory	7	4	7	4
Satisfactory	40	38	34	36
Moderately satisfactory	36	38	42	38
Moderately unsatisfactory	13	15	12	17
Unsatisfactory	4	5	4	5
Highly unsatisfactory	1	<1	1	1

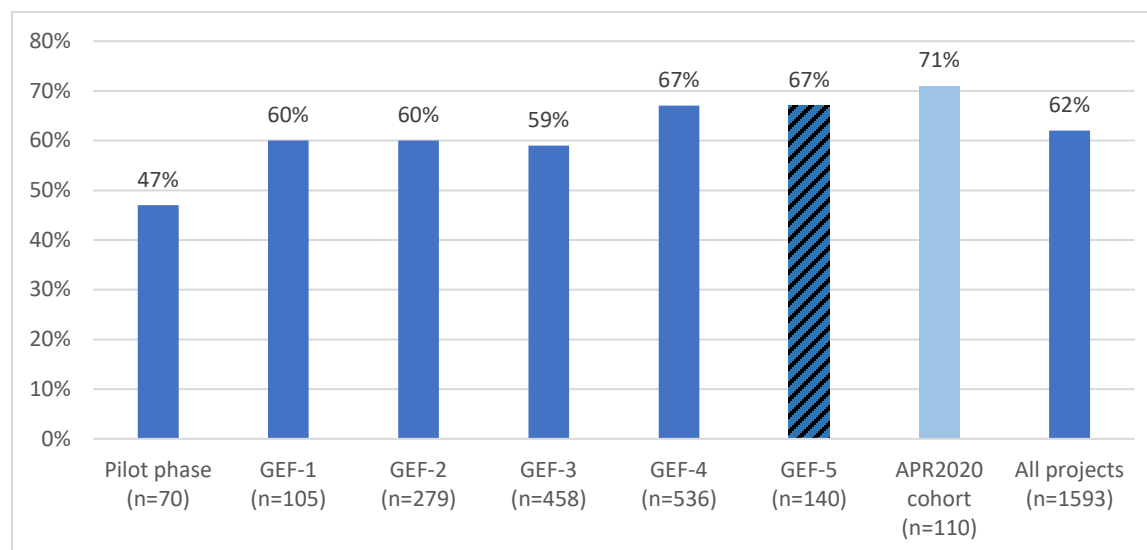
Source: GEF IEO terminal evaluation review data set 2020

Figure 1: Projects with outcomes rated in the satisfactory range by GEF replenishment period



15 Project sustainability ratings have improved over time. The sustainability rating estimates the extent to which a project's outcomes are durable and the project is likely to achieve its expected long-term impact. Cumulatively, 62 percent of the completed projects are rated in the likely range for sustainability (Figure 2). In comparison, 71 percent of the projects in the APR 2020 cohort are so rated—the difference is statistically significant at the 90 percent confidence level. The findings also indicate that projects that were approved in GEF-4 and GEF-5 have higher sustainability ratings than the preceding periods. However, the sustainability rating of a third of projects from GEF-4 and GEF-5 is in the unlikely range. GEF IEO will undertake a more detailed analysis of the topic for APR 2021.

Figure 2: Projects with sustainability of outcomes rated in the likely range by replenishment period

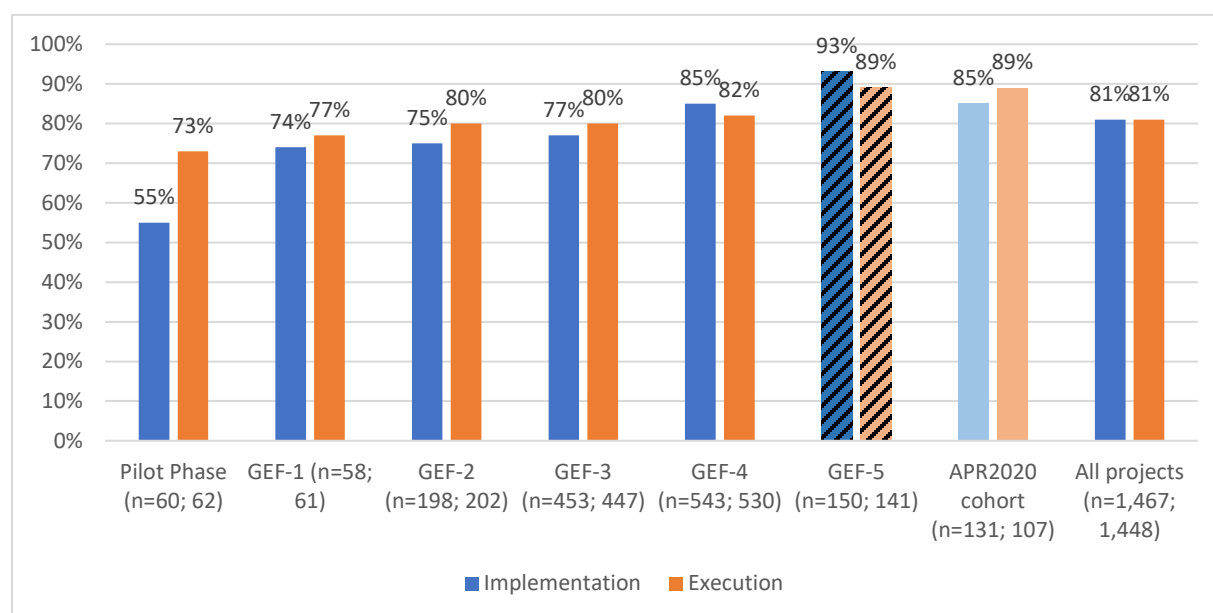


Implementation and Execution

16 The percentage of projects in the APR 2020 cohort rated in the satisfactory range for implementation is consistent with the long-term average. The quality of the implementation rating is based on an assessment of the performance of the GEF Agency. It reflects the extent to which the GEF Agency has performed its role satisfactorily in project identification and preparation, start-up, supervision, application of the GEF policies and procedures, and project M&E. It also reflects the extent to which the GEF Agency took timely corrective actions when gaps in project design and implementation were found. To date, 1,467 completed projects have been rated on quality of project implementation; of these projects, 131 are from the APR 2020 cohort. Of the rated projects, 81 percent are in the satisfactory range (figure 8). In comparison, 85 percent of the projects of the 2020 cohort are in the satisfactory range. Although the ratings for the APR 2020 cohort are higher, the difference is not statistically significant.

17 Project execution has significantly improved over time with a higher percentage of projects in the APR 2020 cohort rated in the satisfactory range for execution than the long-term average. Executing agencies are responsible for execution of the project activities on the ground under the supervision of the GEF Agency. This involves activities such as execution of project design, procurement, stakeholder consultations, and project monitoring. The quality of the execution rating reflects the extent to which an executing agency performed its responsibilities satisfactorily. To date, 1,448 completed projects have been rated on quality of project execution; of these, 81 percent have been rated in the satisfactory range (Figure 3). Of the 107 projects from the APR 2020 cohort that were rated for quality of project execution, 89 percent were rated in the satisfactory range. The ratings for the APR 2020 cohort are significantly different than other completed projects in the GEF portfolio. Overall, the improving trend in the execution rating continues.

Figure 3: Projects with implementation/execution quality rated in the satisfactory range by GEF replenishment period



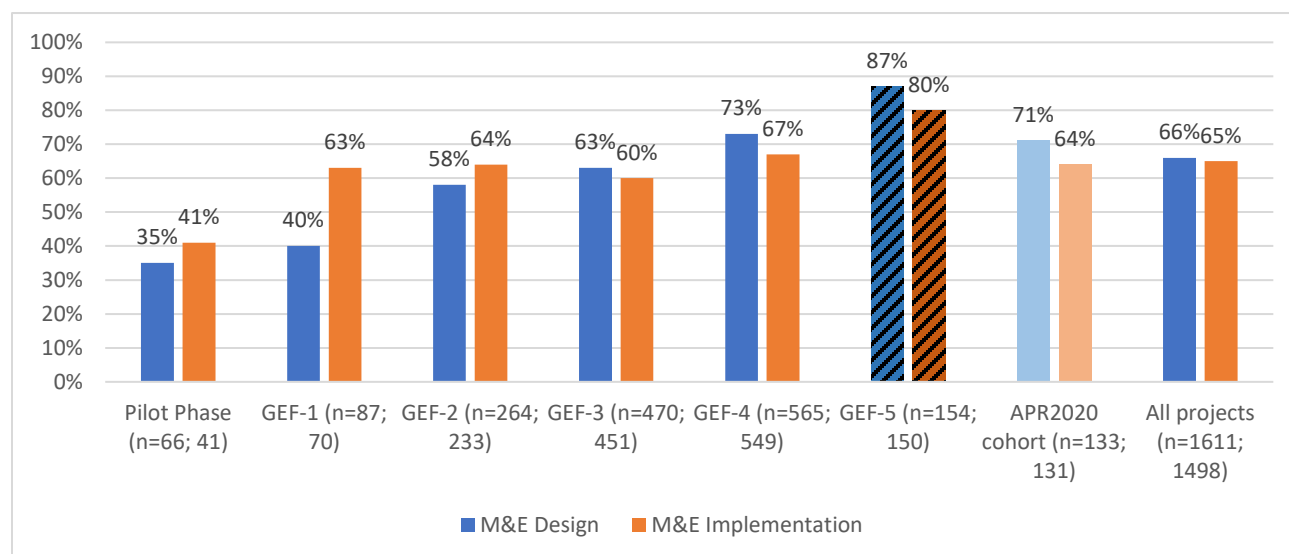
Project Monitoring

18 Project monitoring involves the design and implementation of an M&E plan to track implementation progress and results. GEF Agencies include an M&E plan in the project design and implement this plan during project implementation. An M&E plan is expected to specify indicators to track processes and results, responsibilities, frequency of data collection, reporting procedures, and budget for monitoring activities. Where required, an M&E plan may need to be updated/modified during implementation. Quality of M&E design and implementation are rated to reflect the extent to which an M&E plan was well designed and well implemented.

19 **The percentage of projects rated in the satisfactory range for quality of M&E design is same as the long-term average.** To date, the quality of M&E design has been rated for 1,611 completed projects of which 66 percent have been rated in the satisfactory range. Of the 133 projects in the APR 2020 cohort that were rated for M&E design, 71 percent rated in the satisfactory range, which is not significantly different from the ratings for other completed projects in the GEF portfolio. However, there has been a steady improvement in the quality of project M&E design over various replenishment periods (Figure 4).

20 **Sixty four percent of projects are rated as satisfactory for M&E implementation and this has not changed significantly.** To date, 1,498 completed projects have been rated for M&E implementation, of which 65 percent are rated in the satisfactory range. Of the 131 projects in the APR 2020 cohort rated for M&E implementation, 64 percent were rated in the satisfactory range, which is not significantly different from the ratings for other completed projects in the GEF portfolio. Although there have been some gains in the ratings across the replenishment periods, they have been relatively modest compared with the gains in the ratings for quality of M&E design.

Figure 4: Projects with M&E design/implementation rated in the satisfactory range by GEF replenishment period



Cofinancing

21 GEF-supported activities generally involve both GEF financing and cofinancing by other partners. Cofinancing may increase the scale of the GEF-supported activities and enhance the ability of GEF to generate global environment benefits. The extent to which GEF projects may mobilize cofinancing depends on factors such as the level of GEF funding, focal area and targeted sectors, and recipient country context. However, what is more important is that the cofinancing expected in the project design materializes in a timely manner.

22 **The average level of materialized cofinancing is 105 percent of the amount promised at project start, and for 61 percent of projects at least 90 percent of the promised cofinancing materialized.** Data on materialized cofinancing is available for 1,430 projects (Table 2). The data show that on average for these projects, cofinancing of \$5.2 per dollar of GEF grant was promised at project start. In comparison, on average, cofinancing of \$6.49 per dollar materialized to yield a materialization ratio of 1.25. Of the completed projects, for two-thirds, at least 90 percent of promised cofinancing materialized during implementation. For 16 percent, the materialization was less than 50 percent—only 63 percent of these projects were rated in the satisfactory range for outcomes compared with 85 percent of projects in which materialization of cofinancing was 50 percent or higher. Data on materialization of cofinancing is available for 116 projects in the APR 2020 cohort. For the APR 2020 cohort, \$7.98 per dollar materialized, compared with \$7.58 promised per dollar of GEF grant, for a materialization ratio of 1.05. Of the projects in the APR 2020 cohort, for 61 percent, at least 90 percent of the promised cofinancing materialized; for 26 percent, materialization was less than 50 percent.

Table 2: Materialization of cofinancing

Phase	Projects	Cofinancing Promised per \$ of GEF Grant	Cofinancing Materialized per \$ of GEF Grant	Cofinancing Materialization Ratio	% of Projects with at Least 90% Materialization	% of Projects with Less Than 50% Materialization
Pilot phase	59	6.80	6.54	0.96	67	16
GEF-1	93	2.86	2.34	0.82	55	17
GEF-2	253	4.91	5.83	1.19	71	11
GEF-3	408	4.59	5.59	1.21	67	16
GEF-4	492	6.53	10.02	1.54	64	19
GEF-5	120	5.96	5.86	0.98	68	19
GEF-6	5	8.94	9.27	1.04	80	0
APR 2020	116	7.58	7.98	1.05	61	26
All projects	1430	5.18	6.49	1.25	66	16

CHAPTER 3. THE MANAGEMENT ACTION RECORD 2020

23 The GEF MAR tracks the level of adoption of the GEF Council and Least Developed Countries Fund (LDCF)/Special Climate Change Fund (SCCF) Council decisions that are based on the recommendations of the evaluations conducted by the GEF IEO. The GEF Secretariat and the GEF Agencies, referred to as GEF management as applicable, are responsible for adoption of the Council's decisions. The MAR serves two purposes: "(1) to provide Council a record of its decisions based on the evaluation reports presented by the GEF IEO, the proposed management actions, and the actual status of these actions; and (2) to increase the accountability of GEF Management regarding Council decisions."⁵

24 MARs are published as a separate document by the GEF IEO. MAR 2020 reports on the level of adoption of decisions based on GEF IEO recommendations included in eight different evaluations:

- (a) "Joint GEF-UNDP Small Grants Programme Evaluation" (GEF/ME/C.48/02) reported in "Semi-Annual Evaluation Report, June 2015" (GEF/ME/C.48/02)
- (b) "Evaluation of the GEF Civil Society Organization Network" (GEF/ME/C.50/02)
- (c) "Review of the GEF Policy on Minimum Standards on Environmental and Social Safeguards" (GEF/ME/C.52/Inf.08) reported in "Semi-Annual Evaluation Report: May 2017" of the GEF IEO (GEF/ME/C.52/01/Rev.02)
- (d) "Review of GEF Support for Transformational Change" (GEF/ME/C.52/Inf.06) reported in "Semi-Annual Evaluation Report: May 2017" of the GEF IEO (GEF/ME/C.52/01/Rev.02)
- (e) "Review of GEF's Engagement with Indigenous Peoples" (GEF/ME/C.53/Inf.07) reported in "Semi-Annual Evaluation Report of the Independent Evaluation Office: November 2017" (GEF/ME/C.53/01)
- (f) "Biodiversity Focal Area Study" (GEF/ME/C.53/Inf.03) reported in "Semi-Annual Evaluation Report of the Independent Evaluation Office: November 2017" (GEF/ME/C.53/01)
- (g) "Program Evaluation of the Least Developed Countries Fund" (GEF/LDCF.SCCF.20/ME/02)
- (h) "Program Evaluation of the Special Climate Change Fund" (GEF/LDCF.SCCF.22/ME/02)

⁵ GEF Council, "Procedures and Format of the GEF Management Action Record." GEF/ME/C.27/3., GEF Council November 2005.

25 Of the eight evaluations, six evaluations were presented to, and their recommendations endorsed by, the GEF Council. The remaining two evaluations were presented to, and their recommendations endorsed by, the LDCF/SCCF Council. These eight evaluations contained 26 recommendations of which 24 pertained to the six evaluations presented to the GEF Council, and two recommendations pertained to the two evaluations presented to the LDCF/SCCF Council.

26 During 2017, the GEF council endorsed 58 recommendations presented to it by the GEF IEO through the May and November 2017 Semi-Annual Evaluation Reports. Progress in adoption of these recommendations was not reported in MAR 2017 because it was still too early to take stock of the progress. MAR 2020 tracks and reports on progress in adoption of 21 of these 58 recommendations. Of these 21 recommendations, five pertain to the “Review of GEF’s Engagement with Indigenous Peoples” (GEF/ME/C.53/Inf.07), three to the “Review of the GEF Policy on Minimum Standards on Environmental and Social Safeguards” (GEF/ME/C.52/Inf.08), 12 to the “Biodiversity Focal Area Study” (GEF/ME/C.53/Inf.03), and one to the “Review of GEF Support for Transformational Change” (GEF/ME/C.52/Inf.06). The remaining 37 recommendations will be covered in future MARs.

27 The Third Professional Peer Review of the Independent Evaluation Function of the Global Environment Facility, being presented concurrently to the June 2020 GEF Council meeting, includes a recommendation that calls for changes to the current MAR approach that involves a higher level of participation by the Council in drafting the Council decisions while considering the Management Response. The GEF IEO will follow up with the Council on this recommendation on improving the system, with a planned rollout of a new MAR system in June 2021.

Rating Approach

28 For each tracked GEF Council and LDCF/SCCF Council decision that is reported on, the GEF management provides self-ratings on the level of adoption along with commentary, as necessary. Ratings and commentary on tracked decisions are also provided by the GEF IEO for verification. The rating categories for the progress of adoption of Council decisions were agreed upon by the GEF IEO, the GEF Secretariat, and the GEF Agencies, through a consultative process. Categories are as follows:

- (a) **High:** Fully adopted and fully incorporated into policy, strategy, or operations
- (b) **Substantial:** Decision largely adopted but not fully incorporated into policy, strategy, or operations, as yet
- (c) **Medium:** Adopted in some operational and policy work, but not to a significant degree in key areas
- (d) **Negligible:** No evidence or plan for adoption, or plan and actions for adoption are in a very preliminary stage.
- (e) **Not rated:** Ratings or verification will have to wait until more data are available or proposals have been further developed.

(f) **N/A:** Not applicable (see commentary)

29 The Council decisions may be graduated or retired from the MAR because of one or more of the following reasons:

- (a) **Graduated** due to high or, where appropriate, substantial level of adoption of Council decision
- (b) **Retired** as the Council decision has become less relevant, or subsequent Council decisions have made high level of adoption of the decision difficult, or further progress on adoption of the decision is likely to be slow and long drawn. An automatic reason for retirement would be if a decision has been reported on in the MAR for five years.

The GEF IEO Keeps Track of the Reasons for Removing a Decision from the MAR

Findings

30 **The GEF IEO and the GEF Secretariat agree on ratings of adoption of council decisions for all eight evaluations tracked in MAR 2020 with the adoption level rated substantial for seven evaluations and medium for one.** The council decision based on the “Program Evaluation of the Least Developed Countries Fund,” for which adoption has been rated substantial, would be retired because a new Council decision is expected on the 2020 Update of the Program Evaluation of the LDCF in December 2020.

GEF Council Decisions with a Substantial Level of Adoption

31 The GEF Council decision based on the “Evaluation of the GEF Civil Society Organization Network” encouraged the network to establish a working group with a balanced representation to interact with the Council Working Group on an updated vision for the network, including governance, policies, and cooperation mechanisms. The GEF IEO and the management agree that there has been substantial progress in adopting the Council decision. During the reporting period, the GEF Secretariat organized three consultations with civil society organizations and presented the Council with information on civil society organization, indigenous peoples and local communities, and private sector engagement in the GEF corporate scorecard. The GEF IEO is currently conducting an evaluation of the institutional policies and engagement, including the stakeholder engagement policy.

32 The “Review of GEF Support for Transformational Change” recommended the development and application for a framework for ex-ante assessments of projects or programs intended to be transformational. GEF IEO and the management agree that progress on adoption of these recommendations has been substantial, as the GEF has embraced the proposed framework and included considerations for transformational change in the programming direction for GEF-7, including impact programs.

33 The “Review of the GEF Policy on Minimum Standards on Environmental and Social Safeguards” recommended that GEF should: review its minimum standards for environmental and social safeguards; improve monitoring of safeguards and reporting; and support capacity

development, convening of experts, and communications. The GEF IEO and the management agree that progress on adoption of these recommendations has been substantial. Following on the preparation of an updated policy on environmental and social safeguards, the GEF Secretariat developed guidelines, as well as an assessment of GEF Agencies' compliance, both presented at the 57th Council meeting. The GEF IEO continues to track the development of an environmental and social safeguards template and GEF Secretariat support in capacity development, expert convening, and communications, which has taken place at Extended Constituency Workshops and Agency's retreats.

34 The biodiversity focal area study included 12 recommendations related to project design and overall strategy for GEF programming in biodiversity. Although some recommendations could not yet be rated, overall progress on adoption of the recommendations is rated substantial by both the GEF IEO and management. Much progress is related to the GEF-7 Global Wildlife Program, which has increased geographic and species coverage of the GEF portfolio of biodiversity-related projects.

35 The GEF IEO recommendations in the "Review of GEF's Engagement with Indigenous Peoples" called for: dedicated funding opportunities for indigenous people's organizations; update of policies and guidelines; review of the role of the indigenous peoples' advisory group; and improved reporting on engagement of indigenous people and relevant results through midterm reviews and terminal evaluations. Both the GEF IEO and the management assessed the overall progress on adoption of the recommendations to be substantial. A new policy for indigenous people was prepared and approved by the Council in the previous reporting period, and the GEF IEO will evaluate its implementation as part of the Seventh Comprehensive Evaluation of the GEF.

36 The Program Evaluation of the Least Developed Countries Fund recommended that the GEF Secretariat should: explore and develop mechanisms that ensure the predictable, adequate, and sustainable financing of the Fund; make efforts to improve consistency regarding their understanding and application of the GEF gender mainstreaming policy and the Gender Equality Action Plan to the LDCF; and ensure that the data in the GEF PMIS is up to date and accurate. Both the GEF IEO and the management assessed the overall progress on adoption of the recommendations to be substantial. The GEF Secretariat has co-organized a successful pledging event where additional resources for the LDCF were pledged by several donors. The GEF Policy on Gender Equality, which is also applicable to LDCF activities, is regarded by the GEF IEO as substantial progress on the adoption of the gender mainstreaming-related recommendation. Although the transition to the portal is still underway, the systems in place to ensure that portal data—including project status, key dates, and financial figures—are continually updated and kept accurately moving forward is not clear. A new Council decision will be made on the 2020 Update of the Program Evaluation of the LDCF in December 2020 and this decision is therefore retired.

37 The Program Evaluation of the Special Climate Change Fund called the GEF Secretariat to prioritize sustainable financing for the Fund; to describe the SCCF's niche within the global adaptation finance landscape; and, ensure that PMIS data is up to date and accurate. The GEF IEO assesses the overall progress on adoption of these recommendations to be substantial.

Much of the progress made is in four terms of the GEF Programming Strategy on Adaptation for the Special Climate Change Fund for 2018–22, which the IEO will continue to track. Progress on ensuring sustainable funding has been medium. Progress on the transition from PMIS to the GEF Portal is described in the preceding paragraph.

Decision with a Medium Level of Adoption

38 The assessment of the GEF IEO and the management on the level of adoption of the decision based on the “Joint GEF-UNDP Small Grant Programme Evaluation” is medium, as it was the previous year. The Council’s decision had called for reconsideration of the criteria for upgradation of the participating countries. The management reports that it has reconsidered the criteria for upgradation but is using it without any changes for the GEF-7 period. One country, Malaysia, was upgraded in the previous period and a full-size project for \$2.5 million for the country was approved by the Council as part the December 2019 Work Program. Discussions between the GEF Secretariat and UNDP on the Small Grants Programme implementation are ongoing, the third joint Small Grants Programme evaluation will also include a focus on country upgrading.

Retired Decisions

39 The council decision based on the “Program Evaluation of the Least Developed Countries Fund” that recommended that the GEF Secretariat explore and develop mechanisms to ensure the predictable, adequate, and sustainable financing of the Fund, makes efforts to improve consistency regarding their understanding and application of the GEF gender mainstreaming policy and the Gender Equality Action Plan to the LDCF, and to ensure that the data in the Project Management Information System is up to date and accurate, has been retired with a substantial rating for adoption. A new Council decision will be made on the 2020 Update of the Program Evaluation of the LDCF in December 2020. All other decisions will continue to be tracked in MAR 2021.

Section B

CHAPTER 4. QUALITY OF REPORTING

40 Comprehensive, candid, and timely reporting on the performance and implementation of GEF activities facilitates a realistic assessment of the health of the GEF portfolio and identification of emerging concerns. It allows the management to draw lessons, identify specific areas where corrective action may be required, and deploy resources to bring such activities back on track. Credible reporting also ensures accountability for use of public funds. Therefore, it is important to give adequate attention to the quality of reporting of GEF activities.

41 Beginning with APR 2004—the first APR—several APRs have reported on the quality of terminal evaluation reports. APR 2020 presents a more detailed analysis on this topic than previously, drawing on a structured review of terminal evaluation content. It provides an assessment of the trends in quality of terminal evaluations, level of compliance with the GEF IEO’s “Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects” (GEF IEO 2017), and candor in reporting through PIRs.

42 The analysis presented in this chapter shows that the quality of terminal evaluations has improved, although the rate of improvement has slowed. It shows that compliance with the GEF IEO’s terminal evaluation guidelines is uneven across Agencies, and, although some Agencies are better at complying with the requirements, all Agencies have areas of weak compliance. The analysis shows that although the guidelines are meant for the terminal evaluations for full-size projects, the evaluations for medium-size projects perform equally well at following the guidelines.

43 The reasons for variation in compliance with the guidelines, both at the GEF Partnership and at the Agency level, are not fully clear. These will be examined in the GEF IEO’s ongoing evaluation of the Self Evaluation Systems of GEF Agencies.

44 The analysis of the documentation for canceled and low-performing projects shows that PIRs and terminal evaluations generally present a more optimistic picture of project performance than validations conducted by the GEF IEO. Although PIRs tend to rate performance more optimistically as compared to terminal evaluations, the analysis shows that for a vast majority of projects, they flag the relevant concerns and challenges in a timely manner for projects that were eventually rated unsatisfactory or lower for outcome. The analysis also shows that for the canceled projects, information on causes that led to cancellation is available in almost all the instances.

Key Questions

45 The analyses presented in this chapter seek to answer the following questions:

- (a) To what extent do terminal evaluation reports meet the criteria used to assess quality?
- (b) To what extent do terminal evaluations follow the GEF IEO’s guidelines for preparation of terminal evaluations?

- (c) To what extent is the reporting through PIRs on implementation challenges and performance candid in terms of timeliness and sufficiency of attention to emerging concerns?

Methodology

Quality of terminal evaluations

46 The first question is answered through an assessment of quality of terminal evaluations submitted to the GEF IEO. These assessments are undertaken by either the GEF IEO or the Agency evaluation offices or both (annex C includes GEF IEO guidelines for the terminal evaluation review). The quality of terminal evaluations is based on the following six criteria:

- (a) Quality of reporting on outcomes
- (b) Consistency of reporting, completeness of evidence, and ratings are well substantiated.
- (c) Quality of reporting on sustainability
- (d) Quality of lessons and recommendations
- (e) Reporting on actual project costs and cofinancing
- (f) Quality of reporting on project M&E

47 The GEF IEO rates the quality of terminal evaluations on each of these six criteria on a six-point scale. These ratings are then used in an index to determine the overall quality of the report, again on a six-point scale⁶. Of the 1,706 terminal evaluation reports submitted to the GEF IEO, the overall quality has been rated for of 1,679. Ninety-eight percent of the (831) terminal evaluation reviews prepared by the GEF IEO rates the quality of the terminal evaluations on each of the six criteria. Although 97 percent of the (875) terminal evaluations in which quality has been rated by the Agency evaluation offices are rated for overall quality on a six-point scale, there are gaps in the ratings for the individual criteria and only 34 percent are rated for all six criteria. As a result, the disaggregated ratings on the six quality criteria are available for 1,109 terminal evaluations. Nonetheless, the number of observations for individual quality criterion range from 1,135 to 1,211. The ratings for overall quality and individual criterion have been used for reporting on broad quality trends and identifying differences across Agencies and between full-size and medium size projects.

⁶ The weights assigned to the index are as follows: outcome (0.3) + consistency (0.3) + sustainability (0.1) + lessons (0.1) + financing (0.1) + M&E (0.1) = overall quality rating.

Compliance with the terminal evaluation guidelines

48 Compliance with the GEF IEO's "Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects" (GEF IEO 2017) was assessed by determining the extent terminal evaluations meet the requirements given in the guidelines. These guidelines were issued in April 2017 and replaced the preceding guidelines that were issued in 2008 (GEF IEO 2008). The new guidelines are applicable to full-size projects. However, the Agencies are encouraged to use these guidelines as a reference to develop their approach to conduct terminal evaluations for medium-size projects and programs. Therefore, data for terminal evaluations for full-size projects are analyzed and presented from a compliance perspective and that for medium-size projects from a normative perspective.

49 The 2017 guidelines describe reporting requirements on topics such as project identification, basic project details, feedback from stakeholders, theory of change, methodology, outcome, sustainability, implementation, project M&E, environmental and social safeguards, gender concerns, lessons, and recommendations. The guidelines also provide directions on timeliness of the conduct of terminal evaluations and their submission to the GEF IEO. The instrument to assess compliance is included in Annex D.

50 Given that the review assesses compliance with the new guidelines, it is restricted to those terminal evaluations of the APR 2020 cohort that were prepared in or after October 2017—at least six months after the issuance of the new guidelines. The six-month threshold was chosen because it presents a reasonable timeframe to provide an opportunity for Agencies to incorporate the new guidance, as they were issued after a consultative process with the GEF Agencies and their details had been shared prior to issuance. However, in practice, the Agencies may differ in terms of their ability to apply the guidelines quickly in their practice for preparing terminal evaluations for completed GEF projects. For example, UNEP updated its generic terms of reference for completed GEF projects in 2017 and UNIDO incorporated the GEF IEO guidelines in its evaluation manual in 2018. Other Agencies have taken additional time or are yet to mainstream it in their practices. This, in turn, may reflect the extent to which the terminal evaluations prepared by them meet the GEF requirements.

51 Of the 134 terminal evaluations submitted during the calendar year 2019, 121 were prepared in or after October 2017. Of these, three terminal evaluations for full-size projects, two prepared by UNDP (GEF IDs 3750 and 4731), and one by the International Fund for Agricultural Development (IFAD; GEF ID 3963) were dropped because they were not submitted in English, which is a requirement. Of the remaining 118 terminal evaluations, 80 report on completed full-size projects and 38 on medium-size projects. The terminal evaluations for these projects were covered through the review.

52 The level of Agency compliance with the terminal evaluation guidelines is assessed at the individual criterion level, and at the reporting dimension level into which these criteria are organized. For example, the assessment of compliance with requirements related to the reporting dimension 'project's theory of change' will include examination of whether the given terminal evaluation provides an overall statement of the project's theory of change; its causal chain, key assumptions, and reports on whether these assumptions were found to be valid

during implementation. First, compliance with each individual criterion is assessed. Second, within a reporting dimension the number of criteria for which a terminal evaluation is in compliance are added and then divided by the total number of criteria assessed for a given cluster. This gives us a compliance score for individual reporting dimension. There is no overall compliance score for the terminal evaluation as this is already captured qualitatively in the rating for quality of terminal evaluation that are provided through the validations conducted by the GEF IEO.

Candor in reporting

53 Question three addresses candor in reporting, mainly through terminal evaluations and PIRs. Candor is generally understood as being frank – open and honest – in reporting. The GEF IEO has addressed the issue of candor, especially to evaluate the supervision related responsibilities of the Agencies, through APR 2006, APR 2008 and OPS-4. Candor in reporting implementation challenges and performance related concerns is answered through assessments that examine Agency reporting on projects that were rated unsatisfactory or lower for outcomes or were canceled after CEO’s endorsement/approval. The premise is that these two sets of projects faced problems that resulted in an unsatisfactory or lower outcome rating or cancellation, or both. Therefore, if the PIRs prepared by the Agencies report these problems and challenges with candor, these concerns should have been flagged by the PIRs in real time with a fair reflection of the importance of the concern. These projects provide an opportunity to assess whether in these cases—in which there may be disincentives for candor—the system worked as it was supposed to work. The review primarily covers reporting through PIRs, although, where available, reporting through midterm reviews has also been covered. The information provided by the PIRs and midterm reviews was checked against that provided by the terminal evaluations. The aim is to assess timeliness and consistency in reporting of challenges and concerns. For example, if a terminal evaluation reports a concern that became evident at a certain point during project implementation, then the PIR that covered the period in question may be expected to report on the given concern. The review ascertains whether this happened.

54 Of the completed GEF projects for which terminal evaluations are available, 1,083 projects were approved in GEF-3 or later with terminal evaluations submitted in calendar year 2018 or earlier. Of these 1083 projects, 1070 were rated for quality of outcomes at completion. Of these, 50 projects were rated unsatisfactory or lower and were selected for examination (**Error! Reference source not found.**). The analysis was restricted to those rated unsatisfactory or lower on outcomes because it may be said with greater certainty that there were shortcomings in either design or implementation or both, or that these projects experienced unforeseen exogenous shocks during implementation.

55 An instrument was used to gather information from PIRs, midterm reviews, terminal evaluation, and other formal communication from the Agency (annex F). The instrument aims at developing a timeline of events and factors that led to unsatisfactory outcomes, when these could have been—and were—reported on, and whether reporting was consistent with the challenge on hand.

Table 3: Unsatisfactory or lower outcome rating—by GEF Agency

Lead Agency	Projects Rated	Rated Unsatisfactory or Lower
ADB	13	0 (0%)
AfDB	1	0 (0%)
CI	1	0 (0%)
EBRD	1	0 (0%)
FAO	25	0 (0%)
IDB	18	1 (6%)
IFAD	26	2 (8%)
Joint Projects	42	0 (0%)
UNDP	584	21 (4%)
UNEP	165	3 (2%)
UNIDO	51	1 (2%)
World Bank	270	22 (8%)
WWF	2	0 (0%)
Total	1070	50 (5%)

Note: ADB=Asian Development Bank. AfDB=African Development Bank. CI=Conservation International. EBRD=European Bank for Reconstruction and Development. IDB=Inter-American Development Bank. WWF= World Wildlife Fund.

56 The duration for which these projects were under implementation ranges from two to 12 years, with an average of 5.8 years. Of the 50 projects PIRs are available for 44 (88 percent). For these 44 projects, on average, 4.4 PIRs are available per project—the median value of available PIRs is 4. For 13 projects, a midterm review is also available.

57 The second set of projects include canceled projects that were approved from GEF-3 onwards as of December 31, 2018. The review was restricted to canceled projects that had received CEO endorsement or CEO approval before cancellation, because a paper trail that highlights reasons for cancellation may be expected for these. Of the 98 canceled projects approved from GEF-3 onwards, 32 had been CEO endorsed. These 32 projects include those prepared by the World Bank (12 projects), UNDP (nine projects), UNEP (three projects), IADB (two projects), Asian Development Bank (ADB; 3 projects), IFAD (two projects), and FAO (one project). Of the 32 projects, 18 were implemented for some time, and for 16 at least one PIR had been submitted before cancellation. The remainder (12 projects) were canceled before the start of implementation. For each of these projects, available documentation such as email communications, PIRs, and cancellation memorandum, was examined. An instrument was used to gather information on characteristics of the project, availability of PIRs and other reports, incidence of unsatisfactory progress in project start up, and reasons for cancellation (annex E). The aim of the examination is to construct a timeline of events leading to project cancellation and to determine whether the given Agency shared information on the experienced concerns with the GEF with candor and without delay.

Limitations

58 The number of terminal evaluations reviewed to assess compliance with the GEF IEO guidelines is enough to give an overview of level of compliance. However, it is generally difficult to compare performance of GEF Agencies, except for dimensions where the difference in performance is substantial. This constraint may be mitigated in the future by including data for more projects in the analysis.

59 As noted, although the terminal evaluation guidelines require that terminal evaluations are prepared in English, three terminal evaluations included in the APR 2020 cohort were prepared in Spanish or French. Given the ongoing pandemic, it was difficult to find suitable reviewers to assess quality of these terminal evaluations. These terminal evaluations had to be dropped from the analysis.

60 Although the use of desk reviews to assess compliance with the terminal evaluation guidelines helps in identifying the areas where compliance is low, it is not effective in examining the reasons that may explain variations. For example, it is difficult to know whether the observed gaps in compliance are due to a low level of awareness of the GEF IEO guidelines or due to other barriers that are internal to the GEF Agencies. This information gap will be addressed by the GEF IEO through additional information gathering efforts as part of its ongoing evaluation on the Agency Self-Evaluation Systems.

61 The assessment of consistency of information in PIRs and terminal evaluations draws on the PIRs that are available within the GEF Portal, or the archived PMIS. For approximately half of the projects, at least one PIR is missing. This gap does not pose a risk to the findings: for any given project there is continuity in the information reported through subsequent PIRs. However, there is a chance that a PIR essential for establishing timeliness is occasionally missing.

62 The projects that were rated unsatisfactory or lower on outcomes are so rated because of the evidence provided in terminal evaluations. The assessment of candor of reporting is based on comparison of the information from PIRs and midterm evaluations with the terminal evaluation. This approach works well when reporting from PIRs and midterm reviews is inconsistent with that of the terminal evaluation. It also works well when reporting is fully consistent across instruments and candid. However, it does not work well when the reporting through PIRs and midterm review is consistent with the information from terminal evaluation, but all information from all these sources lacks candor. Identification and examination of these projects are beyond the scope of the analysis that was conducted. This is, therefore, a limitation of the review design.

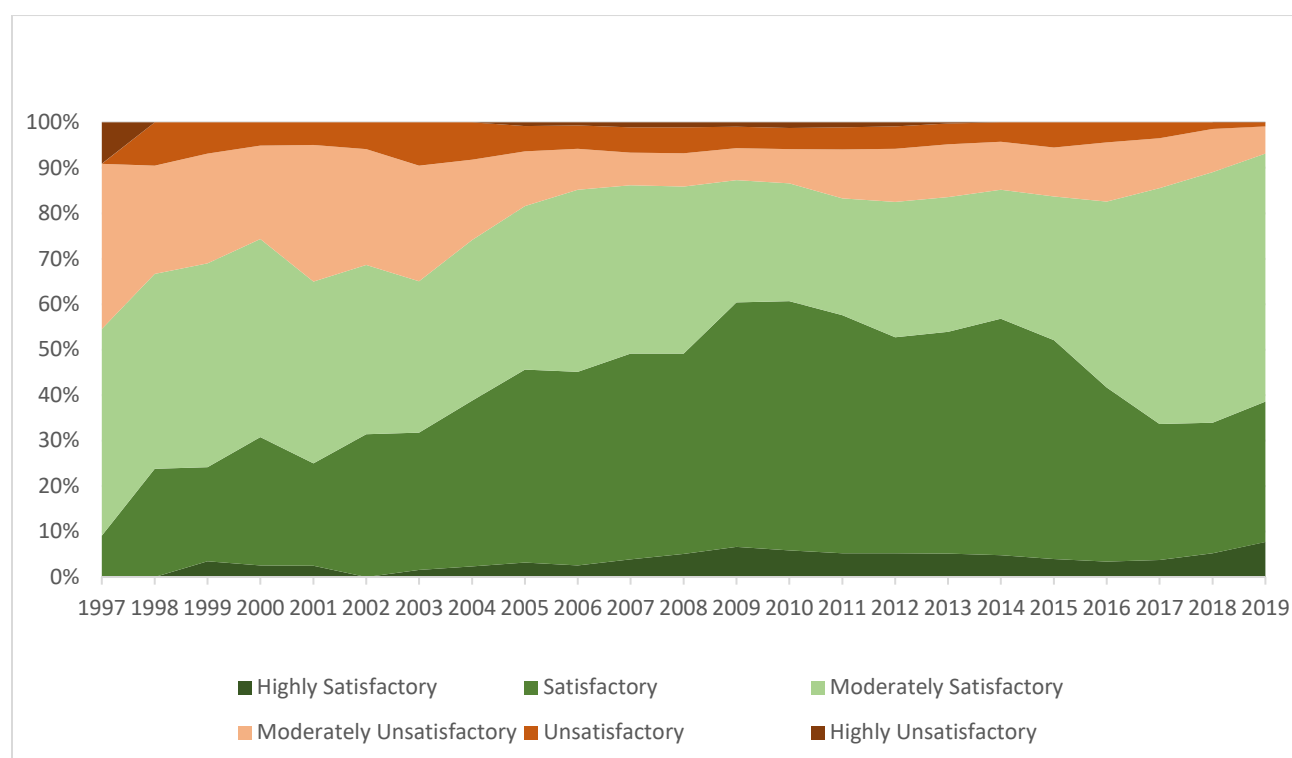
Findings

Quality of terminal evaluations

63 **The quality of terminal evaluation reports improved rapidly from 1997 to 2009 but the gains have slowed thereafter.** Up to 2004, approximately two-thirds of the terminal evaluations were rated in the unsatisfactory range (figure 5). APR 2004 reported concerns

related to the quality of terminal evaluations submitted by UNDP and UNEP and recommended that the two Agencies establish terminal evaluation review processes for GEF projects to improve the quality of their reports. This, along with measures such as the issuance of, and consultations for, the “GEF Monitoring and Evaluation Policy” (GEF IEO 2006) and the “Guidelines for GEF Agencies in Conducting Terminal Evaluations” (GEF IEO 2008), contributed to an improvement in the quality of terminal evaluations (GEF IEO 2008—APR 2007). Up to 2009, almost all the terminal evaluations submitted (99 percent) were for projects that were implemented by the World Bank, UNDP, and UNEP. From 2010 onwards, at least one out of six terminal evaluations has been submitted by the Agencies that gained direct access as part of the first or second round of expansion of the GEF Partnership. During this period, although the number of terminal evaluations rated in the satisfactory range for quality has been high, the share of terminal evaluations that are rated satisfactory or higher has declined.

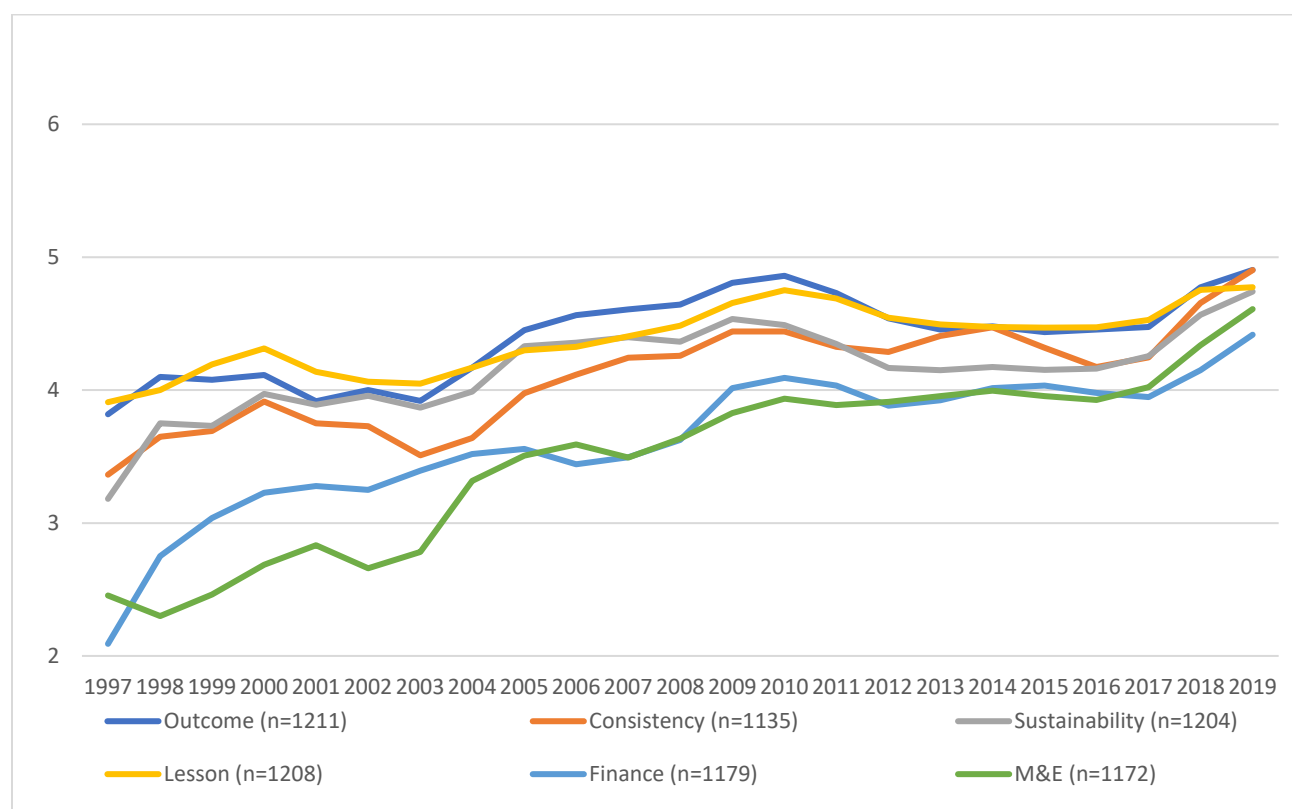
Figure 5: Three-year rolling average of terminal evaluation quality ratings presented—based on year of terminal evaluation completion (N=1,677)



64 **The ratings for quality of terminal evaluations have improved on all six dimensions, and dimensions where a lower percentage of the evaluations were rated in the satisfactory range at baseline have shown greater progress.** The analysis on quality dimensions considered the ratings on a six-point scale for different dimensions of terminal evaluation quality (figure 6). The analysis shows that terminal evaluations on average rate higher on dimensions, such as reporting on project outcome, sustainability, consistency, and lessons, compared with reporting on utilization of finance and project M&E. The difference in performance was more salient for terminal evaluations that were prepared before 2010. Thereafter, there has been convergence in ratings on the quality dimensions. The terminal evaluations that have been prepared from

2010 onwards have higher ratings for reporting consistency, coverage of M&E, and financing, compared with those prepared up to 2009—the difference on these dimensions is statistically significant after controlling for project size and Agency type.

Figure 6: Three-year rolling average on terminal evaluation quality dimensions—by year of terminal evaluation completion



65 **The net disconnect between GEFI IEO ratings and Agency ratings for terminal evaluation quality is only one percent, but 13 percent of the terminal evaluations have different ratings.** There are 66 terminal evaluations that have been rated for quality by both the GEF IEO and Agency evaluation offices on a six-point scale. Terminal evaluations prepared by the World Bank (33), UNDP (14), UNEP (10), UNIDO (six), joint projects (two), and ADB (one) are represented in this group. The rating scale is ordinal, but it may be treated as an index to facilitate comparison of average ratings. For these 66 terminal evaluations, the average overall quality rating provided by the GEF IEO and Agency evaluation offices is the same at 4.4. Eighty-six percent of terminal evaluations (57 out of 66) were rated the same on a binary scale by both GEF IEO and the Agency evaluation offices—whereas the remainder (nine terminal evaluations) were rated differently. However, the net disconnect is negligible because four of the terminal evaluations that were rated in the satisfactory range by the GEF IEO were rated in the unsatisfactory range by the Agency evaluation office, whereas five of the terminal evaluations rated in the unsatisfactory range by the GEF IEO were rated in the satisfactory range by the latter. Overall, GEF IEO rated 83 percent in the satisfactory range for quality, whereas the Agency evaluation offices rated 82 percent in this range.

66 **The quality of terminal evaluations prepared by UNDP and UNEP does not vary based on project size, whereas terminal evaluations of full-size projects prepared by the World Bank are generally of higher quality than for medium-size projects.** For UNDP, of the terminal evaluations prepared from 2010 onwards, 84 percent of those for full-size projects and 83 percent of those for medium-size projects rate in the satisfactory range for quality (table 4). Similarly, for UNEP, the percentage of terminal evaluations from the recent period rated in the satisfactory range are 97 percent for both types of projects. In contrast, of the terminal evaluations prepared by the World Bank there is substantial difference in the percentage rated in the satisfactory range for quality by project type: 88 percent for the terminal evaluations for full-size projects prepared from 2010 onwards versus 66 percent for the evaluations for medium-size projects prepared during the same period. For other Agencies, the number of observations is insufficient to assess difference in terminal evaluation quality ratings for full- and medium-size projects.

67 **The difference in quality of terminal evaluations for full-size and medium-size projects prepared by the World Bank may be a result of the difference in level of effort invested.** The team leaders of the World Bank–implemented projects prepare a detailed terminal evaluation—called implementation completion report—for a full-size project that is then reviewed by the IEG of the World Bank. However, for a medium-size project the project team leaders usually prepare a brief implementation completion memorandum. These memorandums generally lack the level of details that the implementation completion reports have. Furthermore, unlike implementation completion reports, these memorandums are not reviewed by the IEG, which may reduce the incentive for higher quality reporting.

Table 4: Percentage of terminal evaluation reports rated in the satisfactory range

GEF Agency	Full-Size Projects			Medium-Size Projects		
	Through 2009	2010 Onwards	All Years	Through 2009	2010 Onwards	All Years
ADB	—	91 (11)	91 (11)	0 (3)	—	0 (3)
FAO	100 (1)	94 (18)	95 (19)	—	83 (6)	83 (6)
IDB	—	75 (16)	75 (16)	—	67 (3)	67 (3)
IFAD	—	80 (20)	80 (20)	—	75 (4)	75 (4)
UNDP	80 (116)	84 (346)	83 (462)	83 (71)	83 (236)	83 (307)
UNEP	81 (21)	97 (66)	93 (87)	84 (43)	97 (73)	92 (116)
UNIDO	100 (1)	95 (21)	95 (22)		83 (30)	83 (30)
World Bank	88 (137)	88 (248)	88 (385)	76 (78)	66 (44)	72 (122)
Joint/others	75 (12)	91 (46)	88 (58)	100 (1)	100 (7)	100 (8)
All Agencies	84 (288)	87 (792)	86 (1080)	79 (196)	84 (403)	82 (599)

Note: Number of observations in parentheses

Compliance with the GEF IEO terminal evaluation guidelines

68 Compared with other Agencies, the terminal evaluations for full-size projects prepared by UNEP and UNIDO generally conform to the terminal evaluation guidelines but they also have some areas of low compliance. **Error! Reference source not found.** Table 5 presents information on the level of overall compliance of terminal evaluations for full-size projects with the requirements by reporting dimension. Terminal evaluations for full-size projects prepared by UNEP and UNIDO generally comply with most of the requirements. However, a low percentage of terminal evaluations prepared by UNEP (64 percent) report on Agency performance and a low percentage of those prepared by UNIDO (29 percent) report on application of social and environmental safeguards. Among other Agencies, terminal evaluations prepared by the World Bank are strong in reporting on project M&E and consistency of the performance ratings provided in terminal evaluations but are weak in reporting on information sources and on sustainability. Terminal evaluations prepared by UNDP are strong in reporting on sustainability and generally provide performance ratings on all dimensions but are weak in reporting on project's theory of change and on application of social and environmental safeguards.

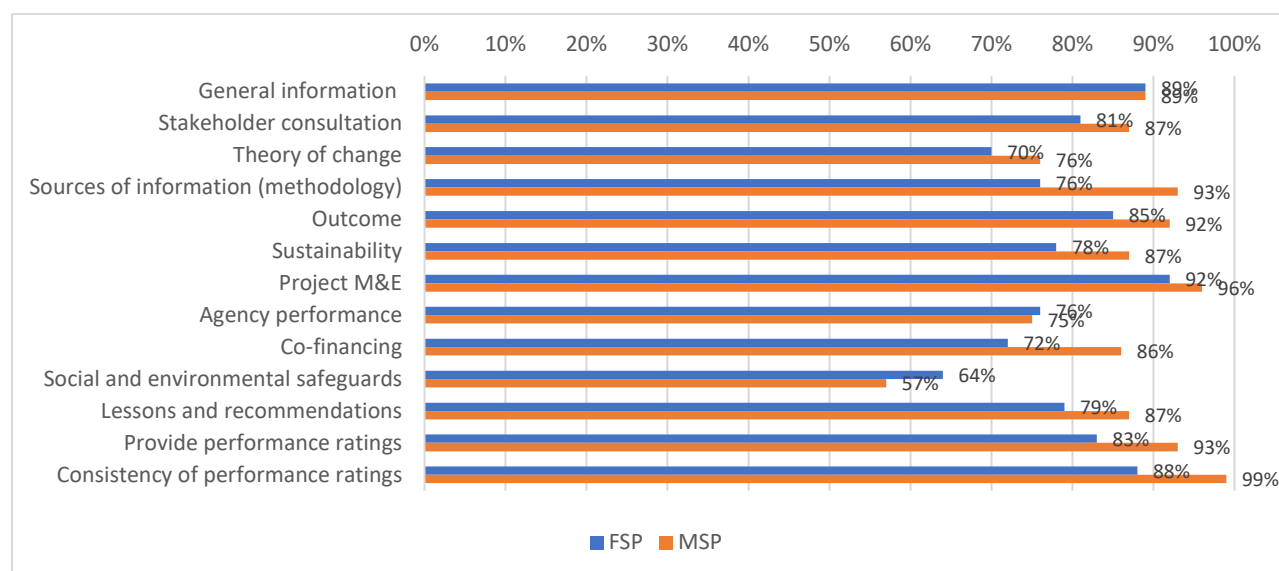
Table 5: Compliance on reporting requirements for terminal evaluations of full-size projects

Reporting dimensions	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
General information	85	93	93	88	86	89	89
Stakeholder consultation	78	89	100	100	72	44	81
Theory of change	67	30	98	83	82	69	70
Sources of information (methodology)	84	88	90	90	27	80	76
Outcome	77	81	94	87	87	84	85
Sustainability	85	91	92	89	57	71	78
Project M&E	84	86	96	92	100	78	92
Agency performance	64	79	64	88	80	78	76
Cofinancing	76	83	77	87	62	65	72
Social and environmental safeguards	67	49	71	29	87	53	64
Lessons and recommendations	72	80	88	81	75	79	79
Provide performance ratings	91	98	92	93	66	80	83
Consistency of performance ratings	82	86	95	85	96	74	88

69 **The variation in compliance may be a result of various factors which will need to be further examined.** The overall effort made by the GEF IEO to disseminate the guidelines may affect the extent to which the GEF Agency staff is aware of the requirements. Internal processes of the Agencies for conducting terminal evaluations may also explain variation. For Agencies such as UNEP and UNIDO, in which preparation of terms of reference for the terminal evaluations, selection of evaluators, and clearance to the finalized report are centralized, there may be greater consistency in terminal evaluation reports, and it may be easier to mainstream the GEF IEO guidance in practice. These two Agencies also mainstreamed the GEF IEO terminal evaluation guidelines in the generic terms of reference (UNEP) and evaluation manual (UNIDO) soon after issuance of the guidelines. On the other hand, for Agencies such as UNDP and the World Bank, in which terminal evaluations are prepared through a decentralized process, the extent to which the GEF terminal evaluation guidelines are implemented depends on whether the internal guidelines of these Agencies for preparation of terminal evaluations also reflect the GEF requirements. This review has not examined these factors in detail as these will be addressed in the ongoing GEF IEO evaluation of the Self Evaluation Systems of GEF Agencies.

70 **Terminal evaluations for medium-size projects meet the guidelines to a similar extent as full-size projects.** Compared with 80 full-size projects, terminal evaluations for 38 medium-size projects were covered in the review. The analysis shows that the terminal evaluations for medium-size projects match or nominally exceed the performance of the terminal evaluations for full-size projects (figure 7). Given that the findings are based on a subset of an annual cohort of terminal evaluations, it is difficult to say whether results for medium-sized projects are representative. More details on comparison of terminal evaluations for full-size and medium-size projects are provided in annex D.

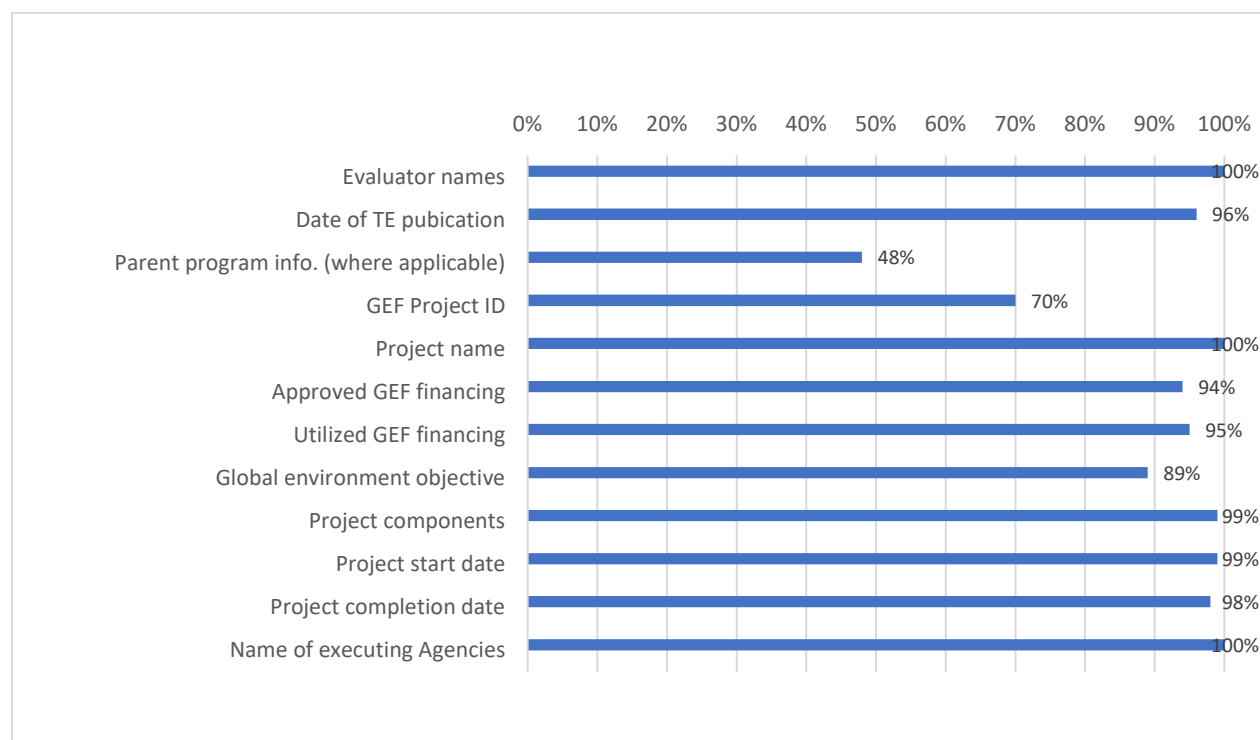
Figure 7: Percentage of terminal evaluations that comply with GEF IEO guidance—by project size



General Information

71 **There are considerable information gaps in the terminal evaluations for projects that are affiliated with programs.** All or almost all terminal evaluations provide project name, executing agency name, name of evaluators, project start and completion date, and project components (**Error! Reference source not found.**). However, there are significant gaps in information on parent programs of projects implemented within the framework of programmatic approaches and GEF ID of the project covered by the terminal evaluation. Most Agencies link the covered project with the respective GEF project ID, but only a few (16 percent) of the terminal evaluations for the World Bank–implemented projects do so. A few terminal evaluations prepared by FAO did not provide their publication date. Several terminal evaluations (11 percent) did not provide information on the global environment objectives pursued by the project, although these terminal evaluations are spread across the Agencies.

Figure 8: Terminal evaluations for full-size projects that provide information on basic project characteristics



Stakeholder Feedback

72 **Most terminal evaluations for full-size projects have incorporated stakeholder feedback.** The terminal evaluation guidelines require that evaluators seek inputs from the project stakeholders and address their inputs in the finalization of the report. Most terminal evaluations report having addressed the feedback of stakeholders: 81 percent for full-size projects and 87 percent for medium-size projects (table 6). Although the guidelines also call for sharing of the draft report of the terminal evaluation with the operational focal points (OFPs) and for addressing their comments in the final version of the report, the guidelines do not explicitly require that the terminal evaluations report whether or not this took place. Only 29

percent of the terminal evaluations report sharing the draft report of the evaluation with the OFP. However, it does not follow that this was due to lack of sharing of the draft report with the OFPs. The GEF IEO will follow up on this finding through its ongoing work on evaluation of the Agency Self-Evaluation Systems.

Table 6: Percentage of terminal evaluations for full-size projects that report how stakeholders were involved in conduct of evaluations

Terminal evaluation reports that...	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
Draft report of terminal evaluation was shared with OFP ⁷	17 (6)	39 (18)	25 (4)	40 (5)	26 (23)	17 (6)	29 (62)
OFP Feedback addressed in finalization	0 (6)	22 (18)	25 (4)	40 (5)	9 (23)	0 (6)	15 (62)
Feedback of other stakeholders addressed	78 (9)	89 (18)	100 (13)	100 (6)	72 (25)	44 (9)	81 (80)

Note: Number of observations in parentheses.

Theory of Change

73 Most terminal evaluations discuss the theory of change of the evaluated projects. The GEF IEO's "Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects." (GEF IEO 2017) requires that the terminal evaluations present a discussion of the project's theory of change. Eighty-three percent of the terminal evaluations for full-size projects discuss the evaluated project's theory of change (table 7). Most terminal evaluations also discuss causal links in the project's theory of change and key assumptions of the theory. Approximately a third do not discuss how causal links are supposed to work and nearly half do not report on whether the key assumptions remained valid during implementation. Among the Agencies, terminal evaluations prepared by UNEP are especially strong in addressing various aspects of the underlying project's theory of change. Those prepared by the World Bank and UNIDO also address most aspects related to project's theory of change.

Table 7: Percentage of terminal evaluations for full-size projects that address theory of change

Report discusses.....	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
Observations	9	18	13	6	25	9	80
Project's theory of change	89	44	100	100	96	78	83
Causal links of theory of change	89	39	100	100	96	78	81
How causal links are supposed to work	78	17	100	83	80	78	69
Assumptions of the theory of change	78	33	100	83	88	56	73
Whether assumptions remained valid	0	17	92	50	48	56	44

⁷ The questions related to sharing of draft report with OFPs and addressing their comments in finalization of the terminal evaluation reports was applied only in case of national projects. Stakeholder feedback related questions was applicable to all projects.

Sources of Information

74 Terminal evaluations for full-size projects prepared by most Agencies discuss their information sources. Transparent discussion about the sources of information of an evaluation adds to its credibility. Eighty-six percent of the terminal evaluations for full-size projects give an overview of the information sources used for the evaluation (table 8). Most of the terminal evaluations mention desk reviews, interviews, and field verification as their information sources. Among the Agencies, the terminal evaluations prepared by UNDP, UNEP, UNIDO, and FAO are strong in describing their data sources and in providing supporting details. In comparison, a lower percentage of terminal evaluations for World Bank–implemented full-size projects (60 percent) discuss their information sources. Furthermore, a lower percentage of their terminal evaluations for full-size projects discuss the who-, why-, and where-related questions for the interviews and field verifications.

Table 8: Percentage of terminal evaluations that discuss information sources—for full-size projects

Discusses/provides...	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
<i>Observations</i>	9	18	13	6	25	9	80
Discusses sources of information	100	100	100	100	60	89	86
<i>Desk review as data source</i>	100	100	100	100	48	89	83
Information on reviewed documents	100	100	100	100	92	100	98
<i>Interviews as a data source</i>	100	100	100	100	44	78	80
Info on who was covered	100	100	100	100	36	100	89
Info on interviewee selection	78	56	77	67	18	57	58
List of interviewees provided	89	100	100	100	0	71	78
<i>Field verification as a data source</i>	100	100	100	100	40	78	80
Info on who visited	79	94	85	100	0	63	72
Info on which sites were visited	79	100	92	67	10	100	78
Info on basis of site selection	44	56	62	83	0	63	50

**The italicized rows are for information that does not indicate terminal evaluation quality. The numbers below the italicized rows are a percentage of the italicized number.*

Reporting on Outcomes and Sustainability

75 Most terminal evaluations for full-size projects discuss relevance of outcomes to country and agency priorities, and relevance of project design for achieving expected outcomes; however, a discussion of relevance from the GEF perspective is missing in approximately a third. Among the Agencies, all terminal evaluations prepared by UNEP discuss project relevance from GEF, country, agency, and design perspectives (table 9). Terminal evaluations for World Bank–implemented full-size projects tend to be strong in discussing some aspects of relevance and weak in others. Although a high percentage of terminal evaluations

discuss relevance of the project from the recipient country and the World Bank perspective, a much lower percentage discuss a project's relevance to GEF priorities and relevance of its design to intended outcomes.

Table 9: Percentage of terminal evaluations that report on given aspect of outcome performance

Discusses	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
Observations	9	18	13	6	25	9	80
Relevance							
Relevance of outcome to GEF priorities	67	56	100	83	44	89	66
Relevance to country priorities	89	83	100	100	96	100	94
Relevance of outcome to Agency Priorities	78	61	100	100	92	78	84
Relevance of project design to outcome	89	83	100	83	64	89	81
Effectiveness							
Output target achievement at completion	67	78	100	100	92	100	88
Outcome achievement at completion	89	94	100	100	100	89	96
Performance on all outcome indicators	11	83	46	50	80	44	61
Discusses overall outcome achievement	89	100	100	100	96	89	98
Detailed discussion of factors that affect outcome achievement	78	72	85	67	96	33	78
Efficiency							
Timeliness of activities	89	94	100	100	92	100	95
Efficiency in use of resources	89	89	100	83	92	100	93
Factors that affect efficiency	89	83	92	83	96	100	91
Sustainability							
Risks that may affect sustainability	89	100	100	100	88	89	94
Likelihood of risk materialization	89	83	100	83	40	56	70
Impact of risk materialization	78	89	77	83	44	67	69

76 Most terminal evaluations for full-size projects provide a detailed account of project performance in terms of expected outputs and outcome, actual outputs and outcome, and performance on key outcome indicators. It is important to discuss whether a project achieved its expected outputs before starting to discuss whether the observed outcomes may be attributed to the project. Furthermore, information on actual outcome and performance on key outcome indicators is necessary to determine the extent to which the project made progress toward achieving its long-term objectives. Reporting on outcome effectiveness is strong, because 88 percent of the terminal evaluations report on output target achievement and 96

percent discuss outcome achievement at completion. Almost all terminal evaluations for full-size projects (98 percent) provide at least some account of the factors that affected outcome achievement. However, 20 percent of the evaluations discuss these factors cursorily.

77 Compliance with reporting requirements on efficiency is strong but reporting on sustainability varies by Agency. Most terminal evaluations for full-size projects provide an account of outcome efficiency, providing information on aspects such as timeliness of activities, efficiency in resource, and factors that affect efficiency. Terminal evaluations prepared by most Agencies also report on aspects related to sustainability with an exception of those prepared by the World Bank. This may reflect the World Bank IEG's switch from rating sustainability from 2008 onwards to rating of risk to development outcomes, which was also dropped from IEG ratings from July 2017 onwards (World Bank IEG 2017).

Monitoring, Implementation, and Cofinancing

78 Most terminal evaluations for full-size projects report on M&E design and implementation, and GEF Agency's performance, but there are gaps in reporting on executing agency performance and on cofinancing. Analysis of the information presented in the terminal evaluations for full-size projects shows that a high percentage of terminal evaluations discuss M&E design (89 percent) and M&E implementation (94 percent) (**Error! Reference source not found.**). Similarly, a high percentage discusses overall performance of the GEF Agency (88 percent) and provides information on materialized cofinancing (88 percent). There are, however, gaps in reporting on performance of executing agencies, which are responsible for activities such as procurement, implementation of physical activities, and gathering project-monitoring data. Forty percent of terminal evaluations do not provide an overview of executing agency performance; this gap in reporting is especially salient in World Bank-implemented projects. Although most terminal evaluations provide information on expected and materialized cofinancing, a substantial percentage does not provide information on who administered the cofinancing, and, where applicable, reasons for shortfall or excess materialization and their effect on project performance.

Table 10: Reporting on M&E, implementation, and cofinancing—for full-size projects

Reporting on....	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
Number of observations	9	18	13	6	25	9	80
M&E strengths-weaknesses							
Design	78	83	92	83	100	78	89
Implementation	89	89	100	100	100	78	94
GEF Agency performance							
Overall Agency performance	78	94	54	100	100	89	88
During difference stages of project cycle	44	56	69	83	100	78	75
Supervising executing agencies	56	67	71	100	100	78	81
Executing agency performance							
Overall performance	79	100	62	67	20	67	60
Cofinancing							
Expected cofinancing	89	89	100	100	80	89	89
Data consistent with project documents	89	72	69	83	40	67	64
Materialized cofinancing	89	89	100	100	76	89	88
Materialized cofinancing data is consistent	89	78	92	83	76	89	83
Whether in-cash or in-kind	89	78	85	100	40	78	70
Who administered cofinancing	33	78	69	83	56	11	58
Provides reasons for excess/shortfall*	100 (7)	100 (6)	50 (8)	75 (4)	73 (11)	57 (7)	74 (43)
Discusses effects of excess/shortfall*	29 (7)	83 (6)	50 (8)	50 (4)	55 (11)	43 (7)	51 (43)

Social and Environmental Safeguards

79 **The terminal evaluations prepared by the World Bank are strong in reporting on the application of social and environmental safeguards, whereas a sizable percentage of those prepared by other Agencies have gaps.** Consistent with the GEF “Policy on Environmental and Social Safeguards” (GEF 2019), the GEF terminal evaluation guidelines require reporting on application of the applicable safeguards by the GEF Agencies. Ninety-six percent of terminal evaluations for full-size projects prepared by the World Bank report on application of social and environmental safeguards during project preparation and 100 percent on application during implementation (table 11). The difference between the World Bank and other Agencies is both significant and substantial. This difference in reporting performance is explained by World Bank’s strong tradition of implementing at various stages of project cycle through its “do not harm” approach (GEF IEO 2008—APR 2006). A smaller percentage of terminal evaluations for the UNDP-implemented projects report on application of these safeguards. Almost all the

terminal evaluations prepared by UNEP, and most prepared by UNIDO, report on whether a gender analysis was conducted. When all Agencies are considered as a group, only two-thirds of terminal evaluations for full-size projects report on whether a gender analysis was conducted as part of the project. However, where applicable, a vast majority report on whether actions identified in the gender analysis were implemented.

Table 11: Percentage of terminal evaluations that report on social and environmental safeguards

Discusses whether....	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
Number of Observations	9	18	13	6	25	9	80
Safeguards applied during preparation	44	17	46	17	96	22	50
Safeguards applied during implementation	44	17	46	33	100	44	55
Reporting on conduct of gender analysis	78	61	92	67	60	44	66
Implementation of actions identified in gender analysis (obs. in parentheses)	100 (1)	100 (4)	100 (1)	0 (2)	92 (12)	100 (2)	86 (22)

**Where actual observations differ from the full set, they are noted in parentheses*

Lessons and Recommendations

80 Almost all the terminal evaluations for full-size projects discuss lessons from the project, and these lessons, for the most part, are based on the evidence presented in the terminal evaluations. Ninety-four percent of the terminal evaluations for full-size projects present lessons and 84 percent discuss applicability of the lessons (table 12). For 78 percent of the terminal evaluations, all lessons are based on the presented evidence. For the remainder, at least some of the lessons were difficult to link with the presented evidence. Almost all terminal evaluations present recommendations, specify the actions that need to be taken, and specify who needs to take the action. However, only three out of 10 specify the time frame for the proposed actions.

Table 12: Percentage that present lessons and recommendations—for full-size projects

	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
Number of observations	9	18	13	6	25	9	80
Lessons							
Lessons are presented	56	100	100	100	100	89	94
All lessons based on presented evidence	33	78	77	100	88	78	78
Explicit/implicit discussion of applicability	56	72	100	50	100	89	84
Recommendations							
Presents recommendations	100	100	100	100	80	100	94
Specifies what needs to be done	100	89	100	100	76	100	90

Specifies action taker for recommendations	100	83	100	100	80	67	87
Time frame for expected action	56	39	38	17	4	33	28

**Where actual observations differ from the full set, they are noted in parentheses*

Performance Ratings

81 A substantial percentage of terminal evaluations do not include ratings for execution and sustainability, and in 12 percent of instances the ratings are not supported by the evidence presented in the evaluations. The gaps in performance ratings for execution and sustainability are related to the Agency that prepared the terminal evaluations (table 13). For example, all Agencies, except the World Bank, rate sustainability. The World Bank, as discussed earlier in the chapter, stopped rating sustainability in 2017. Similarly, only a small percentage of terminal evaluations prepared by the World Bank (8 percent) rate performance of the executing agency. While a substantial percentage of terminal evaluations prepared by UNDP and UNEP also do not rate performance of the executing agency, the gap is not of the same magnitude as that for the World Bank. For other Agencies the number of observations are too few to generalize, but gaps in ratings are noted for them as well. The percentage of instances in which a performance rating provided in a terminal evaluation was assessed to be consistent with the evidence ranged from 84 percent for sustainability to 92 percent for outcomes. Overall, ratings for full-size projects were assessed to be consistent with evidence in 88 percent of instances.

Table 13: Percentage of terminal evaluations that provide performance ratings and their consistency

	FAO	UNDP	UNEP	UNIDO	World Bank	Others	All
Observations	9	18	13	6	25	9	80
Ratings are provided for							
Outcome	100	100	100	100	100	67	96
Sustainability	100	100	100	100	24	100	76
M&E	100	100	100	100	96	78	96
Implementation	78	94	92	100	100	89	93
Execution	78	94	69	67	8	67	56
Provided ratings are consistent with evidence							
Outcome	78	89	92	100	100	83	92
Sustainability	78	83	100	67	83	78	84
M&E	67	94	92	100	96	51	88
Implementation	86	88	91	83	100	75	91
Execution	100	76	100	75	100	83	87

Comparison of performance rating from different sources

82 Both PIRs and terminal evaluations provide a more optimistic appraisal of project performance through performance ratings than the terminal evaluation validations conducted by the GEF IEO. There is a greater disconnect in ratings between GEF IEO validations and PIRs, versus GEF IEO validations and terminal evaluations. Although PIRs do not rate project outcome, they do rate the likelihood of achievement of development objectives, which is a close proxy for project outcomes. By the time the last PIR of a project is prepared, there is greater clarity on achievement of the project's development objective and the evidence considered in preparing the PIR is comparable to that considered in preparation of the terminal evaluation. The development objective rating in the last PIR may, therefore, be compared with the outcome rating provided in the terminal evaluation and the validations conducted by the GEF IEO. To ensure comparability, analysis was restricted to projects for which the ratings from all the three sources are available. The comparison shows that whereas 84 percent of projects from all periods were rated in the satisfactory range for outcome by the GEF IEO, 96 percent are so rated by the last PIR, and 92 percent by the terminal evaluation (table 14). The trend remains similar even when we focus only on a smaller subset of more recent projects.

Table 14: Percentage of projects in the satisfactory range for outcome (or equivalent)- by rating source

Year of terminal evaluation completion Rating source	Through 2009 (n=269)	2010 Onwards (n=327)	All (n=596)
Last PIR submitted (Development Objectives)	97	95	96
Terminal evaluation (Outcome)	93	91	92
GEF IEO terminal evaluation validation (Outcome)	83	85	84

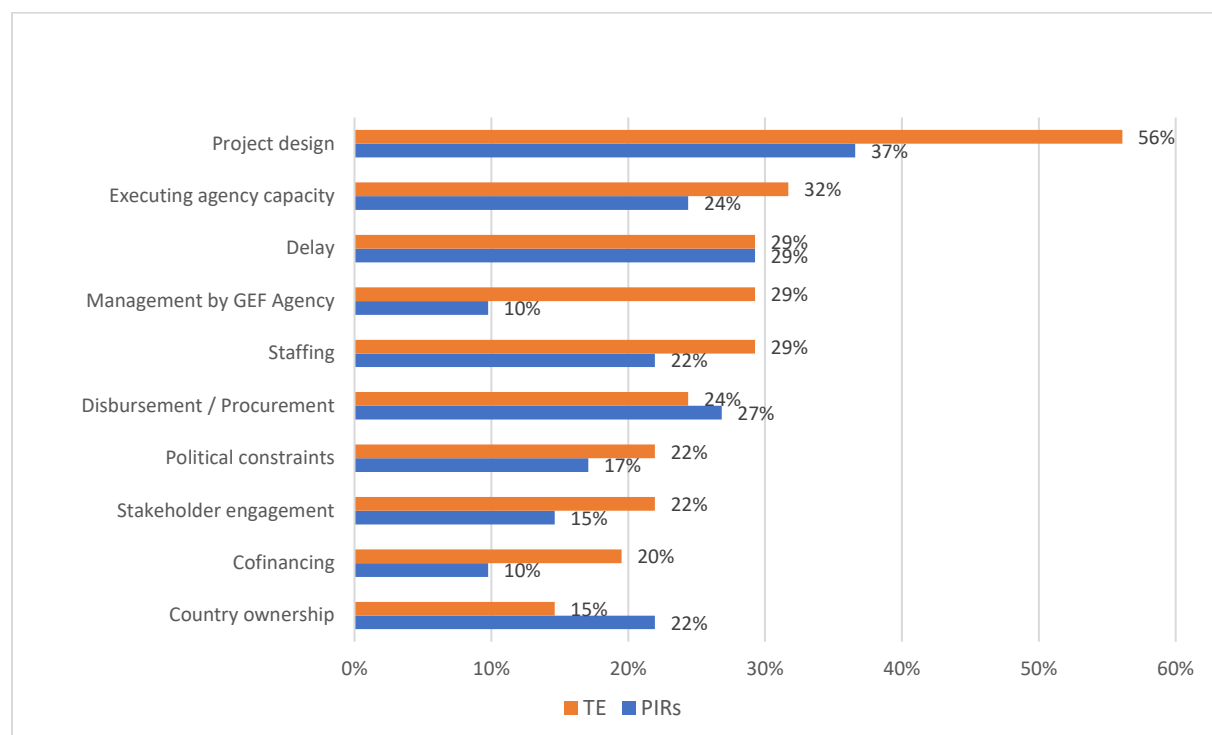
83 Unavailability of PIRs is associated with unsatisfactory or lower outcome rating, although the magnitude of effect is not substantial. Of the completed projects that were approved from GEF-3 onwards, 82 percent of the projects that had PIRs were rated in the satisfactory range for outcomes, whereas of those without a PIR only 70 percent were so rated. This difference is statistically significant. Among the projects that were rated unsatisfactory or lower, PIRs were not available for 12 percent of projects, whereas PIRs were not available for 6 percent of those rated moderately unsatisfactory or higher.

Reporting on projects rated by the GEF IEO as unsatisfactory or lower for outcome

84 PIRs raise early flags on project performance. For most projects that ultimately received unsatisfactory or highly unsatisfactory ratings, PIRs did note issues and risk factors that might jeopardize the achievement of planned outcomes. For any given project, all PIRs together were treated as an information source. Compared with PIRs, terminal evaluations are more likely to frame concerns that affect project performance as rooted in project design or

management by GEF agency (figure 9).⁸ Both PIRs and terminal evaluations are equally likely to discuss concerns such as delays and difficulties in disbursement and procurement. When concerns related to project design, delay, or management by the GEF Agency are raised in the PIRs, more than half of the times it is also picked up by the terminal evaluations. Which shows that several of these issues continued to be relevant in terminal evaluations, which are prepared with an added benefit of hindsight.

Figure 9: Incidence of discussion of concerns that affect project performance—by information source (n=41 projects)



85 There are instances in which the PIRs either do not adequately raise concerns that affect performance or raise it too late. The terminal evaluation for the World Bank–implemented project SIP: Scaling up SLM Practice, Knowledge, and Coordination in Key Nigerian States (GEF ID 3384) discusses concerns related to lack or realism in project design—too few resources for an ambitious agenda and lack of analysis to guide resource use—but these concerns were not raised in the two PIRs or in the midterm review of the project. The terminal evaluation for UNDP-implemented Capacity Development for Improved National and International Environmental Management in Seychelles notes concerns such as frequent staffing changes and lack of back stopping from the UNDP country office, which led to delayed implementation as some of the activities depended on timely completion of other activities.

⁸ Of the 50 projects approved from GEF-3 onward that were included in the review, for 44 PIRs were available. In two other cases – both World Bank implemented medium-size projects – were dropped as same document had been submitted as a PIR and as a terminal evaluation, and for an IFAD implemented project all PIRs were in Spanish. These were excluded from the review, and documents for the remaining 41 projects were reviewed to assess consistency.

These concerns are noted in the PIRs when the project was close to completion. Similarly, terminal evaluation of the UNDP-implemented project ESCO Moldova - Transforming the market for Urban Energy Efficiency in Moldova by Introducing Energy Service Companies (ESCO) (GEF ID 5157) notes that reporting leading up to the midterm did not raise any concerns that would affect performance. Although PIRs after the midterm review did note several concerns related to implementation, temporally these were reported later than what one would expect based on when these concerns were first observed.

Reporting on canceled projects

86 Agencies are required to implement projects that have been approved or endorsed by the GEF CEO. However, there may be situations that may warrant cancellation of a project – the GEF “Project Cancellation” policy (2018) explains how a project may be canceled. After CEO Endorsement/Approval, the decision to cancel rests with the respective Agency, which is required to inform the GEF and other stakeholders of the cancellation. There is no explicit requirement that Agencies provide a reason for cancellation. However, it is important that these reasons are known to facilitate learning. Of the projects that are canceled after CEO Endorsement/Approval, those that are canceled at least after any year of start of implementation are especially important because for these PIRs may be available and consistency in reporting through the cancellation memorandum and PIRs may be assessed. The analysis on reporting on canceled projects assesses the extent to which relevant information is provided and where PIRs are available, whether these report on the reasons that eventually led to cancellation.

87 **Of the 32 canceled projects reviewed, the available documents for 94 percent (30 projects) provide reasons for project cancellation.** Of the 32 projects that were canceled after CEO endorsement/approval, 14 were cancelled before implementation began and the remainder (18 projects) were canceled after start of implementation. For the 14 projects that were canceled before the start of implementation, a cancellation memorandum was available in all but two cases. In these two cases the project status is reported as canceled, although no formal record of cancellation is available at the GEF PMIS/Portal. Of the two, one project was prepared by ADB (Central Asian Countries Initiative for Land Management [CACILM]: Southern Agricultural Area Development Project; GEF ID 3233) and the other by UNDP (Andaman and Nicobar Islands: Ecologically-Sustainable Island Development; GEF ID 1143). An explanation for cancellation is available for all 18 projects that were canceled after start of implementation—in 16 instances this was through PIRs and in the remaining two, in which projects were canceled after implementation start but before submission of the first PIR, through a cancellation memorandum.

88 **For almost all projects that were canceled after implementation start, either PIRs or cancellation memos or both appropriately discuss concerns that ultimately led to cancellation.** Of the 18 projects canceled after implementation, PIRs are available for 16 and cancellation memorandum for the remaining two. These canceled projects were under implementation for an average of 3.6 years, with a range of 1.5 years to 8.3 years. Of the 18 projects, reasons for cancellation have been discussed for 17, and they vary from project to project. Some pertain to weaknesses in project preparation and project design. The World

Bank–implemented International Lighting Efficiency Facility project was canceled because the nongrant funding model was subsequently found to be unsuitable for the project (GEF ID 6980, World Bank). Another World Bank–implemented project, the Environmental Infrastructure Project in the Black Sea-Danube River Basin (GEF ID 1542, World Bank) was canceled because of opposition of the local communities to construction of wetland and issues around the legal status of land plots assigned to construct project facilities. The UNDP–implemented Electric 3-Wheeler Market Launch Phase project in India (GEF ID 2395) was canceled because the changed market conditions led to the loss of the project partner’s interest. In several other cases the projects were canceled due to exogenous political shocks. In one case, the implementing agency, UNDP, lost formal representation in Nicaragua (GEF ID 5306, UNDP) where the project was being implemented, and suspended all activities. Two projects—in Syria and Yemen—were canceled due to deterioration in the security situations in country. The one case in which no performance concerns have been flagged in PIRs is the medium-size project ASTUD: Regional Knowledge Sharing (GEF ID 5527, ADB), for which three PIRs are available, with the most recent rating the project as moderately satisfactory. Although no cancellation memo is available, according to the GEF Trustee dataset the project was canceled on June 30, 2017.⁹

89 For most canceled projects, more than one PIR is available and concerns that led to cancellation are often discussed in the first PIR itself. Thus, in most instances PIRs do report major risks to projects in a timely manner. The concerns that are discussed include issues related to a slow startup, delays in fund disbursement, or stumbling blocks in moving forward due to difficulties in meeting preliminary conditions for the start of activities. For example, for the UNEP–implemented Integrated Approach for Zero Emission Project Development in the New Town of Boughzoul project in Algeria (GEF ID 3927), the first PIR notes a lack of progress as project activities were contingent on the construction of a new city by the government. The project was eventually canceled because of lack of progress in implementation of the baseline activities. The Geothermal Development Project in Yemen (GEF ID 3474, UNEP) experienced significant delays in approval of the contracts for drilling to prove geothermic resources. These concerns were noted in the first PIR, and along with the uncertainty due to the Arab spring that delayed legal and regulatory work, are outlined as the reasons for cancellation in the cancellation memorandum. The World Bank–implemented The Poison Dart Frog Ranching to Protect Rainforest and Alleviate Poverty project (GEF ID 1485) was canceled three years after start of implementation. The first PIR notes project delays due to problems acquiring forest concession, and the slow process in receiving permits and approvals from the National Institute of Natural Resources. The second PIR notes the same concerns and observes that the initial implementation arrangements were ineffective and that the relationship between the two main executing partners has deteriorated. The cancellation memo details the dysfunctional working relationship between the project partners that brought implementation to a standstill.

90 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. The PIRs for UNDP–implemented projects are prepared by the project managers but the country officers and regional technical

⁹ GEF Trustee dataset, 9.19.19

advisors—who are involved in the project in a supervisory and/or advisory role—also rate performance and provide commentary. A comparison of their ratings and accompanying justification shows that those in the supervisory and advisory role tend to be more candid in their reporting. For example, in the case of the Electric 3-Wheeler Market Launch Phase project (GEF ID 2395), the national project manager rates project implementation as moderately satisfactory in the first PIR, whereas the UNDP country office and regional technical advisor rate it as unsatisfactory and moderately unsatisfactory, respectively. The PIR notes a delay in signing the agreement for technology transfer, as well as additional finance requirements by customers due to higher price of the product introduced by project compared with gas vehicles. By the second PIR the UNDP regional technical advisor has downgraded project rating to highly unsatisfactory, as the project had not yet produced a workplan or spent any of the GEF budget (UNDP country office and national project manager ratings were the same as the previous year). Although low performance rating does not necessarily equate with candor, in instances where the project was eventually canceled low performance ratings are more realistic.

91 In a few cases cancellation memorandums provide additional information or explanation that was not provided in the PIRs. For example, for The Low-Carbon and Efficient National Freight Logistics Initiative Project (GEF ID 5842, Inter-American Development Bank), two PIRs are available. The first PIR (fiscal year 2016) notes delays caused by a workers strike and a change in the Minister of Transportation—a key stakeholder of the project focused on the freight transport sector. Despite delays and being behind scheduled execution of the 2016 budget, the project is rated as satisfactory and implementation progress is rated as moderately satisfactory in this PIR. The 2017 PIR rates the project as unsatisfactory overall and implementation progress as moderately unsatisfactory, and notes the strike, lack of political will, and low execution of budget as difficulties that may warrant cancellation of the project. A section on management actions taken to improve progress notes strengthening of the execution team, and a proposal to change the executing agency, but also explains that there is a lack of political will to perform the technical cooperation components of the project. The cancellation memo notes all of these issues and elaborates that “during 2015, the Ministry of Transport (MT) did not carry out the necessary procedures for the incorporation of the TC (technical cooperation) resources into its budget”, leading to an initial delay not noted in the PIRs.

REFERENCES

Global Environment Facility (GEF). 2019. "Policy on Environmental and Social Safeguards." SD/PL/03. Available at:

https://www.thegef.org/sites/default/files/documents/gef_environmental_social_safeguards_policy.pdf

Global Environment Facility (GEF). 2018. "Project Cancellation." Policy: OP/PL/02. Available at:

https://www.thegef.org/sites/default/files/documents/Project_Cancellation_Policy_20181220.pdf

GEF Independent Evaluation Office (IEO). 2005. "Procedures and Format of the GEF Management Action Record." GEF/ME/C.27/3. Available at:

https://www.thegef.org/sites/default/files/council-meeting-documents/C.27.ME_3_Management_Action_Record_4.pdf

_____. 2006. "The GEF Monitoring and Evaluation Policy." Evaluation Document No. 1.

Available at: <https://www.gefio.org/sites/default/files/ieo/evaluations/gef-me-policy-2006-eng.pdf>

_____. 2008. "Guidelines for GEF Agencies in Conducting Terminal Evaluations." Evaluation Document No. 3. Available at:

<https://www.gefio.org/sites/default/files/ieo/evaluations/te-guidelines-2008.pdf>

_____. 2008. "Annual Performance Report 2006." Evaluation Document No. 38. Available at:

<https://www.gefio.org/sites/default/files/ieo/evaluations/apr-2006.pdf>

_____. 2010. "OPS4 Progress Toward Impact." GEF IEO Document. Available at:

https://www.gefio.org/sites/default/files/ieo/evaluations/ops4_0.pdf

_____. 2009. "Annual Performance Report 2008." Evaluation Document No. 49. Available at:

<https://www.gefio.org/evaluations/annual-performance-report-apr-2008>

_____. 2017. "Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects." Available at: <https://www.gefio.org/sites/default/files/ieo/evaluations/files/gef-guidelines-te-fsp-2017.pdf>

World Bank Independent Evaluation Group (IEG). 2017. "Draft Guidelines for Reviewing World Bank Implementation Completion and Results Reports: A Manual for IEG ICR Reviewers."

Available at: https://ieg.worldbankgroup.org/sites/default/files/Data/More-on-IEG/icrr_evaluatormanual_august2018.pdf

ANNEX A: APR 2020 PROJECT COHORT

GEF ID	Project Title	Agency	Country	GEF Period	Outcome Rating
793	Program for the Management of Forests and Adjacent Lands	WB	Benin	GEF - 2	MS
1071	Rural Power	WB	Philippines	GEF - 2	MU
1609	Renewable Energy Enterprise Development - Seed Capital Access Facility	UNEP	Global	GEF - 3	S
1909	Protection of the Canary Current Large Marine Ecosystem (LME)	FAO	Regional	GEF - 4	MS
2096	Development of a National Implementation Plan in Brazil as a First Step to Implement the Stockholm Convention on Persistent Organic Pollutants (POPs)	UNEP	Brazil	GEF - 3	HS
2143	DBSB Water Quality Protection Project - under WB-GEF Strategic Partnership for Nutrient Reduction in the Danube River and Black Sea	WB	Bosnia-Herzegovina	GEF - 3	MU
2368	Hanoi Urban Transport Development	WB	Vietnam	GEF - 3	MS
2444	Biodiversity Conservation and Rural Livelihoods Improvement	WB	India	GEF - 3	MU
2489	Rural Infrastructure (Electrification Sector)	WB	Honduras	GEF - 3	MS
2625	Energy Efficiency Project	WB	Argentina	GEF - 3	MS
2706	Implementing Integrated Water Resource and Wastewater Management in Atlantic and Indian Ocean SIDS	UNEP	Regional	GEF - 4	MS
2759	WB/GEF POL: Manila Third Sewerage Project (MTSP) - under WB/GEF Partnership Investment Fund for Pollution Reduction in the LME of East Asia	WB	Philippines	GEF - 3	MU
2765	Espirito Santo Biodiversity and Watershed Conservation and Restoration Project	WB	Brazil	GEF - 4	MU
2770	Demonstration of a Regional Approach to Environmentally Sound Management of PCB Liquid Wastes and Transformers and Capacitors Containing PCBs	UNEP	Regional	GEF - 4	S
2776	Sustainable Transport	UNDP	Egypt	GEF - 3	MU
2913	Human-Wildlife-Coexistence Management Project in Northern Botswana.	WB	Botswana	GEF - 3	MU
2918	Sustainable Energy Development Project (SEDP)	WB	Rwanda	GEF - 3	S
2926	Environmentally Sound Management and Disposal of Obsolete POPs Pesticides and Other POPs Wastes	UNIDO	China	GEF - 4	HS
2946	Coal Fired Generation Rehabilitation Project	WB	India	GEF - 3	MS
2967	BS Regional Project for Implementing National Biosafety Frameworks in the Caribbean Sub-region - under the GEF Biosafety Program	UNEP	Regional	GEF - 4	MS
3077	Greening the Cocoa Industry	UNEP	Global	GEF - 4	S
3266	Management of Chimborazo's Natural Resources	FAO	Ecuador	GEF - 4	MS

GEF ID	Project Title	Agency	Country	GEF Period	Outcome Rating
3267	MENARID: Adaptation to Climate Change Using Agrobiodiversity Resources in the Rainfed Highlands of Yemen	WB	Yemen	GEF - 4	MS
3302	Climate Adaptation for Rural Livelihoods and Agriculture (CARLA)	AfDB	Malawi	GEF - 4	MS
3332	BS Contributing to the Safe use of Biotechnology	UNEP	El Salvador	GEF - 4	S
3367	SIP: Community-Based Integrated Natural Resources Management in Lake Tana Watershed	IFAD	Ethiopia	GEF - 4	MS
3399	SIP: Lake Victoria Environmental Management Project II	WB	Regional	GEF - 4	MU
3461	Promoting Sustainable Transport Solutions for East Africa	UNEP	Regional	GEF - 4	S
3535	Creating Markets for Renewable Power in Ukraine	EBRD	Ukraine	GEF - 4	S
3537	Mexico Rural Development	WB	Mexico	GEF - 4	S
3541	TT-Pilot (GEF 4): Phase Out HCFCs and Promotion of HFC-free Energy Efficient Refrigeration and Air-Conditioning Systems in the Russian Federation Through Technology Transfer	UNIDO	Global	GEF - 4	S
3589	CTI Coastal and Marine Resources Management in the Coral Triangle: Southeast Asia under Coral Triangle Initiative	ADB	Regional	GEF - 4	S
3591	PAS: Strengthening Coastal and Marine Resources Management in the Coral Triangle of the Pacific - under the Pacific Alliance for Sustainability Program	ADB	Regional	GEF - 4	S
3622	Integrated POPs Management Project: Dioxins and Furans, PCB and Contaminated Sites Management	WB	Philippines	GEF - 4	HU
3669	MENARID: Second Natural Resources Management Project	WB	Tunisia	GEF - 4	MS
3750	CBSP Catalyzing Sustainable Forest Management in the Lake Tele-Lake Tumba (LTLT) Transboundary Wetland Landscape	UNDP	Regional	GEF - 4	MS
3751	BS Capacity Building on Biosafety for Implementation of the Cartagena Protocol - Phase II under the Biosafety Program	UNEP	India	GEF - 4	MS
3777	CBSP Sustainable Management of the Wildlife and Bushmeat Sector in Central Africa	FAO	Regional	GEF - 4	U
3779	CBSP Enhancing Institutional Capacities on REDD issues for Sustainable Forest Management in the Congo Basin	WB	Regional	GEF - 4	MS
3786	CF: Industrial Energy Efficiency	UNIDO	Thailand	GEF - 4	S
3788	LGGE Promoting Energy Efficiency in Buildings in Eastern Africa	UNEP	Regional	GEF - 4	MS
3813	Integrating Trade-offs between Supply of Ecosystem Services and Land Use Options into Poverty Alleviation Efforts and Development Planning	UNEP	Mexico	GEF - 4	HS
3814	Best Practices for PCB Management in the Mining Sector of South America	UNEP	Regional	GEF - 4	MS

GEF ID	Project Title	Agency	Country	GEF Period	Outcome Rating
3822	CBSP - A Regional Focus on Sustainable Timber Management in the Congo Basin	UNEP	Regional	GEF - 4	U
3832	Improving the Energy Efficiency of Lighting and Building Appliances	UNDP	Egypt	GEF - 4	S
3848	Integrated Management of the Ilha Grande Bay Ecosystem	FAO	Brazil	GEF - 4	MS
3917	Improving Energy Efficiency and Promoting Renewable Energy in the Agro-Food and other Small and Medium Enterprises (SMEs) in Ukraine	UNIDO	Ukraine	GEF - 4	S
3963	Social Integral Development and its Interrelation with Climate Change in Watersheds in Lara and Falcon States (Venezuela) (PDELAFA)	IFAD	Venezuela	GEF - 4	U
3981	Integrated Management in Lakes Apanas and Asturias Watershed	IADB	Nicaragua	GEF - 4	S
3985	Demonstration Project for Decontamination of POPs Contaminated Soils Using Non-thermal Treatment Methods	FAO	Botswana	GEF - 4	MS
3986	Disposal of POPs Wastes and Obsolete Pesticides	FAO	Mozambique	GEF - 4	MU
3987	Eritrea: Prevention and Disposal of POPs and Obsolete Pesticides	FAO	Eritrea	GEF - 4	MS
4004	Mini-Grids Based on Small Hydropower Sources to Augment Rural Electrification	UNIDO	Tanzania	GEF - 4	S
4018	Sao Tome and Principe Adaptation to Climate Change	WB	Sao Tome and Principe	GEF - 4	S
4019	Strengthening Resilience and Adaptive Capacity to Climate Change in Guinea-Bissau's Agrarian and Water Sectors	UNDP	Guinea-Bissau	GEF - 4	MS
4036	TT-Pilot (GEF-4) DHRS: Irrigation Technology Pilot Project to face Climate Change Impact	IFAD	Jordan	GEF - 4	MU
4037	TT-Pilot (GEF-4): Overcoming Policy, Market and Technological Barriers to Support Technological Innovation and South-South Technology Transfer: The Pilot Case of Ethanol Production from Cassava	UNIDO	Thailand	GEF - 4	S
4068	Increasing Resilience to Climate Variability and Hazards	WB	Kiribati	GEF - 4	S
4084	CBSP Conservation and Sustainable Use of the Ngoyla Mintom Forest	WB	Cameroon	GEF - 4	MS
4096	Promoting Sustainable Biomass Energy Production and Modern Bio-Energy Technologies	UNDP	Sri Lanka	GEF - 4	S
4105	SPWA-BD: Wetlands Conservation Project	WB	Sierra Leone	GEF - 4	MS
4113	Mainstreaming Biodiversity in Palm Cropping in Colombia with an Ecosystem Approach	IADB	Colombia	GEF - 4	MS
4116	Lighting and Appliances Efficiency Project	WB	Mexico	GEF - 4	S
4173	Lighting Market Transformation in Peru	UNEP	Peru	GEF - 4	MS

GEF ID	Project Title	Agency	Country	GEF Period	Outcome Rating
4198	MED: Integrated Coastal Zone Management-Mediterranean Coast	WB	Morocco	GEF - 4	MS
4332	Sustainable Land and Forest Management in the Greater Caucasus Landscape	UNDP	Azerbaijan	GEF - 5	MU
4336	Lighting One Million Lives in Liberia	WB	Liberia	GEF - 5	S
4340	Strategic Planning and Action to Strengthen Climate Resilience of Rural Communities in Nusa Tenggara Timor Province (SPARC)	UNDP	Indonesia	GEF - 5	S
4352	Environmental Land Management and Rural Livelihoods	WB	Tajikistan	GEF - 5	S
4412	Establishing the tools and methods to include the nine new POPs into the Global Monitoring Plan	UNEP	Global	GEF - 5	S
4470	Building a Multiple-Use Forest Management Framework to Conserve Biodiversity in the Caspian Hyrcanian Forest Landscape	UNDP	Iran	GEF - 5	S
4492	Adaptation of Nicaragua's Water Supplies to Climate Change	WB	Nicaragua	GEF - 5	MS
4500	GEF Large-City Congestion and Carbon Reduction Project	WB	China	GEF - 5	HS
4517	Reducing Barriers to Accelerate the Development of Biomass Markets in Serbia	UNDP	Serbia	GEF - 5	S
4528	West Africa Regional Fisheries Program in Ghana	WB	Ghana	GEF - 4	U
4551	Community Based Flood and Glacial Lake Outburst Risk Reduction	UNDP	Nepal	GEF - 5	MS
4616	Climate Change Adaptation to Reduce Land Degradation in Fragile Micro-Watersheds Located in the Municipalities of Texistepeque and Candelaria de la Frontera	FAO	El Salvador	GEF - 5	MS
4618	Access to and Benefit Sharing and Protection of Traditional Knowledge to Promote Biodiversity Conservation and Sustainable Use	UNEP	Guatemala	GEF - 5	MU
4631	Watershed Approach to Sustainable Coffee Production in Burundi	WB	Burundi	GEF - 5	S
4642	Sustainable Agriculture and Climate Change Mitigation Project	WB	Uzbekistan	GEF - 5	MS
4653	CBPF-MSL: Strengthening the Management Effectiveness of the Protected Area Landscape in Altai Mountains and Wetlands	UNDP	China	GEF - 5	MS
4655	CBPF-MSL: Strengthening the Management Effectiveness of the Sub-system of Wetland Protected Areas for Conservation of Globally Significant Biodiversity	UNDP	China	GEF - 5	MS
4701	Scaling up Community-Based Adaptation (CBA) in Niger	UNDP	Niger	GEF - 5	S
4716	Conservation and Sustainable Use of Biodiversity in Coastal and Marine Protected Areas (MPAs)	UNDP	Guatemala	GEF - 5	MS

GEF ID	Project Title	Agency	Country	GEF Period	Outcome Rating
4731	Advancing Landscape Approaches in Ecuador's National Protected Area System to Improve Conservation of Globally Endangered Wildlife	UNDP	Ecuador	GEF - 5	S
4750	Multiplying Environmental and Carbon Benefits in High Andean Ecosystems	UNEP	Regional	GEF - 5	S
4811	CBPF-MSL: Strengthening the Management Effectiveness of the Wetland Protected Area System in Hainan for Conservation of Globally Significant Biodiversity	UNDP	China	GEF - 5	MS
4835	Expansion and Improved Management Effectiveness of the Achara Region's Protected Areas	UNDP	Georgia	GEF - 5	S
4908	GGW: Agriculture Production Support Project (with Sustainable Land and Water Management)	WB	Chad	GEF - 5	U
4922	Decision Support for Mainstreaming and Scaling up of Sustainable Land Management	FAO	Global	GEF - 5	MS
4960	Scaling up Adaptation in Zimbabwe, with a Focus on Rural Livelihoods, by Strengthening Integrated Planning Systems	UNDP	Zimbabwe	GEF - 5	MS
4976	Addressing the Risk of Climate-induced Disasters through Enhanced National and Local Capacity for Effective Actions	UNDP	Bhutan	GEF - 5	MS
4985	Reducing global and local environmental risks from primary mercury mining in Khaidarkan the Kyrgyz Republic	UNEP	Kyrgyz Republic	GEF - 5	MU
4995	Strengthening Climate Information and Early Warning Systems in Eastern and Southern Africa for Climate Resilient Development and Adaptation to Climate Change – Zambia	UNDP	Zambia	GEF - 5	S
4999	Integrated Responses to Short lived Climate Forcers Promoting Clean Energy and Energy Efficiency	UNEP	Mexico	GEF - 5	S
5003	Strengthening Climate Information and Early Warning Systems in Africa for Climate Resilient Development and Adaptation to Climate Change - Burkina Faso	UNDP	Burkina Faso	GEF - 5	MS
5031	Ensuring Global Environmental Concerns and Best Practices Mainstreamed in the Post-Conflict Rapid Development Process of Sri Lanka Through Improved Information Management	UNDP	Sri Lanka	GEF - 5	MU
5038	Implementation of BAT and BEP for Reduction of UP-POPs Releases from Open Burning Sources in Armenia	UNIDO	Armenia	GEF - 5	S
5041	Strengthening Decentralized Management of the Environment to Meet Rio Convention Objectives	UNDP	Guinea	GEF - 5	MS
5048	Capacity Building for the Strategic Planning and Management of Natural Resources in Belize	UNDP	Belize	GEF - 5	MS
5056	Strengthening Community Resilience to Climate Induced Natural Disasters in the Dili to Ainaro Road Development Corridor, Timor Leste	UNDP	Timor Leste	GEF - 5	S

GEF ID	Project Title	Agency	Country	GEF Period	Outcome Rating
5096	Payment for Watershed Services in the Chishui River Basin for the Conservation of Globally Significant Biodiversity	UNDP	China	GEF - 5	MS
5097	Capacity for Implementing Rio Conventions in Vietnam	UNDP	Vietnam	GEF - 5	MS
5110	LME-EA: Applying Knowledge Management to Scale up Partnership Investments for Sustainable Development of Large Marine Ecosystems of East Asia and their Coasts	WB	Regional	GEF - 5	S
5126	Mainstreaming Global Environment Commitments for Effective National Environmental Management	UNDP	Suriname	GEF - 5	MU
5170	Discovering Nature-based Products and Build National Capacities for the Application of the Nagoya Protocol on Access to Genetic Resources and Benefit Sharing	UNDP	Fiji	GEF - 5	MU
5178	Strengthening Capacities to Measure, Report and Verify Indicators of Global Environment Benefits	UNDP	Papua New Guinea	GEF - 5	MU
5186	MENA: Desert Ecosystems and Livelihoods Knowledge Sharing and Coordination Project	WB	Regional	GEF - 5	S
5215	GGW: Forests and Adjacent Lands Management Project	WB	Benin	GEF - 5	MS
5218	Cleantech Programme for SMEs in India	UNIDO	India	GEF - 5	MS
5295	Generating, Accessing and Using Information and Knowledge Related to the Three Rio Conventions	UNDP	Cambodia	GEF - 5	MS
5297	Promoting access to clean energy services in Saint Vincent	UNDP	St. Vincent and Grenadines	GEF - 5	MS
5307	Global Project on the Updating of National Implementation Plans for POPs	UNEP	Global	GEF - 5	S
5448	Implementing the Nagoya Protocol on Access to Genetic Resources and Benefit Sharing	UNDP	Bhutan	GEF - 5	S
5467	Harmonization of Information Management for Improved Knowledge and Monitoring of the Global Environment in Georgia	UNDP	Georgia	GEF - 5	S
5494	Development of Mercury Risk Management Approaches in Latin America	UNEP	Regional	GEF - 5	S
5505	GEF UNIDO Cleantech Programme for SMEs in Turkey	UNIDO	Turkey	GEF - 5	S
5515	GEF UNIDO Cleantech Programme for SMEs in South Africa	UNIDO	South Africa	GEF - 5	S
5525	Global Project on the Updating of National Implementation Plans for POPs	UNEP	Global	GEF - 5	S
5553	GEF UNIDO Cleantech Programme for SMEs	UNIDO	Pakistan	GEF - 5	S
5570	Mainstreaming Rio Convention Provisions into National Sectoral Policies	UNDP	Jordan	GEF - 5	HS
5601	Support to GEF Eligible Countries for Achieving Aichi Biodiversity Target 17 Through a Globally Guided NBSAPs Update Process	UNDP	Global	GEF - 5	S

GEF ID	Project Title	Agency	Country	GEF Period	Outcome Rating
5627	ASTUD PRC Clean Bus Leasing	ADB	China	GEF - 5	S
5633	Lead Paint Elimination Project in Africa	UNEP	Regional	GEF - 5	S
5700	SLM Offset in Western Mongolia	UNDP	Mongolia	GEF - 5	HS
5750	Mainstreaming Sustainable Management of Tea Production Landscapes	UNEP	Global	GEF - 5	S
5771	Improving Mangrove Conservation across the Eastern Tropical Pacific Seascape (ETPS) through Coordinated Regional and National Strategy Development and Implementation	WWF-US	Regional	GEF - 5	S
5797	Securing Tenure Rights for Forest Landscape Dependent Communities: Linking Science with Policy to Advance Tenure Security, Sustainable Forest Management and People's Livelihoods	FAO	Global	GEF - 5	S
5812	Geothermal Resource Development in Saint Lucia	WB	St. Lucia	GEF - 5	S
5826	Strengthening National Systems to Improve Governance and Management of Indigenous Peoples and Local Communities Conserved Areas and Territories	UNDP	Philippines	GEF - 5	MS
5831	Establishing the Foundations of a Partnership to Accelerate the Global Market Transformation for Efficient Appliances and Equipment	UNEP	Global	GEF - 5	HS
9329	Scaling up the SE4ALL Building Efficiency Accelerator (BEA)	UNEP	Global	GEF - 6	HS
9365	Land Degradation Neutrality Target Setting Project	IUCN	Global	GEF - 6	S
9567	Renewable Energy for the City of Marrakech's Bus Rapid Transit System	UNDP	Morocco	GEF - 6	HS

ANNEX B: PILOT PHASE AND GEF-1 PROJECTS ADDED TO GEF IEO TER DATABASE

GEF ID	Project Title	Agency	Country	GEF Period	Outcome Rating
74	Ozone Depleting Substance Consumption Phase-out (first tranche)	WB	Russian Federation	GEF - 1	6
76	Alternate Energy	WB	India	Pilot Phase	5
93	Ozone Depleting Substances Phase-out Project	WB	Bulgaria	GEF - 1	5
94	Technical Support and Investment Project for the Phaseout of Ozone Depleting Substances	WB	Hungary	GEF - 1	6
115	Phaseout of Ozone Depleting Substances	WB	Poland	GEF - 1	5
543	Transcarpathian Biodiversity Protection	WB	Ukraine	Pilot Phase	5

ANNEX C: TERMINAL EVALUATION REPORT REVIEW GUIDELINES

1. The assessments in the terminal evaluation reviews will be based largely on the information presented in the terminal evaluation report. If insufficient information is presented in a terminal evaluation report to assess a specific issue such as, for example, quality of the project's monitoring and evaluation system or a specific aspect of sustainability, then the preparer of the terminal evaluation reviews will briefly indicate so in that section and elaborate more if appropriate in the section of the review that addresses quality of report. If the review's preparer possesses other first-hand information such as, for example, from a field visit to the project, and this information is relevant to the terminal evaluation reviews, then it should be included in the reviews only under the heading "Additional independent information available to the reviewer." The preparer of the terminal evaluation review will take into account all the independent relevant information when verifying ratings.

B.1 Criteria for Outcome Ratings

2. Based on the information provided in the terminal evaluation report, the terminal evaluation review will make an assessment of the extent to which the project's major relevant objectives were achieved or are expected to be achieved¹⁰, relevance of the project results, and the project's cost-effectiveness. The ratings on the outcomes of the project will be based on performance on the following criteria:¹¹

- (a) **Relevance.** Were project outcomes consistent with the focal area/operational program strategies and country priorities? Explain.
- (b) **Effectiveness.** Are project outcomes commensurate with the expected outcomes (as described in the project document) and the problems the project was intended to address (that is, the original or modified project objectives)?
- (c) **Efficiency.** Include an assessment of outcomes and impacts in relation to inputs, costs, and implementation times based on the following questions: Was the project cost-effective? How does the project's cost/time versus outcomes equation compare to that of similar projects? Was the project implementation delayed due to any bureaucratic, administrative, or political problems and did that affect cost-effectiveness?

3. An overall rating will be provided according to the achievement and shortcomings in the three criteria ranging from highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, highly unsatisfactory, and unable to assess.

¹⁰ *Objectives* are the intended physical, financial, institutional, social, environmental, or other development results to which a project or program is expected to contribute (OECD DAC 2002).

¹¹ *Outcomes* are the likely or achieved short-term and medium-term effects of an intervention's outputs. Outputs are the products, capital goods, and services that result from a development intervention; these may also include changes resulting from the intervention that are relevant to the achievement of outcomes (OECD DAC 2002). For the GEF, environmental outcomes are the main focus.

4. The reviewer of the terminal evaluation will provide a rating under each of the three criteria (relevance, effectiveness, and efficiency). Relevance of outcomes will be rated on a binary scale: a 'satisfactory' or an 'unsatisfactory' rating will be provided. If an 'unsatisfactory' rating has been provided on this criterion, the overall outcome achievement rating may not be higher than "unsatisfactory". Effectiveness and Efficiency will be rated as following:

- **Highly satisfactory.** The project had no shortcomings.
- **Satisfactory.** The project had minor shortcomings.
- **Moderately satisfactory.** The project had moderate shortcomings.
- **Moderately unsatisfactory.** The project had noticeable shortcomings.
- **Unsatisfactory.** The project had major shortcomings.
- **Highly unsatisfactory.** The project had severe shortcomings.
- **Unable to assess.** The reviewer was unable to assess outcomes on this dimension.

5. The calculation of the overall outcomes score of projects will consider all three criteria, of which relevance criterion will be applied first - the overall outcome achievement rating may not be higher than "unsatisfactory". The second constraint that is applied is that the overall outcome achievement rating may not be higher than the "effectiveness" rating. The third constraint that is applied is that the overall rating may not be higher than the average score of effectiveness and efficiency criteria calculated using the following formula:

$$\text{Outcomes} = (b + c) \div 2$$

6. In case the average score is lower than the score obtained after application of the first two constraints, then the average score will be the overall score. The score will then be converted into an overall rating with mid values being rounded up upwards.

B.2 Impacts

7. Has the project achieved impacts, or is it likely that outcomes will lead to the expected impacts? Impacts will be understood to include positive and negative, primary and secondary long-term effects produced by a development intervention. They could be produced directly or indirectly and could be intended or unintended. The terminal evaluation review's preparer will take note of any mention of impacts, especially global environmental benefits, in the terminal evaluation report including the likelihood that the project outcomes will contribute to their achievement. Negative impacts mentioned in the terminal evaluation report should be noted and recorded in section 2 of the terminal evaluation reviews template in the subsection on "Issues that require follow-up." Although project impacts will be described, they will not be rated.

B.3 Criteria for Sustainability Ratings

8. Sustainability will be understood as the likelihood of continuation of project benefits after completion of project implementation (GEF 2000). To assess sustainability, the terminal evaluation reviewer will identify and assess the key risks that could undermine continuation of benefits at the time of the evaluation. Some of these risks might include the absence of or inadequate financial resources, an enabling legal framework, commitment from key stakeholders, and enabling economy. The following four types of risk factors will be assessed by the terminal evaluation reviewer to rate the likelihood of sustainability of project outcomes: financial, sociopolitical, institutional frameworks and governance, and environmental.

9. The following questions provide guidance to assess if the factors are met:

- **Financial resources.** What is the likelihood that financial resources will be available to continue the activities that result in the continuation of benefits (income-generating activities, and trends that may indicate that it is likely that in future there will be adequate financial resources for sustaining project outcomes)?
- **Sociopolitical.** Are there any social or political risks that can undermine the longevity of project outcomes? What is the risk that the level of stakeholder ownership is insufficient to allow for project outcomes/benefits to be sustained? Do the various key stakeholders see in their interest that the project benefits continue to flow? Is there sufficient public/stakeholder awareness in support of the long-term objectives of the project?
- **Institutional framework and governance.** Do the legal frameworks, policies, and governance structures and processes pose any threat to the continuation of project benefits? While assessing this parameter, consider if the required systems for accountability and transparency, and the required technical know-how, are in place.
- **Environmental.** Are there any environmental risks that can undermine the future flow of project environmental benefits? The terminal evaluation should assess whether certain activities in the project area will pose a threat to the sustainability of project outcomes. For example, construction of dam in a protected area could inundate a sizable area and thereby neutralize the biodiversity-related gains made by the project.

10. The reviewer will provide a rating as follows:

- **Likely.** There are no risks affecting that criterion of sustainability.
- **Moderately likely.** There are moderate risks that affect that criterion of sustainability.

- **Moderately unlikely.** There are significant risks that affect that criterion of sustainability.
- **Unlikely.** There are severe risks affecting that criterion of sustainability.
- **Unable to assess.** Unable to assess risk on this dimension.
- **Not applicable.** This dimension is not applicable to the project.

B.4 Criteria for Assessment of Quality of Project M&E Systems

11. GEF projects are required to develop M&E plans by the time of work program inclusion, to appropriately budget M&E plans, and to fully carry out the M&E plan during implementation. Project managers are also expected to use the information generated by the M&E system during project implementation to improve and adapt the project to changing situations. Given the long-term nature of many GEF projects, projects are also encouraged to include long-term monitoring plans that measure results (such as environmental results) after project completion. Terminal evaluation reviews will include an assessment of the achievement and shortcomings of M&E systems.

- (a) **M&E design.** Project should have a sound M&E plan to monitor results and track progress in achieving project objectives. An M&E plan should include a baseline (including data, methodology, and so on), SMART (specific, measurable, achievable, realistic, and timely) indicators and data analysis systems, and evaluation studies at specific times to assess results. The time frame for various M&E activities and standards for outputs should have been specified. Questions to guide this assessment include: In retrospect, was the M&E plan at entry practicable and sufficient (sufficient and practical indicators identified; timely baseline; targets created; effective use of data collection; analysis systems including studies and reports; practical organization and logistics in terms of what, who, and when for M&E activities)?
- (b) **M&E plan implementation.** The M&E system was in place and allowed the timely tracking of results and progress toward project objectives throughout the project. Annual project reports were complete, accurate, and with well-justified ratings. The information provided by the M&E system was used to improve and adapt project performance. An M&E system should be in place with proper training for parties responsible for M&E activities to ensure that data will continue to be collected and used after project closure. Question to guide this assessment include: Did the project M&E system operate throughout the project? How was M&E information used during the project? Did it allow for tracking of progress toward project objectives? Did the project provide proper training for parties responsible for M&E activities to ensure data will continue to be collected and used after project closure?
- (c) **Other questions.** This includes questions on funding and whether the M&E system was a good practice.

- Was sufficient funding provided for M&E — in the budget included in the project document?
- Was sufficient and timely funding provided – for M&E during project implementation?
- Can the project M&E system be considered – a good practice?

12. A number rating 1–6 will be provided for each criterion according to the achievement and shortcomings with highly satisfactory = 6, satisfactory = 5, moderately satisfactory = 4, moderately unsatisfactory = 3, unsatisfactory = 2, highly unsatisfactory = 1, and unable to assess = no rating. The reviewer of the terminal evaluation will provide a rating under each of the three criteria (M&E design, M&E plan implementation, and M&E properly budgeted and funded) as follows:

- Highly satisfactory. There were no shortcomings in that criterion of the project M&E system.
- Satisfactory. There were minor shortcomings in that criterion of the project M&E system.
- Moderately satisfactory. There were moderate shortcomings in that criterion of the project M&E system.
- Moderately unsatisfactory. There were significant shortcomings in that criterion of the project M&E system.
- Unsatisfactory. There were major shortcomings in that criterion of the project M&E system.
- Highly unsatisfactory. There was no project M&E system.

The rating for M&E during implementation will be the overall rating of the M&E system:

Rating on the Quality of the Project Monitoring and Evaluation System = b

B.5 Criteria for Assessment of Quality of Terminal Evaluation Reports

13. The ratings on quality of terminal evaluation reports will be assessed using the following criteria:

- (a) The report presents an assessment of all relevant outcomes and achievement of project objectives in the context of the focal area program indicators if applicable.
- (b) The report was consistent, the evidence presented was complete and convincing, and ratings were well substantiated.
- (c) The report presented a sound assessment of sustainability of outcomes.
- (d) The lessons and recommendations are supported by the evidence presented and are relevant to the portfolio and future projects.

- (e) The report included the actual project costs (totals, per activity and per source) and actual cofinancing used.
- (f) The report included an assessment of the quality of the M&E plan at entry, the M&E system used during implementation, and whether the information generated by the M&E system was used for project management.

14. A number rating 1–6 will be provided for each criterion according to the achievement and shortcomings with highly satisfactory = 6, satisfactory = 5, moderately satisfactory = 4, moderately unsatisfactory = 3, unsatisfactory = 2, highly unsatisfactory = 1, and unable to assess = no rating.

Each criterion to assess the quality of the terminal evaluation will be rated as follows:

- Highly satisfactory. There were no shortcomings in the terminal evaluation on this criterion.
- Satisfactory. There were minor shortcomings in the terminal evaluation on this criterion.
- Moderately satisfactory. There were moderate shortcomings in the terminal evaluation on this criterion.
- Moderately unsatisfactory. There were significant shortcomings in the terminal evaluation on this criterion.
- Unsatisfactory. There were major shortcomings in the terminal evaluation on this criterion.
- **Highly** unsatisfactory. There were severe shortcomings in the terminal evaluation on this criterion.

15. The first two criteria (of all relevant outcomes and achievement of project objectives and report consistency and substantiation of claims with proper evidence) are more important and have therefore been assigned a greater weight. The quality of the terminal evaluation reports will be calculated by the following formula:

Quality of the Terminal Evaluation Report = $0.3 \times (a + b) + 0.1 \times (c + d + e + f)$

The total number will be rounded and converted to the scale of highly satisfactory to highly unsatisfactory.

B.6 Assessment of Processes Affecting Attainment of Project Outcomes and Sustainability

16. This section of the terminal evaluation review will summarize the factors or processes related to implementation delays and cofinancing that may have affected attainment of project results. This section will summarize the description in the terminal evaluation on key causal linkages of these factors:

- Cofinancing and project outcomes and sustainability. If there was a difference in the level of expected cofinancing and actual cofinancing, what were the reasons for it? To what extent did materialization of cofinancing affect project outcomes and/or sustainability? What were the causal linkages of these effects?
- Delays and project outcomes and sustainability. If there were delays, what were the reasons for them? To what extent did the delay affect project outcomes and/or sustainability? What were the causal linkages of these effects?
- Country ownership and sustainability. Assess the extent to which country ownership has affected project outcomes and sustainability. Describe the ways in which it affected outcomes and sustainability, highlighting the causal links.

ANNEX D: INSTRUMENT FOR REVIEW OF TE QUALITY

Quality and compliance assessment instrument

1. Project ID
2. Name of the reviewer
3. Lead GEF Agency
4. Who prepared the terminal evaluation report?
 - Agency operations/consultants hired by the operations
 - Agency evaluation unit / consultants hired by the evaluation unit
 - Unable to assess
 - Other (please specify)
5. Was terminal evaluation prepared within six months before to six months after project completion?
 - Yes
 - No
 - Unable to assess
6. Does the report include the names of evaluators that prepared it?
 - Yes
 - No
- * 7. Does the report provide the date of its publication?
 - Yes
 - No
8. Is this a joint project? (A joint project is implemented by two or more GEF Agencies. Usually, the collaborating Agencies have distinguished responsibilities for activities, components, or geographical units covered by the project.)
 - Yes
 - No
9. If this is a joint project, does the report cover all of the GEF-supported activities? (Even though a joint project is implemented by two or more GEF Agencies, the terminal evaluation report submitted to the GEF should be unified and should cover all the project activities.)
 - Yes
 - No
 - Unable to assess
10. Was the project designed as a child project under a programmatic approach?

- Yes
- No

11. If the project is a child project, does terminal evaluation provide information on the parent program under which the project was developed?

- Yes
- No

This set of questions assess the extent to which the terminal evaluation report provides basic information on the covered project.

12. The report provides information on the following: Yes / No / Unable to Assess

- GEF Project ID: Yes / No / Unable to Assess
- Project Name: Yes / No / Unable to Assess
- GEF financing at CEO: Yes / No / Unable to Assess
- Endorsement: Yes / No / Unable to Assess
- GEF financing utilized: Yes / No / Unable to Assess
- Project's global environmental objective: Yes / No / Unable to Assess
- Description of project components/activities: Yes / No / Unable to Assess
- Project start /effectiveness date: Yes / No / Unable to Assess
- Project completion date (actual or expected): Yes / No / Unable to Assess
- Name of the executing agencies of the project: Yes / No / Unable to Assess

13. Does the report clarify the extent to which evaluators that prepared the terminal evaluation were involved in project operations? (not applicable when terminal evaluation has been prepared by or commissioned by project staff directly involved in project implementation)

- Yes
- No
- Unable to Assess
- Not applicable, terminal evaluation prepared by the Agency operations

Operation Focal Point and Stakeholder Feedback

14. Is this a single country project

- Yes
- No

15. Did the terminal evaluation note that the evaluators shared the draft report of the evaluation with the operational focal point

- Yes
- No

16. Does the report describe how evaluators incorporated feedback from the operational * focal point

- Yes
- No

17. Does the terminal evaluation provide information on efforts made to seek and address feedback from the project's stakeholders

- Yes
- No

Theory of Change

18. The report provides an explicit (or implicit) statement on project's theory of change - i.e. how project activities through a causal chain would lead to project outcomes and long-term impact?

- Yes
- No

19. Report discusses project outputs, outcomes and long-term impacts as part of projects theory of change?

- Yes
- No

20. Report describes how causal links among the outputs, outcomes and long-term impacts are supposed to work

- Yes
- No

21. Report describes the assumptions made in the project's theory of change

- Yes
- No

22. Report discusses the extent to which assumptions made in project's theory of change held true during implementation.

- Yes
- No

Methodology - General information

23. Terminal evaluation describes the sources * of information used

- Yes
- No

24. Terminal evaluation uses information gathered through the project monitoring system

- Yes
- No
- Unable to assess

Methodology - Desk Review

25. Desk review was used for data gathering and analysis

- Yes
- No
- Unable to assess

26. Report provides information on the documents that were reviewed for information collection:

- Yes
- No

Methodology – Interviews

27. Interviews were a source of information for the terminal * evaluation?

- Yes
- No
- Unable to assess

28. Report discusses who were covered through the interviews

- Yes
- No

29. Report discusses how those that were interviewed were chosen.

- Yes
- No

30. Report provides a list of those that were interviewed

- Yes
- No

Methodology – online survey

31. Online survey/s was/were used for data collection

- Yes
- No
- Unable to assess

32. Report provides information on who all were covered through the online survey

- Yes
- No

33. Report discusses the rationale for targeting those that were covered through the online survey.

- Yes
- No

34. Report provides information on when the online survey was administered.

- Yes
- No

35. Report provides information on the response rates

- Yes
- No
- Methodology - field verification / visit / observation

36. Information for terminal evaluation was also gathered through field verification observation

- Yes
- No
- Unable to assess

Methodology – field verification / visit / observation

37. Report provides information on who conducted the field verifications / visits/ observations

- Yes
- No

38. Report provides information on where these field verifications / visits / observations were conducted (what was covered through these verifications)

- Yes
- No

39. Report provides information on the basis for site selection for field verification / visits / observations

- Yes
- No

40. Terminal evaluation assesses project outcome's **relevance** * to following:

- GEF Priorities: Yes / No
- Country Priorities: Yes / No
- Agency Priorities: Yes / No

41. Terminal evaluation assesses whether the project design is **relevant** to achievement of the project objectives

- Yes
- No

42. Outcome **relevance** rating provided by the terminal evaluation is consistent with the evidence? (including evidence provided in the report and - where applicable - available through other independent sources)

- Yes
- No
- NA, rating not provided

43. Report presents the **output targets** that were expected at project-start

- Yes
- No

44. The report discusses the extent to which the **output targets** were met.

- Yes
- No

45. The report describes the **outcomes expected** at project start.

- Yes
- No
- Other (please specify)

46. The report describes the extent to which **expected outcomes** were achieved at implementation completion?

- Yes
- No
- Other (please specify)

47. The report clearly describes the extent to which targeted performance was achieved for **each** of the outcome indicators.

- Yes, for all indicators
- Yes, for most indicators
- Yes, for some indicators
- No, does not clearly describe target achievement for any indicator
- Other (please specify)

48. Report discusses the overall level of outcome achievement

- Yes
- No

49. Report discusses the factors that affected achievement of outcomes

- Yes, a detailed discussion

- Yes, but only a cursory discussion
- No

50. **Outcome effectiveness rating** is consistent with the evidence. (including evidence provided in the report and - where applicable - that available through other independent sources)

- Yes
- No
- NA, rating not provided

51. **Outcome efficiency:** Report provides information on timeliness of execution of project activities

- Yes
- No

52. **Outcome efficiency:** discusses the extent to which project's use of resources - funds, staff, processes - was efficient.

- Yes
- No

53. **Outcome efficiency:** The report discusses the factors that affected efficiency.

- Yes
- No

54. **Outcome efficiency rating** provided by the terminal evaluation is consistent with the evidence. (including evidence provided in the report and - where applicable - available through other independent sources)

- Yes
- No
- NA, rating not provided

55. **Overall outcome rating** provided by the terminal evaluation is consistent with the evidence. (this includes evidence provided in the report and - where applicable - that available from other independent sources)

- Yes
- No
- NA, rating not provided

Sustainability

56. Report describes the risks that may affect project **sustainability** (sustainability is understood as: the extent to which the net benefits of the intervention continue, or are likely to continue. (OECD, 2020))

- Yes
- No

57. Report discusses probability/likelihood of materialization of the major risks that may affect **sustainability**?

- Yes
- No

58. Report discusses the likely impact of a major risk's materialization on **sustainability**

- Yes
- No

59. Report provides an overall assessment of the likelihood of project **sustainability**.

- Yes
- No

60. **Sustainability rating** is consistent with the evidence. (including evidence provided in the report and - where applicable - available through other independent sources)

- Yes
- No
- NA, rating not provided

Project M&E

61. Describes the strengths and weaknesses of * project **M&E design**

- Yes
- No

62. Describes strengths and weaknesses of the **M&E implementation**

- Yes
- No

63. **M&E rating** is consistent with the evidence. (this includes evidence provided in the terminal evaluation and - where applicable - that available from other independent sources).

- Yes
- No
- NA, rating not provided

Project Implementation and Execution

64. The report provides an overall assessment of the performance of the **GEF Agency** in project implementation.

- Yes
- No

65. The report discusses the performance of the **GEF Agency** during different stages of the project cycle (preparation, start up, implementation).

- Yes
- No

66. The report discusses the performance of the **GEF Agency** in supervising the work of the executing agency.

- Yes
- No

67. The **quality of implementation (or equivalent) rating** of **GEF Agency** is consistent with the evidence.

(this includes evidence provided in the report and - where applicable - that available through independent sources).

- Yes
- No
- NA, no rating provided

68. Report discusses the overall performance of the * **executing agency**.

- Yes
- No

69. Report describes the performance of the **executing agency** discussing its performance in terms of execution of project's day to day activity, use of funds, staffing, procurement, etc.

- Yes
- No

70. The **quality of execution (or equivalent) rating** of the **executing agency** is consistent with the available evidence (including that provided in the report and - where applicable - that provided by other independent sources)

- Yes
- No
- NA, no rating provided

Cofinancing

71. Terminal evaluation provides information on expected cofinancing

- Yes
- No

72. Information on expected cofinancing provided in the report is consistent with the information in CEO Endorsement documents

- Yes
- No (please specify)

73. Provides information on materialized cofinancing

- Yes
- No
- Materialized Cofinancing

74. Information on materialized cofinancing is internally consistent (figures tally)

- Yes
- No (please specify)

75. Provides following details on materialized cofinancing

- Type of cofinancing - Cash or In-kind: Yes/No/Not Applicable
- Type of cofinancing - Loan, grant, equity, guarantee: Yes/No/Not Applicable
- Who administered cofinancing (implemented activities funded by cofinancing): Yes/No/Not Applicable

76. Does the overall materialized cofinancing match with the level of expected cofinancing? (within $\pm 10\%$)

- Yes
- No
- Unable to assess- some cofinancing information provided but it is difficult to know the extent to which cofinancing met the targets

77. If the overall materialized cofinancing was below (or above) expectations does the report provide the reasons for it?

- Yes
- No

78. If the overall materialized cofinancing was below (or above) expectations, does the report discuss the effects of the shortfall/or excess materialization on project results.

- Yes
- No

Social and Environmental Safeguards

79. Provides information on whether environmental and social safeguards were applied during project preparation

- Yes

- No

80. Provides information on whether environmental and social safeguards were applied during project implementation.

- Yes
- No

Gender analysis

81. Provides information on whether a gender analysis was conducted to identify concerns?

- Yes
- No

82. Provides information on application of the actions identified in the gender analysis.

- Yes
- No
- Not Applicable, gender analysis was not conducted.

Lessons

83. The report presents lessons

- Yes
- No

84. The lessons are based on the evidence presented in the report.

- Yes
- Some of the lessons are based on the evidence whereas others are not
- No

85. The report indicates situations where the lessons will be applicable.

- Yes, explicitly
- Yes, implicitly
- No

Recommendations

86. Terminal evaluation presents recommendations.

- Yes
- No

87. The recommendations (including supporting text) state what * needs to be done

- Yes
- No

88. Recommendations (including supporting text) make it clear as to who needs to take the recommended action.

- Yes, explicitly
- Yes, implicitly
- No

89. Recommendations (including supporting text) make the time frame for recommended actions clear.

- Yes
- No

Overall assessment of the report

90. To what extent do you agree or disagree with the following statements about the terminal * evaluation report

- The report is easy to understand (lucid): Strongly Agree / Agree / Disagree / Strongly Disagree
- The data on project outputs and outcomes is presented in manner that facilitates comparison of achievement with targeted performance: Strongly Agree / Agree / Disagree / Strongly Disagree
- The report is comprehensive (covers all of the important aspects): Strongly Agree / Agree / Disagree / Strongly Disagree
- The report is concise (to the point/not repetitive): Strongly Agree / Agree / Disagree / Strongly Disagree
- The conclusions are consistent with the evidence: Strongly Agree / Agree / Disagree / Strongly Disagree
- The performance ratings are consistent with the evidence: Strongly Agree / Agree / Disagree / Strongly Disagree

91. Is this terminal evaluation a good practice?

- Yes, overall a good practice
- Yes, some sections are a good practice
- No, but not an example of poor practice.
- No. TE is an example of poor practice. Following are the areas where it is very weak: ____

92. If project is a good practice, explain why this terminal evaluation or sections of this terminal evaluation are a good practice: _____

93. If project is a good practice, would you recommend that this terminal evaluation is used as an example of good practice for terminal evaluation preparation?

- Yes
- No

ANNEX E: INSTRUMENT FOR REVIEW OF CANCELED PROJECTS

Questions for CEO Endorsed/Approved Projects

1. Project ID:
2. Project Title
3. Lead Agency:
4. Project CEO Endorsement/Approval Date (If applicable)
5. Project Implementation Start Date
6. Project Cancellation Date (date of GEF SEC notification of cancellation)
7. Did implementation start before cancellation/dropping?
 - No
 - Yes. (Please provide the explanation given.)

Projects for which Implementation Never Began

8. If implementation for project never began, was a reason provided for cancellation in a cancellation memo?
 - No
 - Yes. (Please provide the explanation given.)
9. If implementation for the project never began, was a reason provided for cancellation in interim reporting?
 - There was interim reporting but it did not explain cancellation.
 - There was no interim reporting.
 - Yes. (Please provide the explanation given.)

Projects for which Implementation Began

10. If implementation of project began, how many years after the start date was the project canceled?
 - 1 year or less after implementation start
 - More than 1 year after implementation start
11. Projects for which Implementation Began, review of Project Implementation Reports (PIRs)

12. If PIRs are available, do any of the PIRs flag problems that may jeopardize achievement of project outcomes/outputs?

- No PIRs available
- PIRs are available but none of the PIRs flag performance related concerns
- Yes and at least one of the PIRs flags performance related concerns. (Please describe the concerns flagged.)

13. If PIRs are available, how many are available?

14. Are they available for every year the project was under implementation? (note a project may have one PIR less than the duration for which it was under implementation, exclude repetitions)

- Yes
- No

Projects for which Implementation Began, review of Midterm Evaluation

15. If a midterm report is available, does it flag problems that may jeopardize achievement of project outcomes/outputs?

- No MTE available
- MTE is available but it does not flag performance related concerns
- Yes, the MTE flags performance related concerns. (Please describe the concerns flagged.)

Projects for which Implementation Began, review of Cancellation Memo

16. If a cancellation memo is available, does it explain the reason for project cancellation?

- There is no Cancellation Memo (by Agency) available.
- Cancellation Memo (by Agency) is available but it does not explain reason for cancellation.
- Cancellation Memo (by Agency) is available and it explains the reason for cancellation (Please provide the reason for cancellation given in the memo.)

17. Any other comments?

Annex F: Instrument for Review of Low Performing Projects

1. Project ID:
2. Project Title
3. Lead Agency:
4. Project CEO Endorsed/Approval Date (If applicable):
5. Project Implementation Start Date
6. Project Completion or Cancellation Date
7. Are PIRs available for the project?
Yes
No

Questions on Project Implementation Reports (PIRs)

8. How many Project Implementation Reports are available?
9. Are PIRs available for every year the project was under implementation? (note a project may have one PIR less than the duration for which it was under implementation, exclude repetitions)
Yes
No
10. Did the ratings for implementation progress reported in any of the PIRs indicate performance in the Unsatisfactory range?
Yes
No
11. Did the ratings for development objectives reported in any of the PIRs indicate performance in the Unsatisfactory range?
Yes
No
12. Did the ratings for risks to project achievement of outcomes reported in any of the PIRs indicate that achievement of outcomes is in jeopardy?
Yes
No

13. Did the last PIR submitted to the GEF Secretariat indicate performance in the Unsatisfactory range for implementation progress?
Yes
No
14. Did the last PIR submitted to the GEF Secretariat indicate performance in the Unsatisfactory range for achievement of Development Objectives?
Yes
No
15. Did the rating for risks to project achievement in the last PIR submitted to the GEF Secretariat indicate that achievement of outcomes was in jeopardy?
Yes
No
16. Please provide a description of any factors noted in the PIRs as leading to the project's poor performance:

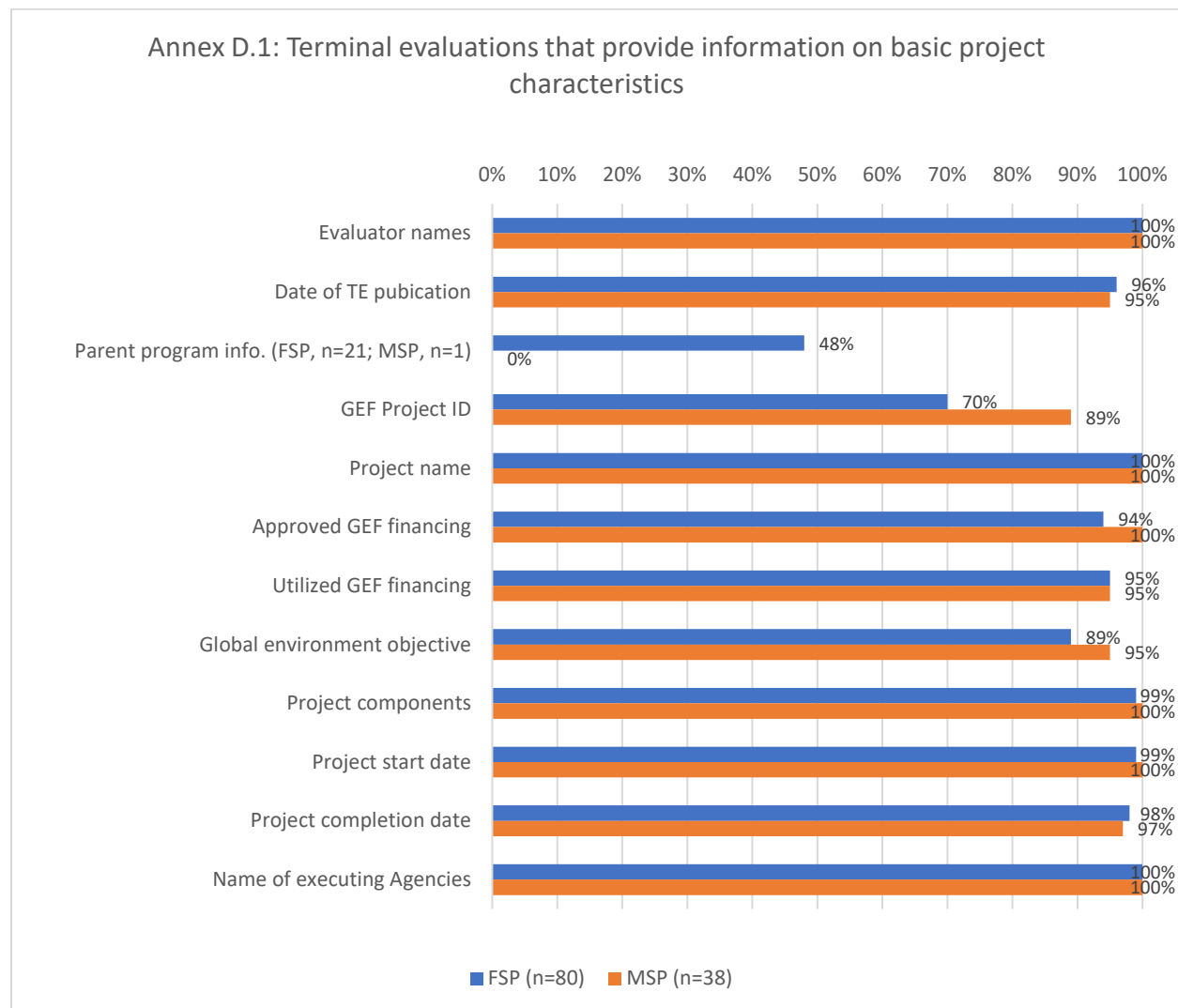
Questions on Midterm Evaluation

17. Is a midterm evaluation available?
Yes
No
18. Do the performance ratings provided in the midterm review indicate low performance?
Yes
No
19. Does the midterm review indicate there was a need to restructure the project?
Yes
No
20. Please provide a description of factors noted in the midterm review as leading to the project's poor performance:

Terminal Evaluation/Cancelation Memo

21. Please provide a description of factors noted in the TE, TE review or Cancelation Memo (if applicable) as leading to the project's poor performance:
22. Any additional comments about the project?

ANNEX D: RESULTS OF REVIEW OF TERMINAL EVALUATION QUALITY



Annex D.2: Percentage that seek and address feedback from stakeholders (observations in parentheses)

TE describes that.....	FAO	UNDP	UNEP	UNIDO	WB	Others	All
Full-size projects							
Draft report was shared with OFP ¹²	17 (6)	39 (18)	25 (4)	40 (5)	26 (23)	17 (6)	29 (62)
OFP Feedback was addressed	0 (6)	22 (18)	25 (4)	40 (5)	9 (23)	0 (6)	15 (62)

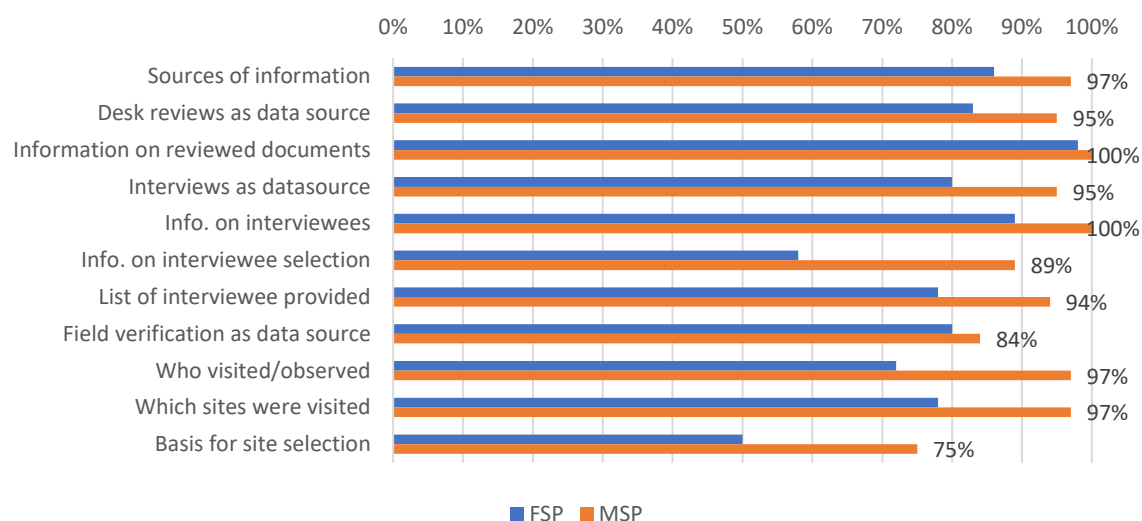
¹² The questions related to sharing of draft report with OFPs and addressing their comments in finalization of the terminal evaluation reports was applied only in case of national projects. Stakeholder feedback related questions was applicable to all projects.

Feedback of stakeholders addressed	78 (9)	89 (18)	100 (13)	100 (6)	72 (25)	44 (9)	81 (80)
Medium-size projects							
Draft report was shared with OFP	—	81 (16)	75 (4)	20 (5)	0 (1)	—	65 (80)
OFP Feedback was addressed	—	81 (16)	75 (4)	20 (5)	0 (1)	—	65 (80)
Feedback of stakeholders addressed	100 (1)	100 (17)	92 (12)	80 (5)	0 (2)	0 (1)	87 (80)

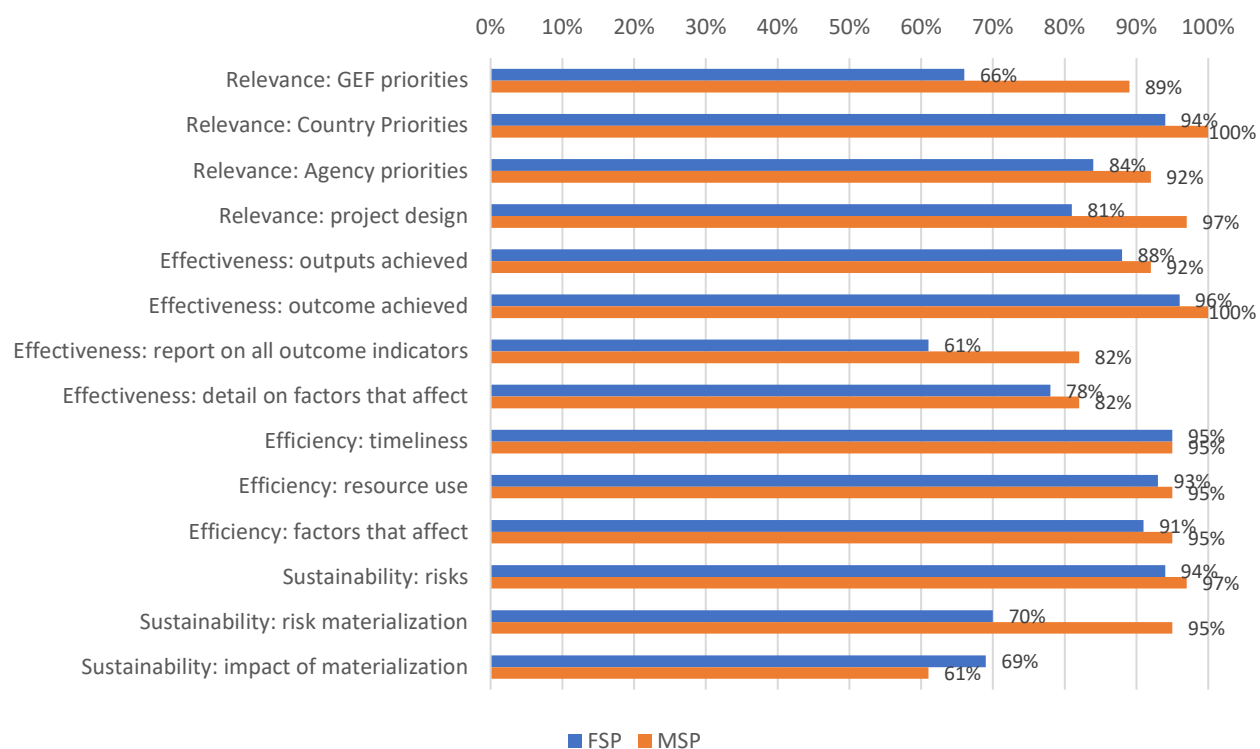
Annex D3: Percentage of Terminal Evaluations that discuss aspects of the Project's Theory of Change

Report discusses.....	FAO	UNDP	UNEP	UNIDO	WB	Others	All
<i>Full-size projects (observations)</i>	<i>9</i>	<i>18</i>	<i>13</i>	<i>6</i>	<i>25</i>	<i>9</i>	<i>80</i>
Project's theory of change	89	44	100	100	96	78	83
Causal links of theory of change	89	39	100	100	96	78	81
How causal links are supposed to work	78	17	100	83	80	78	69
Assumptions of the theory of change	78	33	100	83	88	56	73
Whether assumptions remained valid	0	17	92	50	48	56	44
<i>Medium-size Projects (observations)</i>	<i>1</i>	<i>17</i>	<i>12</i>	<i>5</i>	<i>2</i>	<i>1</i>	<i>38</i>
Projects theory of change	100	76	100	100	50	100	87
Projects causal links	100	59	100	100	50	100	79
How causal links are supposed to work	100	29	100	80	50	100	63
Assumptions of the theory of change	100	76	100	80	50	100	84
Whether assumptions remained valid	100	65	92	40	50	0	68

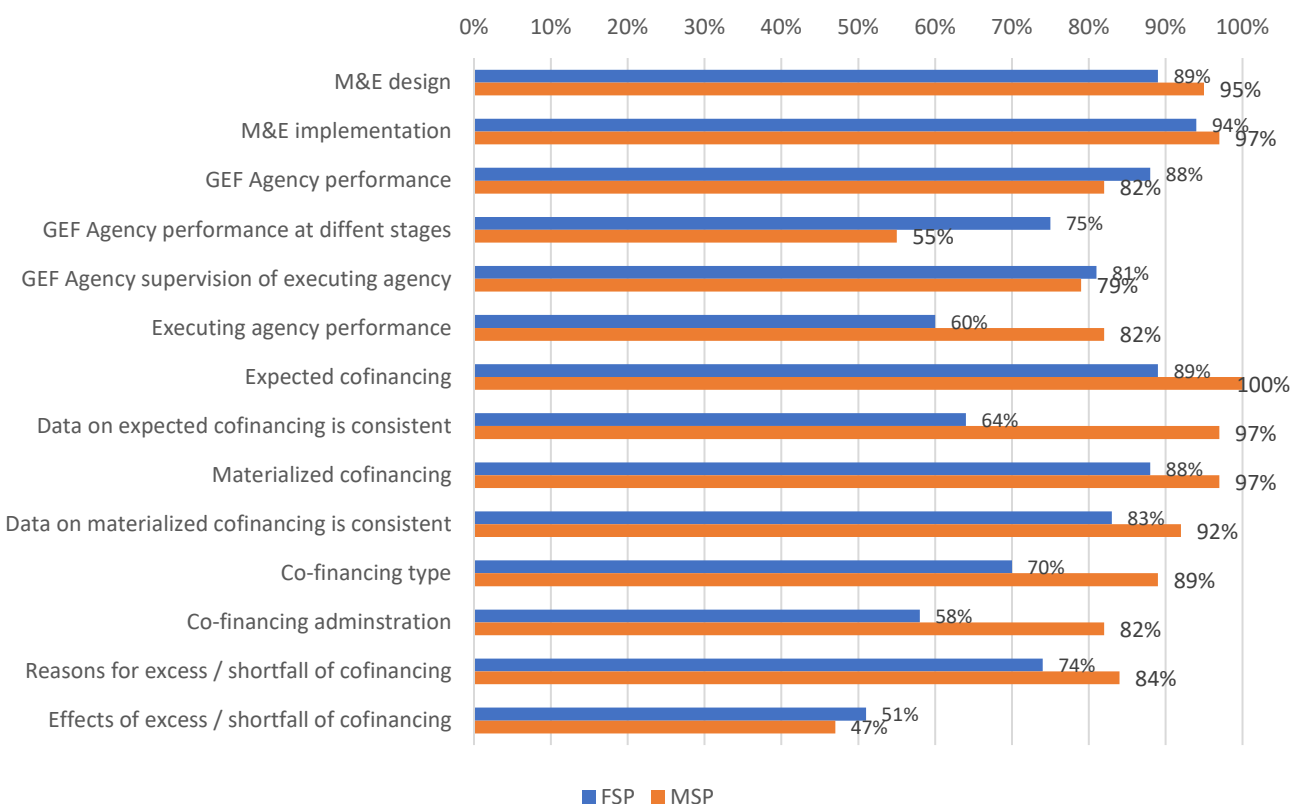
Annex D.4: Percentage of terminal evaluations that discuss data sources as part of the discussion on methodology



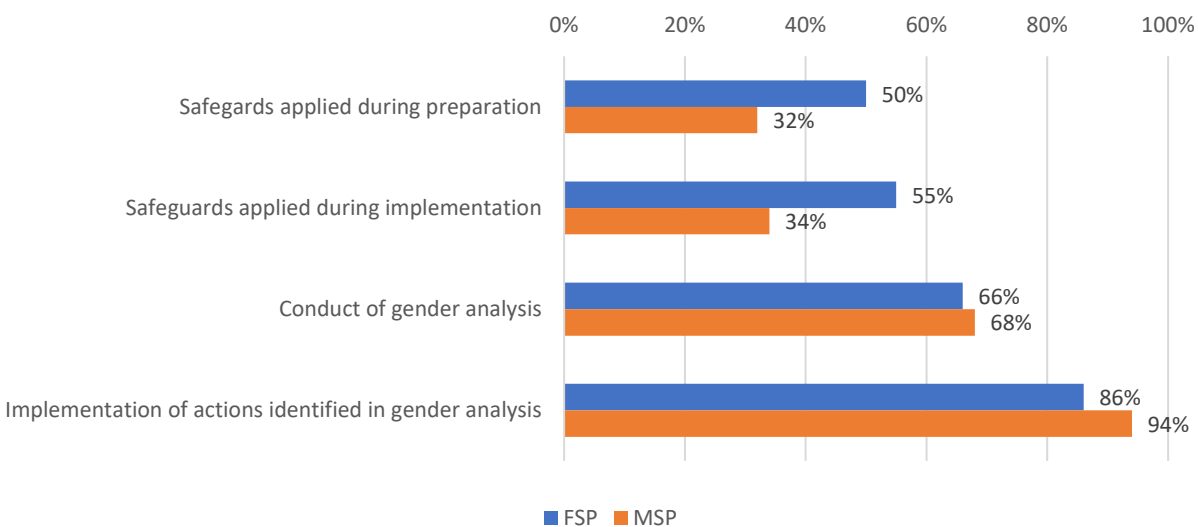
Annex D.5. Percentage of Terminal Evaluations Reporting on various facets of outcome performance



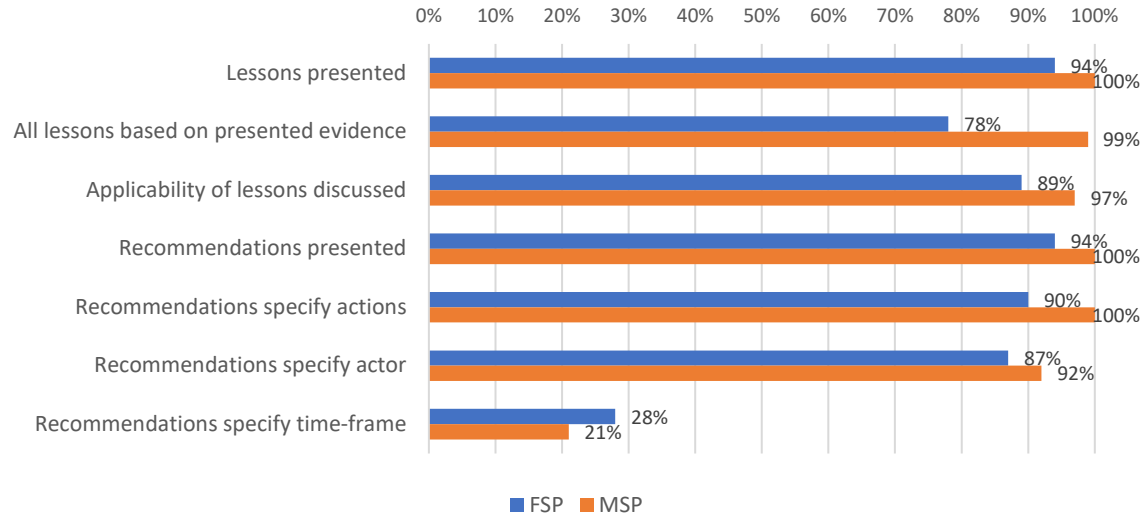
Annex D.6: Percentage of terminal evaluations that report on M&E, Agency performance, co-financing



Annex D.7. Terminal evaluations reporting on social and environmental safeguards



Annex D.8: Percentage of terminal evaluations that present lessons and recommendations



Annex D.9: Performance ratings in terminal evaluations provided and consistency of performance ratings - by project size

