

Feedback on the Draft Approach Paper and Evaluation Team Response

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
Sonja Teelucksingh, Senior Environmental Specialist, GEF Secretariat, and peer reviewer for the evaluation			
1	<i>The Agency self-evaluation systems are expected to facilitate learning and accountability across the GEF partnership.</i>	Sonja Teelucksingh: More detail would be useful in the context of the value of the Agency self-evaluation systems to the GEF's portfolio monitoring and evaluation, and its importance in providing early warnings and possible remedial actions on any implementation challenges.	More detail been added on this specific topic in the introduction section of the revised paper.
2	<i>The evaluation will assess the factors that affect these systems, and to identify areas for improvement.</i>	Sonja Teelucksingh: Why should the evaluation spend time on the factors that affect these systems, especially if these factors are not in our control? Perhaps an identification of areas for improvement in the areas that we CAN control should be more the focus.	The purpose of this line of questioning is to understand how these factors affect achievement of outcomes. Even if the factors are not in our control, we could take these into account in designing of our activities.
3	<i>The expectations from the self-evaluation systems of the Agencies are outlined in several GEF policy documents and policies of the GEF Agencies. For example,...</i>	Sonja Teelucksingh: The paper mentions a number of policies that deal with self-evaluations, but omits some key ones such as gender, stakeholder engagement. Was this an oversight, or are these policies beyond the scope of the exercise? If the latter, then the paper needs to clearly state this as one of the limitations.	A reference to reporting on gender and safeguards has been added in the discussion on quality of reporting.
4	<i>However, there are variations in self-evaluations needs and practices of the GEF Agencies given the differences in their mandates, scale of operation, level of independence of their evaluation function, and their self-evaluation traditions. Therefore, their self-evaluation practices for GEF supported activities may vary.</i>	Sonja Teelucksingh: The paper acknowledges that there are variations in self-evaluation practices, capabilities and experiences across the GEF Agencies. However, there will inevitably be similarities amongst some, and this would also be a valuable addition to the discussion.	Both similarities and variations are there. Similarity is a built-in assumption in the discussion. The point that the systems vary is emphasized because the evaluation needs to address these variations as well.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
3	<i>The GEF IEO is undertaking this evaluation in response to requests from the GEF Council and the GEF Secretariat to assess the performance of the self-evaluation systems of the Agencies . The GEF Council and the Secretariat are interested in ensuring that the self-evaluation systems of the Agencies monitor their GEF portfolios well, facilitate learning, and are harmonized . The evaluation will focus on how the Agency self-evaluation systems address the GEF supported activities .</i>	Sonja Teelucksingh: It would be helpful to examine the nature of these requests, and appropriate reference and links to the source Council document would assist in this regard.	No need for change. GEF IEO's relationship with Council and Secretariat is codified in the GEF Instrument. Communications among the three take place outside the realm of Council meetings and documents as well.
5-6	<i>Sub-Section on Knowledge Management</i>	Sonja Teelucksingh: This sub-section gives interesting information, but it would be helpful to elaborate a little more on its relevance to this study.	This discussion is important because we want the evaluation to adequately address the learning dimension of outcomes of the self-evaluation system. Knowledge management is the key mechanism through which this is supposed to happen.
6	<i>For example, evaluation units of several GEF Agencies such as UNDP, UNEP, and IFAD assess and report on quality of self-evaluations through their annual reports. However, in these reports' coverage of topics such as candor in reporting and learning is not detailed .</i>	Sonja Teelucksingh: It would be good for this paper to outline the methodologies these Agency evaluation units employ in assessing and reporting on the quality of the Agency self-evaluations.	No change made. The evaluation does not cover the independent evaluation system of the Agencies. The information provided in the paper is to lay out what is available. Greater details on the methodologies deployed may accessed at the cited sources.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
6	<i>EBRD, ADB and UNDP are presently undertaking an assessment of their self-evaluation systems</i>	Sonja Teelucksingh: The AfDB could be added to this list of agencies currently undertaking an evaluation of the self-assessment system: https://idev.afdb.org/en/document/ongoing-evaluation-bank%E2%80%99s-self-evaluation-systems-and-processes What are the respective timelines of these evaluations? Is there any way for those reports to have synergies with this one?	The list of Agencies that have conducted or are conducting evaluation of their self-evaluation systems has been updated. If these evaluations become available before the analysis of the GEF IEO's evaluation is complete, then the findings of the former would be considered.
7	<i>APR 2006 (GEF IEO 2007) included an assessment of project supervision practices of World Bank, UNDP and UNEP, and – among other topics – covered quality of reporting through the annual project implementation reports. APR 2015 covered gaps in submission of tracking tools by the GEF Agencies. However, GEF IEO is yet to conduct a comprehensive assessment of the self-evaluation systems of GEF Agencies.</i>	Sonja Teelucksingh: 2006 is a very long time ago. Is there anything more recent, say, in OPS-6? “Gaps in submission” covers compliance issues only. This is not the only target of self-evaluation assessments. An assessment of quality (as opposed to quantity) could be highly informative as well.	APR 2020 includes a detailed section on quality of terminal evaluations. APRs in general contain a section on quality of terminal evaluations, although it is usually quite brief. This information has been updated in the revised approach paper.
7	<i>The GEF Independent Evaluation Office (IEO) is undertaking the ‘Evaluation of the Agency Self-Evaluation Systems’ to assess the extent to which Agency self-evaluation systems meet the GEF requirements and provide information that is sufficient, timely, credible, and useful.</i>	Sonja Teelucksingh: Can you be clearer on what you mean by the “GEF requirements”? To my knowledge, beyond the provision of key implementation data points, and compliance with submission of reports, there have been no clear requirements in the past (though this has changed in GEF-7).	GEF Evaluation Policy has minimum standards on M&E, GEF IEO guidelines on terminal evaluations also specify requirements. These are mentioned in the paper. More details have been added in the revised paper.
7	<i>The evaluation will be based on a theory of change presented in Figure 1</i>	Sonja Teelucksingh: A valuable component of a ToC is an identification of potential risk factors and associated mitigating measures. This is missing from this diagram.	We use a parsimonious approach to theory of change. Not everything is included – nor everything needs to be included. We have included more information on assumptions, which also covers risks, in the revised paper.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
8	<i>Figure 1</i>	Sonja Teelucksingh: Figure 1 does not align well with the text in the section on system effectiveness. How are the dependent variables illustrated in Figure 1? The text in three boxes aligns with some of the outcome descriptions, but not all. For example, Information quality is noted in the text as a key intermediate outcome, but is not listed in the Figure 1. Figure 1 lists “learning” as an outcome. This is an all-encompassing term; as such it would be beneficial to discuss this outcome with more precision. For example - who are the targeted groups for learning? Via what media, which may vary according to group? What of information sharing? For which (all?) of these steps is the self-evaluation accountable?	No change. Figure 1 is not meant to provide a detailed description. It’s just a heuristic tool. The narrative explains the figure – and therefore provides more detail.
9	<i>When targets and milestones are not met, an Agency clearly communicates non-achievement and, where applicable, facilitates corrective actions</i>	Sonja Teelucksingh: From where has this statement come? Is it Policy driven? Are there examples that can be given?	It’s a description of what accountability related outcome would look like. The description lays the normative expectation of what it means when we say that a self-evaluation system is contributing to enhancing accountability.
11	<i>Among the organizational characteristics, variables such as business model, scale of operation, organizational culture, and relationship with independent evaluation affect the design and performance of the self-evaluation system.</i>	Sonja Teelucksingh: The section on organizational characteristics includes multiple examples of potential <u>differences</u> among the GEF agencies, when on most of the listed criteria the agencies seem remarkably <u>similar</u>: “international organizations that work at scale” seems to cover all but a few (newer) GEF agencies.	No response required.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
11	<i>How do GEF Agencies address self-evaluations through their policy framework ?</i>	Sonja Teelucksingh: Some key additional questions to be added here: • To what extent are Agencies' self-evaluation systems, underlying policy frameworks, and resultant ratings comparable? • To what extent, and along what parameters, is it possible to unify the self-assessment systems across Agencies? • What has been the Agencies' own development path to ratings in their own timeline? Have they changed over time to the extent that past data is rendered non-comparable? Are we able to look at a single Agency's ratings trajectory over time using historical implementation data? Furthermore, the discussion of the policy framework should also include how Agencies implement these frameworks.	Thanks for the inputs. These will be considered in developing the questions further.
11-12	<i>The evaluation will assess the extent to which policies explain the purpose and role of self-evaluations, provide guidance on how the self-evaluations ought to be conducted, and clarify relationship with independent evaluation.</i>	Sonja Teelucksingh: A key additional question to be added here: • Are there any international evaluation standards that can be applied to self-evaluation, and how do the Agencies' practices reflect these standards?	The discussion on literature already covers UNEG, ECG guidance on good practices, and work of MOPAN. We will address this as necessary.
12	<i>The evaluation will assess how GEF Agencies address the credibility of information generated by their self-evaluation system.</i>	Sonja Teelucksingh: This is a vital point that is deserving of its own "key question".	The issue is discussed at an appropriate level in the listed questions in the revised paper.

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
12	<i>To what extent are the Agency self-evaluation systems meeting the needs of GEF partnership? The evaluation will record perceptions of the Agency staff, national counterparts, and consultants, on the extent to which the Agency self-evaluation systems are effective in supporting the learning and accountability needs of the GEF partnership.</i>	Sonja Teelucksingh: Within the GEF Partnership, the GEF Secretariat (and the GEF Council) are heavy users of the data from the Agencies' self-evaluation systems, and yet the Secretariat is largely absent from the focus of the paper. The Secretariat should therefore be <u>explicitly</u> considered - the analysis and recommendations should make clear how the Secretariat uses (or can potentially use) the self-evaluation systems as a mean to enhance the overall performance of projects, policies and programming strategies. The question on "meeting the needs of the GEF partnership" does not explicitly include comments/input from the GEF secretariat. In addition to the "Agency staff, national counterparts and consultants", the Secretariat's views should also be sought. In addition, the Evaluation Office is itself a user group of this data and its views should also be included as a data point. This section would benefit from clarification that individual agency systems will be assessed, as opposed to the implication that all Agencies will be considered together.	The role of Secretariat is addressed through the review of RBM, which also covers the GEF Portal. The GEF IEO role in using the self-evaluation information is being captured through the review on RBM/Portal and through review of the GEF IEO's terminal evaluation validation process. These evaluations, including the evaluation of the Agency self-evaluation systems, will contribute to the OPS-7. Therefore, we don't intend to enhance the scope of this evaluation – which focuses on the Agency self-evaluation systems. As clearly mentioned in the draft approach paper, the individual Agency is the unit of analysis. Therefore, each Agency would be covered separately.

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
12	<i>It will also assess effectiveness by determining the extent to which information provided by the system is comprehensive, timely, credible, accessible, and useful, and in line with the GEF requirements .</i>	Sonja Teelucksingh: One key additional question to consider is - what are the best practices of any agencies in terms of self-evaluations? For example, UNDP changed their structure recently – is their current system an example of best practice? A related question to be considered here is, given that much of the ratings data is aggregated across different aspects of the portfolio for monitoring purposes, how (or is it even possible) to set a benchmark for comparison among the data points that emerge from all of these different self-evaluation systems across Agencies? “Ratings” is not an homogenous category - there are different ratings that focus on different aspects of a project throughout its life cycle – IP, DO, and risk ratings, and also sustainability ratings at TE. Will the IEO evaluate them all? If yes, then this paper should be expanded to cover objective setting, implementation, and risk management. An examination of the self-ratings relative to implementation parameters and an analysis of any clear deviations between them can be very valuable as a methodological approach to determining “credibility” - suggested implementation parameters include duration, disbursement/endorsement speed, M+E document submission, and financial closure. An analysis of the outliers of each of those parameters would be useful – for example, a project rated highly but with a low disbursement speed, a project rated highly but in its 15th PIR, etc. An analysis of low rated projects and/or low performing projects themselves would also be instructive.	Best practices: The evaluation will capture good practices. It will also facilitate sharing of these practices through workshops and focus groups. Ratings: APR 2020 presented an analysis on the performance ratings given in the terminal evaluations. This work is being deepened further and will be reflected in the evaluation on self-evaluation system. We don’t intend to expand the review to the ratings in PIRs and MTRs. However, the review has assessed the extent to which PIR and MTR reporting captures the emerging risks and is timely.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
12	<i>What are the factors that affect effectiveness of the self-evaluation systems?</i>	Sonja Teelucksingh: Level of resources allocated for self-evaluation, in theory, is directly related to agency fees (which are standardized) and to M&E funding as a project line item approved by Council. Does the evaluation propose to assess the level of resources for self-evaluation at both the project level and the agency level?	We will address this topic in the report in terms of sufficiency of resources for self-evaluation.
12	<i>It will assess how presence of a robust independent evaluation function affects a self-evaluation system's effectiveness .</i>	Sonja Teelucksingh: "It will assess how..." – this suggests a predetermined conclusion. Please replace "how" with "if". What is the methodology for this assessment?	Thanks for the suggestion. We have incorporated it.
12	<i>These hypothesis address variables where substantial variations may be expected among GEF Agencies .</i>	Sonja Teelucksingh: How will the assessment deal with these variations?	In a case study approach, variations in independent variables across cases help us understand the effect of these on the dependent variable.

13	Section on Evaluation Design	Sonja Teelucksingh: This section is very scant and needs further scoping; the reviewers should also be given the opportunity to comment on that expanded material before the approach paper becomes final. Some points to consider are as follows: <u>Project Types and Project Focus:</u> It would be useful to provide a clearer sense as to how the evaluation will look at different projects which may have different self-reporting requirements or requirements (such as programs, or capacity building projects). <u>Reporting differences:</u> on the assumption that many more than 2 projects will be considered (as per the comment directly below), analyses of projects must especially be careful to note monitoring differences (that are based on policy) – for example between the GEF-6 IAP Pilots and regular GEF-6 projects; the differences before and after the performance indicator Table was required at PIF stage; the requirements for PFDs, PIFs, MSPs, and child projects; and the requirements for focal area projects versus global set-aside projects.. The evaluation should segment projects into appropriate phases that align with GEF monitoring policies in effect at the time, and conduct case studies accordingly. <u>Project stage:</u> will the selection be limited to completed projects only for which all reports including TEs are available? <u>Rating Type:</u> As mentioned in an earlier comment, “ratings” is not an homogenous category - there are different ratings that focus on different aspects of a project throughout its life cycle	The section has been strengthened with new information on evaluation criteria, system’s design thinking approach, which will be deployed for data gathering and analysis. The evaluation does not focus on project types – activity type. Its interest in these is only to the extent the self-evaluation processes for these are the same or different. Where the processes are different, the rationale and effect of different processes will be assessed. One or two projects are sufficient to understand the process deployed for self-evaluation, as long as the process is the same. Number of projects covered per Agency will differ based on whether an Agency uses the same process for all projects or different processes based on the type of project. Sampling in this case is for illustration and not for ensuring representativeness. The sampling will be restricted to recently completed and under implementation projects. Ratings: As noted earlier, only ratings provided in the terminal evaluations will be considered for assessment. Comparisons: comparisons among Agencies will be made. The evaluation will document there experiences and share their good practices.
----	------------------------------	--	---

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
		– IP, DO, and risk ratings, and also sustainability ratings at TE. Will the IEO evaluate them all? If yes, then this paper should be expanded to cover objective setting, implementation, and risk management. <u>Project age</u>: the selected projects will likely be those that were approved and under implementation prior to the 2019 updates to the Monitoring Policy and the Evaluation Policy. Therefore, the analysis will by definition be a little dated. Whilst this is unavoidable, the evaluation should make explicit note of the new Policies and cast any forthcoming recommendations within that framework. <u>A “Compare and Contrast” Element</u>: The design should include some elements related to compare and contrast between the agencies. Some agencies may have rigorous policies but lack compliance; other agencies may lack policies but have robust compliance; other agencies may have great production but poor utilization. The evaluation design should include techniques for making appropriate comparisons.	

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
13	<i>For each of the Agencies, two GEF supported projects will be selected to assess operation of the self-evaluation system at the project level.</i>	Sonja Teelucksingh: The sample size of 2 projects per Agency is too small. There are also different lessons to be learned by project type as well (FSPs, MSPs, EAs, etc.) as they each present different evaluation opportunities and challenges. For instance, would PPGs and Enabling Activities not present a unique set of self-assessment opportunities? Whatever the final number (and it should be more than two!), there should also be an indication in this paper of how this sample will be spread across the different cohorts of the GEF portfolio: replenishment phase, focal area, region, project type, project size, etc. Country context will also be particularly important – for example, projects in FCV countries may face particular implementation challenges.	Sample size: we don't agree with the argument. The sampling is for illustration and not for ensuring representativeness. The approach followed here is of qualitative, not quantitative, research. We don't agree with the suggestion that this needs to be done in the approach paper. This information may be provided as annex in the final report.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
13	<i>Desk Reviews : the source material from GEF Agencies will be reviewed.</i>	Sonja Teelucksingh: It should be made clear what source material will be GEF specific, and what is not. The Desk Review will use GEF Agency sourced material, but there may be useful material from other sources as well: e.g. reviews of GEF Agencies and practice done by NGOs, individual GEF donors, COP bodies, think tanks, etc. The materials to be reviewed are important, but it would be helpful to clarify first if agencies have developed the materials; and then clarify how and how often the materials are actually used. It would be helpful to ascertain the utilization rate for various materials both with the agency (e.g., who are the agency audiences for various materials and are they using the materials) and outside the agencies (e.g., who are the non-agency audiences and are they using the materials.) Rate of utilization could be a powerful indicator for system effectiveness.	What source materials were used will be described in the evaluation report. All that is publicly available or submitted to the GEF is already noted. Upfront it is difficult to know what all material will become available. Assessing utilization rates of different types of report is beyond the scope of this evaluation. The focus of the report is on products that are prepared and submitted by the GEF Agency to the GEF.
13	<i>Review of a sample of annual project implementation reports, mid-term reviews, and implementation completion reports , along with relevant guidance will facilitate a comparison of the information being gathered through these tools and quality of information provided</i>	Sonja Teelucksingh: As mentioned in an earlier comment, this review should be expanded to cover more projects than only “two GEF supported projects” per Agency. PIRs are mentioned only twice in the paper, whereas they form the basis on which much of the reporting on project implementation is taking place.	This has been addressed in the response to an earlier question. PIR is an important product of the GEF Agency self-evaluation systems. Its importance is adequately recognized in both the draft approach paper and the revised paper.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
13	<i>Reports prepared by UNEG, ECG, MOPAN, and JIU, that cover at least some aspects of self-evaluation in GEF Agencies will also be reviewed.</i>	Sonja Teelucksingh: What about the World Bank study referenced earlier?	It is an assessment that will be useful for understanding the World Bank systems.
13	<i>Datasets: The evaluation will draw on different datasets maintained by the GEF IEO. This includes data on project performance and quality of reporting.</i>	Sonja Teelucksingh: I suggest that the datasets maintained by the GEF Secretariat also be consulted.	Change has been made in the revised paper to reflect the suggestion.
14	<i>GEF Secretariat staff involved in coordination of the self-evaluations at the GEF corporate level will be interviewed .</i>	Sonja Teelucksingh: Interviews should include GEF Secretariat staff that both coordinate and utilize the agency self-evaluation systems. For example, one part of GEF coordinates the requirements for performance indicators – but all GEF staff are users of the information provided by the Agencies, albeit in different ways. For example, the programs team can be asked if agencies have used their self-evaluation systems to improve capacity to generate global environmental benefits; whilst the operations team can be asked if agencies are submitted required implementation reports on time, operating within the relevant GEF policies, etc. What is the process for the selection of the interviewees?	This issue has been discussed earlier. A separate GEF IEO review on RBM system and the portal will address GEF Secretariat's role in detail. Its not necessary to discuss this in detail in the paper. We start with an assumption that within the GEF Secretariat there will be a team that will be involved in coordinating self-evaluation submissions, their analysis, and its use. The GEF IEO will request the GEF Secretariat management to provide us access to these resource persons.
14	<i>National counterparts will be interviewed to gather their perceptions on the performance of the Agency self-evaluation systems.</i>	Sonja Teelucksingh: Could you please provide more information regarding the interviewees? Will they include, for example, GEF OFPs? Executing agencies?	Information from the GEF OFPs, executing agencies, and CSOs, will be gathered through an online survey. Some of these may also be invited for detailed interview. However, these details need to be fleshed out further.
14	<i>Different modules will be developed to gather information from the different sets of interviewees.</i>	Sonja Teelucksingh: What do you mean by “modules”? Do you mean, questionnaires?	Yes, different interview questionnaires for different interviewee type.

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
14	<i>Online survey : An online survey will be conducted to gather perception on credibility and use of information provided by the self-evaluation system of the organization.</i>	Sonja Teelucksingh: Will the peer reviewers have the opportunity to review the questions, the mode of sampling and the sample composition across the 18 Agencies and the rest of the GEF partnership? Clarification is needed on the proposed participants in the survey. Are agency and country partners the only participants? What about GEF Secretariat, GEF IEO, GEF Council, GEF Trustee, conventions, and other users of agency M&E materials? A full stakeholder analysis may be helpful to inform both the interviews and the survey.	No. Advice from the peer reviewers is sought on broader design issues. Later their suggestions will be solicited when we present the emerging finding. Peer reviewers are not expected to contribute in detailed designing and execution of the evaluation. We will cover GEF Secretariat, OFP, and CSO network members through the online survey. Response rates to the online surveys are usually in the range of 5 to 10 percent. Therefore, online surveys usually do not make sense where the universe of the respondents is small – so some categories of relevant respondents may not be targeted by it.
15	<i>Workshops : Two workshops are planned. The first workshop will be to kick off the evaluation and to gather information from key informants from GEF Agencies. Subsequently, towards the end of the evaluation, a workshop with participants from the GEF Agencies will be conducted. The aim of the second workshop will be to share the preliminary data, to interpret the observed patterns and explore the reasons for emerging findings.</i>	Sonja Teelucksingh: The workshops appear focused only on agency staff, yet other stakeholders are listed in the review section – should these not be added to the workshop section?	Yes, the workshops are focused on those involved in conducting self-evaluations and are the one that are expected to take corrective actions. The GEF Secretariat and other stakeholders would also be invited to participate.
15	<i>Despite their support, it may still be difficult to execute all the planned activities of the evaluation given the level of complexity in the required coordination .</i>	Sonja Teelucksingh: And what are the contingency plans?	It's a risk that we are aware of. It may lead to delays or dropping of some activities.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
15	<i>The information may be used to identify issues that are of concern and need to be explored further through interviews and focus groups .</i>	Sonja Teelucksingh: A focus group discussion among key stakeholders could be very valuable. This is not an avenue often (ever?) utilized by the IEO.	No change suggested or required.
15	<i>The peer reviewers will provide feedback on the draft approach paper, the intermediary products , and the draft report of the evaluation.</i>	Sonja Teelucksingh: It would be useful to list what will be the intermediary products of the exercise.	Analysis on specific topics – which eventually get incorporated in the draft report of the evaluation.
15	<i>The draft approach paper of the evaluation will benefit from the feedback from the key stakeholders. While the first workshop is planned as an information sharing and gathering event, the second workshop will provide an opportunity to the key stakeholders such as the GEF Agencies (operations and evaluation), the Secretariat, STAP, and the CSO Network , to provide feedback on the emerging findings of the evaluation.</i>	Sonja Teelucksingh: The text implies a workshop will inform the draft approach paper. Is this a typo? The schedule does not include a revised approach paper after the workshop. If this is indeed a typo, then when will these key stakeholders be targeted for feedback on the draft paper?	This is not relevant now. Due to the pandemic there has been a reorientation of the purpose of the workshops. Purpose of different workshops has been explained in the revised text.
15-16	<i>While the first workshop is planned as an information sharing and gathering event, the second workshop will provide an opportunity to the key stakeholders such as the GEF Agencies (operations and evaluation), the Secretariat, STAP, and the CSO Network , to provide feedback on the emerging findings of the evaluation.</i>	Sonja Teelucksingh: The CSO network as a stakeholder has been absent from the paper until this point. Can you expand on if they are indeed considered a stakeholder in this exercise and, if so, what is the plan to engage them?	They are a stakeholder and their engagement is primarily through information sharing and consultations. Given the focus is on the subject of Agency self-evaluation systems, the narrative gives more attention to the GEF Agencies and their perspectives.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
17	<i>Table 1. Schedule of work activities</i>	Sonja Teelucksingh: Can we have a revised timeline, in light of the new era that is upon us? Will the report still be ready for the December 2020 Council? Can there be explicit mention of engagement with the Secretariat in terms of interviews, participation in workshops, and a review of the draft report? When will the Agencies get to review the draft report?	Revised timeline is included in the revised paper.
Garrett Kilroy, Senior Evaluation Specialist, Asian Development Bank, and Peer reviewer for the Evaluation			
2	The expectations from the self-evaluation systems of the Agencies are outlined in several GEF policy documents and policies of the GEF Agencies.	Garrett Kilroy: The reader would benefit a more comprehensive definition of what is understood in this evaluation by this term – and what are its component parts. In addition, throughout the draft sometimes there is reference to ‘self-evaluation systems’ and sometimes ‘self-evaluations’, the latter could be inferred as the end products, i.e. completion reports, whereas the former infers a wider more comprehensive, full project cycle. Finally, sometimes M&E is cited, which also needs to be put in context of the wider self-evaluation system. A short section describing the component parts of the self-evaluation systems may help contextualize each of these elements.	Agreed. An explanation of what constitutes the self-evaluation system has been added.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
3	 there are variations in self-evaluations needs and practices of the GEF Agencies given the differences in their mandates, scale of operation, level of independence of their evaluation function, and their self-evaluation traditions. Therefore, their self-evaluation practices for GEF supported activities may vary.	Very insightful. It is something that may influence the level of attention needed during the evaluation for differing agencies. Agencies with long established policies, practices and independent evaluation functions, may not warrant as much attention as agencies without such policies or independent evaluation functions.	Thanks for the endorsement. No action required.
3	The GEF Council and the Secretariat are interested in ensuring that the self-evaluation systems of the Agencies monitor their GEF portfolios well, facilitate learning, and are harmonized.	Garrett Kilroy: (On learning): And accountability? (On harmonization): Is this a reasonable expectation, given the variability of capacities and mandates of the Agencies outlined in the previous para?	Accountability added along with learning. Reference to harmonization removed because of restructuring of the paragraph. However, harmonization remains an expectation not only of the GEF Council but of the larger development community. It is not a question that needs to be seen in binary terms but as a gradient and based on what is feasible and realistic.
3	On literature review	Garrett Kilroy: This is an excellent review. The identified lack of studies on M&E in international development organizations is an important lacuna. If the evaluation, can progress this aspect, even to better harness what is available, we will all benefit.	Thanks for the endorsement.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
6	Several GEF Agencies already assess performance of their self-evaluation systems, although such assessments are usually limited in their scope. For example, evaluation units of several GEF Agencies such as UNDP, UNEP, and IFAD assess and report on quality of self-evaluations through their annual reports. However, in these reports' coverage of topics such as candor in reporting and learning is not detailed.	Garrett Kilroy: At the project level, this may be something that could be captured for GEF financed projects in the Agencies. IED, for example, provides a rating for completion report quality in our validations. We have criteria in our guidelines for this assessment. Candor and quality of lessons forms part of the criteria. But, I agree, the assessment may not cover these topics in detail.	No action required.
6	Behind the Mirror: A Report on the Self-Evaluation Systems of the World Bank Group' stands out as an exception to the rule.	Garrett Kilroy: Perhaps some key findings/lessons relevant to this evaluation could be mentioned. I am not sure if relevant for GEF, but we found the following report useful on the M&E side: 2013 Biennial Report on Operations Evaluation Assessing the Monitoring and Evaluation Systems of IFC and MIGA	Information on key findings and lessons included – not only for the World Bank report, but for reports that cover other Agencies. We have not reviewed the 2013 Biennial Report separately, because its findings and lessons are already incorporated in 'Behind the Mirror' report. However, it is definitely a document that will be covered during the conduct of the evaluation. Thanks.
7	APR regularly presents analysis of terminal evaluation quality and submission gaps.	Garrett Kilroy: For the outside reader, it would be useful to define this term to distinguish between an agency's self-evaluation and independent evaluation report.	The term has been defined in the revised approach paper.
7	APR 2015 covered gaps in submission of tracking tools by the GEF Agencies. However, GEF IEO is yet to conduct a comprehensive assessment of the self-evaluation systems of GEF Agencies.	Garrett Kilroy: Has the GEF tracking tool ever been evaluated?	Tracking tools have been covered in detail OPS-5 and in APR2015.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
7	The GEF Independent Evaluation Office (IEO) is undertaking the 'Evaluation of the Agency Self-Evaluation Systems' to assess the extent to which Agency self-evaluation systems meet the GEF requirements and provide information that is sufficient, timely, credible, and useful.	Garrett Kilroy: This sentence reads like the evaluation's over-arching question. But it should be made clear what GEF requirements are. There is no escaping this characteristic of credibility, all the agencies must grapple with it. But there may be resistance where a third party makes a judgement and the evaluation turns into a benchmarking exercise. In ADB our approach was to make a distinction between the credibility of the design of the self-evaluation system, i.e. the architecture of its component parts and associated policies, and the credibility of actual implementation of the system. Credibility we found was more of a higher level principle than the others listed, which are more objective and easier to measure.	GEF requirements has been made clearer in the narrative. The broad areas where there are specific requirements have been mentioned. However, each specific requirement is not listed. We focus on how the self-evaluation systems apply to the GEF supported activities, and credibility of these arrangements and the information on GEF activities. We don't intend to evaluate the self-evaluation systems of Agencies in its entirety and similarly benchmarking is with an intent to facilitate better understanding of the systems and not to identify Agencies which have the best systems and so on.
7	The evaluation will assess the factors that affect these systems, and identify areas for improvement.	Garrett Kilroy: Looking beyond identifying areas for improvement, what can be reasonable channel for the results? Is there a GEF review or event that the report could influence?	The evaluation will rely on collaborative data analysis and information sharing as a mechanism for improving system performance. Workshops and focus groups undertaken as part of the evaluation will provide a platform to do so.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
7	The evaluation will be based on a theory of change presented in Figure 1.	This is useful device to communicate the conceptual framework. One query is whether GEF should be identified within this framework? Or is there a reason the TOC decouples GEF's own self-evaluation systems from that of the agencies? Will all five boxes fall within the remit of the evaluation?	Figure 1 is generic and may be extended to the GEF as well. We feel that making it more detailed makes it more accurate, but we lose on simplicity. We have adopted a parsimonious approach in developing the framework.
7	A well performing system (intermediate outcome) is expected to lead to learning and accountability (final outcomes).	Garrett Kilroy: On learning: (and decision making?)	Decision making is implicit in both the terms – learning and accountability
9	While information generated by a self-evaluation system may be regarded as the key system output, information quality may be regarded as its key intermediate outcome. Merits of the information generated by the system may be assessed on dimensions such as comprehensiveness, timeliness, candor and credibility, accessibility, and utility.	Garrett Kilroy: Some explanation of the different types of information generated in a typical self-evaluation system is needed. I am not sure the credibility dimension fits at the same level as these other	A brief discussion on different information products from the self-evaluation systems has been added. Credibility is addressed both at the level of arrangements for self-evaluation of GEF supported activities and at the level of self-evaluation products. We do agree that different evaluators may choose to address it differently.
9	If the knowledge generated by the self-evaluation system is relevant and covers important areas of institutional performance, it may be regarded as comprehensive. An effective self-evaluation system will track what is important and track it well without overburdening the organization. Comprehensive coverage of issues that are of concern may be expected to facilitate learning and accountability.	Garrett Kilroy: Information and knowledge need to be differentiated. This is a very good point as we find that government capacity to effectively implement a project M&E system can be an issue. You may need to distinguish between the centralized M&E systems of the Agencies and the individual project level M&E systems implemented by the borrowing government agency.	Where feasible, this distinction has now been made. This distinction is imbedded in the evaluation. The focus of the evaluation is only on the Agency self-evaluation systems. It has been made clearer that the centralized system coordinated by the GEF Secretariat will be covered through another review on RBM.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
9	Candor and credibility of information provided by the self-evaluation system is an important dimension of its effectiveness: the higher the level of candor and credibility, the greater will be the trust in self-evaluations. Such evaluations may be expected to facilitate learning. While candor in reporting may be useful for accountability at the higher scales, it might be disincentivized at lower scales due to repercussions from reporting issues or concerns.	Garrett Kilroy: Project completion reports? Or the whole self-evaluation system? Not clear what is meant by higher and lower scales	The paragraph has been revised. It discusses only candor of reporting. The issue of credibility of the self-evaluation system is addresses only as it relates to the self-evaluation of the GEF supported activities. The term higher and lower scale has been removed in the revised paragraph.
10	Outcomes and performance of a self-evaluation system may be affected by several factors such those related to system design and implementation, organizational characteristics, and broader context.	Garrett Kilroy: There is one factor missing here – the capacity and interest of the government executing and implementing agencies. In ADB it is the borrower that must establish and maintain the project M&E system throughout the project, following ADB's guidelines and reporting regularly. The government also prepa10res a draft of the completion report. The M&E system and the government completion report are key inputs to the ADB completion report.	Thanks for the input. We have included reference to the capacities of the executing agencies in the revised narrative.
11	Provision of adequate resources to self-evaluations is important. Lack of (staff) time and budget affects implementation of a self-evaluation system. Under resourced systems are unlikely to ensure quality, timeliness and accessibility of the information generated.	Garrett Kilroy: Resources are needed through the project cycle, not just at time of preparing the completion report.	Agreed. The revised text discusses the system not just the products.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
11	An independent evaluation unit may build capacities for self-evaluation by providing guidance and training, and by providing feedback on the quality of self-evaluation. Self-evaluation, on the other hand, may be a source of quality data for independent evaluations. We may expect the self-evaluation system to benefit from a well-functioning independent evaluation system.	Garrett Kilroy: Absolutely, the better the self-evaluation system is the better independent evaluation can be. Absolutely, and the corollary is also very true – the better the self-evaluation system is the better independent evaluation can be.	Agreed. No action required.
11	Key Questions and Hypothesis The evaluation aims to answer the following questions:	Garrett Kilroy: Again, may need to clarify whether these questions concern the end-products of the self-evaluation system (i.e. completion reports) or the system as a whole.	Both system (as it relates to GEF activities) and products of the system are addressed in the questions. The questions have been made clearer to indicate which aspects pertain to products and which to the system itself.
12	The evaluation will assess how GEF Agencies address the credibility of information generated by their self-evaluation system.	Garrett Kilroy: In this context, perhaps it is the quality of information	Agreed. But because of the changes in the paragraph the suggested change is not necessary.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
12	To what extent are the Agency self-evaluation systems meeting the needs of GEF partnership? The evaluation will record perceptions of the Agency staff, national counterparts, and consultants, on the extent to which the Agency self-evaluation systems are effective in supporting the learning and accountability needs of the GEF partnership. It will also assess effectiveness by determining the extent to which information provided by the system is comprehensive, timely, credible, accessible, and useful, and in line with the GEF requirements.	Garrett Kilroy: These needs/requirements need to be clearly articulated early in the paper. This infers a type of performance assessment – it may need to be tailored to the different agencies to reflect their varying organizational setup, mandates and goals	As discussed earlier, the types of requirement have been noted in the paper although specific requirements are not listed. Yes, the paper intends to assess the systems considering their respective characteristics.
12	What are the factors that affect effectiveness of the self-evaluation systems?	Garrett Kilroy: The capacity and interest of the government executing and implementing agencies may be a factor affecting effectiveness	Agreed. As discussed earlier, this point has also been incorporated.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
13	Evaluation Design The evaluation will use a multiple-case design and cover all the GEF Agencies (Yin 2018). Self-evaluation system of a GEF Agency – as it related to the GEF supported activities – will be the unit of analysis. For each of the Agencies, two GEF supported projects will be selected to assess operation of the self-evaluation system at the project level.	Garrett Kilroy: This is section is quite short and the reader is left wondering how project level assessments will feed into the desk review and other sources of information to address the key evaluation questions. Aside from the project assessments, could the whole portfolio for each agency be examined for relevant metrics that are available, e.g. completion report quality? The component parts of a typical system could be outlined here or earlier in the TOC It may be useful to look at both closed and active projects. The closed projects will allow you to unpack the quality of the completion report and elements of the system contributing to it. However, the closed project could 4+ years old and may not represent current self-evaluation systems. Looking at a project in the active portfolio may facilitate this more up-to date assessment and real-time examination of the M&E systems.	Agreed, more details have been added to the section. Some analysis of the entire GEF portfolio of the Agencies is presented in APR 2020. Where already available data allows, this will be done. Agreed, System components have been outlined in the introduction section. We will look at about to be completed and recently completed projects.
13	Datasets: The evaluation will draw on different datasets maintained by the GEF IEO. This includes data on project performance and quality of reporting.	Garrett Kilroy: Is this quality of the Agency completion report or quality of the independent evaluation unit's validation?	Its of the Agency completion report, and excludes validation.

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
Michael Spilsbury, Director of Evaluation Office, UNEP			
NA.	On renaming the evaluation and on use of the term 'self-evaluation'	Michael Spilsbury: This evaluation should be re-titled, perhaps as follows: IEO Evaluation of GEF Agency Arrangements and Systems for Evaluation Or Evaluation of Agency Arrangements and Systems for Evaluation in the GEF Partnership The suggestion to change the title is made because the current approach paper <u>incorrectly</u> assumes all evaluation of GEF projects across the Agencies can be considered 'self-evaluation'. <u>This is most certainly not the case</u> and should not be an embedded assumption. The evaluation arrangement in several GEF agencies do not meet the definition of 'self-evaluation'. E.g. UNEP, FAO, UNIDO and probably others. <u>I would strongly assert that the concept of 'self-evaluation' is not at all compatible with the UNEG norms and standards of Independence, Credibility and Utility and therefore is not a useful starting point for this evaluation</u>	The evaluations conducted by the operations are called by different terms such as decentralized evaluation, auto evaluation, and self-evaluations. Of these, the term self-evaluation is used most commonly. Both suggested names enhance the scope of the evaluation and would require us to assess performance of the Agency evaluation offices as well – which the evaluation does not intend to do. We have clearly defined what the term self-evaluation means in the context of this evaluation. For the purposes of this evaluation we also include the situations where the evaluation office of an Agency commissions or conducts terminal evaluation.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
NA.	'Self-evaluation system of a GEF Agency – as it related to the GEF supported activities – will be the unit of analysis.'	Michael Spilsbury: In my view, the approach paper should be looking to describe the evaluation arrangements across agencies in a systematic way, it should not place such an emphasis on labelling evaluation systems and arrangements as 'self-evaluation' from the outset. The evaluation should be exploring the characteristics of the GEF Agency evaluation arrangements with respect to international standards for evaluation. A set of standards should rather form the framework against which Agency arrangements are examined. Adopting a process perspective by agency would be very illustrative and would allow IEO to consider whether the Agency arrangements and processes conform to the international standards mentioned. In this regard the approach paper should summarise the relevant DAC, UNEG, ECG, and MDBs standards referred to and be explicit about which standards will be selected, - Agency evaluation arrangements should then be assessed against them.	The suggestion is welcome. We will be doing this in the evaluation. The focus indeed will be on systems and processes for GEF supported activities, and less on products. Agency arrangements for reporting on GEF activities will be considered, including arrangements where the evaluation office is involved in conducting terminal evaluations. The standards and minimum requirements noted in the GEF Evaluation Policy (2019) will be followed.

NA	On application of standards noted in GEF Evaluation Policy.	Michael Spilsbury: “Evaluation in the GEF context is guided by internationally recognized principles, norms, and standards. Specifically, the GEF and its Agencies refer to those principles, norms, and standards produced by the United Nations Evaluation Group (UNEG), the Development Assistance Committee of the Organization for Economic Co-operation and Development (OECD-DAC), and the Evaluation Cooperation Group (ECG) of the Multilateral Development Banks (MDBs). Although there is general agreement around internationally recognized norms and standards, there is also a divergence resulting from the diverse goals and objectives of the individual Agencies. These differing goals lead to differences in emphasis and differences in the application of standards across Agencies. Guided by international norms and standards, the GEF Evaluation Policy considers these differences and establishes a set of key principles and criteria common across the GEF partnership. The Policy also establishes four mandatory minimum requirements Agencies must follow in conducting evaluations for GEF-financed activities.” The Approach paper should clearly explain how these standards are to be applied in the evaluation and how these will be used to explicitly identify the <i>‘differences in emphasis and differences in application of standards across the Agencies’</i> are manifested. Whether or not the definition of self-evaluation vs evaluation proper are relevant to describe the evaluation arrangements across the partnership is of much less importance. I would expect that some of the issues mentioned in Minimum Requirement 3 of the GEF Evaluation Policy would be mentioned as aspects against which this evaluation would assess agency compliance.	As noted in the cited text, the GEF Evaluation Policy establishes key principles and criteria that are guided by the international norms and standards. The Evaluation Policy of GEF establishes four mandatory minimum requirements that the Agency must follow. The evaluation will assess the extent to which this is being done. The work on compliance with minimum requirement is being deepened. Some of it has already been presented in APR2020.
----	---	---	--

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
NA.	Criteria	Michael Spilsbury: The approach paper should also refer to the information used to accredit GEF agencies (evaluation arrangements feature among accreditation criteria). Presumably IEO has access to the Agency level information? https://www.thegef.org/sites/default/files/council-meeting-documents/C.40.Inf_.04_Draft_Procedures_Manual_For_Accr_editing_GEF_Agencies.May_10_4.pdf See Annex 2 <i>18. Are Independent Evaluations undertaken by an established body or function as part of a systematic program of assessing results, consistent with the requirements of the GEF monitoring and evaluation policy?</i> <i>19. Does the evaluation function follows impartial, widely recognized, documented Annex 2 25 and professional standards and methods?</i> <i>20. Is the evaluation body or function structured to have the maximum independence possible from the organization's operations, consistent with the structure of the agency, ideally reporting directly to the board of directors or comparable body? If its structural independence is limited, the evaluations body or function has transparent reporting to senior management.</i> <i>21. Is an evaluation disclosure policy in place? Evaluation reports are disseminated as widely as possible, and at a minimum to all parties directly or indirectly involved with the project? To enhance transparency, to the extent possible, are reports made available to the public?</i>	Where appropriate such references have been added. The evaluation is of the self-evaluation system – the assessment of the evaluation body or function is based on as it relates to the self-evaluation system. While the criteria used for accreditation is useful, it is not sufficient because it does not enquire in depth as to how the system is functioning in delivering reporting of high quality on GEF supported activities.

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
NA.	On the extent to which the term 'self-evaluation' appears in the GEF Evaluation Policy (2019).	Michael Spilsbury: It is interesting to note that the term 'self-evaluation' does not appear either in the 2019 GEF Evaluation policy or in the "Guidelines for GEF Agencies in Conducting Terminal Evaluation for Full-sized Projects". Yet it has been embedded as assumption in all the key questions posed in the approach paper. Why? • <i>How do GEF Agencies address self-evaluations through their policy framework?</i> • <i>What arrangements are in place in Agencies to conduct self-evaluations?</i> • <i>To what extent are the Agency self-evaluation systems meeting the needs of GEF partnership?</i> • <i>What are the factors that affect effectiveness of the self-evaluation systems?</i>	Thanks for the observation. The Evaluation Policy may not include the exact term (of self-evaluation), but products of self-evaluation are noted, and Agency role has been described. In any case, non-inclusion of a term in the evaluation policy does not mean that the GEF IEO cannot conduct an evaluation on that topic.
13	On Evaluation design	Michael Spilsbury: The section on evaluation design in this approach paper needs to be significantly strengthened. <i>"The evaluation will use a multiple-case design and cover all the GEF Agencies (Yin 2018). Self-evaluation system of a GEF Agency – as it related to the GEF supported activities – will be the unit of analysis. For each of the Agencies, two GEF supported projects will be selected to assess operation of the self-evaluation system at the project level.."</i> A much more detailed articulation of how the analysis in these case studies will be conducted should be presented. Specifically, it needs to identify the standards against which agency evaluation arrangements will be assessed and how this assessment will be undertaken. It is unclear how the analysis of 'two GEF supported projects' will be used.	Thanks for the suggestion. The section on evaluation design has been strengthened. The sampled projects will be used to understand and illustrate the self-evaluation process at the project level in different Agencies. This has been clarified further.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
8-10	On Theory of Change discussion	Michael Spilsbury: The sections on 'Factors that affect the system performance (the Xs')' and System Effectiveness (The Ys) are rather theoretical and it is not clear how this discussion relates to the proposed evaluation method.	The discussion is indeed theoretical. It lays out the key relationships and mechanisms through which independent variables lead to the outcomes – dependent variable. It helps in identifying what needs to be captured through various instruments deployed by the evaluation.
NA	On UNEP's evaluation arrangement meeting the OECD definition.	Michael Spilsbury: Just for the record. UNEP's evaluation arrangement meets the OECD- definition of evaluation " <i>evaluation carried out by entities and persons free of the control of those responsible for the design and implementation of the development intervention</i> " UNEP's GEF Evaluations are planned, managed and quality-assured by the Evaluation Office of UNEP. UNEP has no decentralized evaluation function. UNEP does not perform 'self-evaluations' of GEF projects.	Thanks for statement. The revised paper acknowledges this type of arrangement for GEF activities and brings it within the ambit of the evaluation.
NA	Comparison of UNEP and UNIDO processes for conduct of terminal evaluation.	Michael Spilsbury: For comparison I have also looked at UNIDO's evaluation manual (p43) to prepare an example of the GEF TE process and also attach the same process table for UNEP (see below). This is intended to be food for thought for this evaluation. (<i>A table comparing the UNEP and UNIDO arrangements for terminal evaluations was attached with the comments</i>).	Thanks – this is very useful. We may use it for reference during the conduct of the evaluation.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
NA	Through email – a separate comment on cost of terminal evaluations.	Michael Spilsbury: Another thought, it might make sense for the IEO to also look into the levels of financial support for the direct costs and indirect costs for TEs across the agencies in the study of evaluation arrangements in the agencies. In other words: How do the agencies budget for the direct costs of Mid-Terms and TEs in GEF projects? (and has this changed over time?) How do the Agencies use GEF fee in support of the part of 'project cycle management' that deals with evaluation? (this would be quite a variable picture I would assume) However, both of these factors may affect the quality of GEF TEs in the agencies, along with the organizational arrangements and procedural approaches I mentioned before. At the very least it would be useful to find out the current norms across the agencies in this regard and IEO might be able to identify some trends / good practices.	Thanks for these inputs. These will be addressed in the evaluation. The question on factors that affect self-evaluation system effectiveness discusses resource sufficiency as an issue to be covered.
Comments by IADB (sent by Annette Bettina Killmer)			
NA	General remarks.	IADB: 1. We commend you on conducting this evaluation, which will surely make an important contribution to the M&E practices of the GEF Partnership. 2. Overall, the Approach Paper is clearly structured and well argued, with a strong logic and solid support from evidence and the literature. We have no comments on the overall theory of change or approach. However, we would like to raise a number of issues (points #3 to 7) for your consideration that, we believe, would further strengthen the approach and its resulting evaluation.	Thanks for the general endorsement for the approach explained in the paper. We have considered your inputs on the specific issues raised by you to revise the paper.

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
NA	On distinguishing between the self evaluation system elements managed by the GEF Secretariat and those by the GEF Agencies.	IADB: 3. In assessing the extent to which Agency self-evaluation systems meet the GEF requirements, and provide information that is sufficient, timely, credible and useful, we consider it essential to make a clear distinction between the system each Agency uses to evaluate the performance of GEF projects, and the system the GEF uses (i.e. the GEF Portal and other information requests) to collect information from Agencies on projects performance and other aspects. Judging from the “Key Questions and Hypothesis”, the approach paper, when it refers to “Agency Self-Evaluation System” refers to the each Agency’s internal system – not the elements of that system that get reported to the GEF through the Portal. However, that should be explicit in the draft approach paper, including in the third Key Question, which is posed as if the GEF partnership had full access to each Agency’s self-evaluation system, which is far from the reality. IADB for example, has a very comprehensive system of M&E that it applies to projects (especially those that finance investments in addition to technical assistance), but since the GEFs Portal is not compatible with Agencies' internal systems and information has to be transferred manually from one system to another (incidentally, this issue speaks not only to the ‘system performance’ but also to the questions of “accessibility” raised in the approach paper), only a fraction of the information that the Agency has at its disposal gets reported to the GEF.	This distinction has been made clearer in the revised paper. The focus of this evaluation is on GEF Agency Systems and not the systems and activities managed by the GEF Secretariat. It has been clarified that a separate evaluation on GEF RBM system is covering the centralized systems and arrangements at the GEF Secretariat.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
9	On whether information quality ought to be treated as an output or as an intermediary outcome.	IADB: 4. (Pg. 9, 2nd paragraph) Information quality should probably be a key system output, as it is something that is very much under the control of the project teams and executing agencies reporting that information. A corresponding (intermediate) outcome could be the useful and timely synthesis of that information. (“Useful” in the sense that, for the GEF Sec and the GEF Partnership as a whole, a portfolio-level perspective/synthesis/analysis would likely be more useful for their decision making. Obviously, for the implementing Agencies and executing agencies, detailed analyses at the project-level are pivotal for decision making; but likely surplus to needs for the GEF Sec.)	This is one of the transition elements that can fit in both categories – as outcome, and as an output – based on how the term is defined. We use the term more as an intermediate outcome because we find that even though availability of information may be ensured its quality is difficult to ensure and is not a given (because of factors internal and/or external to the GEF Agency).

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
NA	On difference in usefulness of the self-evaluation for different stakeholders.	IADB: 5. On a related point to the one just made about “usefulness”, the draft approach paper highlights the link between decision making, timely information and the effectiveness of a self-evaluation system. We agree with the approach paper on this point, and would like to add that this link raises the question ‘What are the decisions that GEF Sec and OPF’s have to take? And hence what information is ‘sufficient, timely, credible and useful’? As compared to ‘What are the decisions GEF Agencies and their project teams have to take?’ The information required by different decision makers within the Partnership varies somewhat depending on their role, meaning that an Agency self-evaluation system may be effective for one decision maker in the GEF Partnership (e.g. the Agency itself), but not for another (e.g. GEF Sec). (For an additional layer of complexity, see comment # 3.) The approach paper should be clear on which decision maker(s) it is focusing in its assessment of effectiveness against “GEF requirements”.	The evaluation will cover all the three groups – Secretariat, Agencies, and OFPs. It will gather perspectives of these three (and other) stakeholders. Its true that a system may work well for one group but not so for the other.

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
13	On use of Agency as a unit of analysis.	IADB: 6. Please keep in mind that the unit of analysis – the “self-evaluation system of a GEF Agency” – with all its policies, quality assurance mechanisms, information sharing, resources, culture, structure, business model etc should be and most likely is designed for each GEF Agency’s entire portfolio of operations. By GEF policy, the share of GEF operations in that portfolio is limited, and in the case of MDB’s, GEF-supported activities tend to account for less than 1% of the portfolio. For other Agencies, that share is 10%, 20% or even higher. Perhaps it would be interesting for the evaluation to explicitly consider the share of GEF operations vis-à-vis an Agency’s total portfolio as one of the factors that affects the effectiveness from a GEF perspective of the Agency’s self-evaluation system.	The focus of the evaluation is on GEF Agency arrangements as they apply to GEF supported activities. The evaluation uses the GEF lens to assess these arrangements although arrangements for other Agency activities may be considered for comparison.
13	On sampling of projects for documenting the self-evaluation process.	IADB: 7. In the selection of the two sample projects, please keep in mind that different types of projects may have to comply with different internal requirements for self-evaluation. For example, in the case of IADB, technical cooperations are governed by a separate policy – and have a specific Monitoring Module in our operations system – from investment grants and equity (NGI) operations. Depending on which two projects are selected for the sample, the ‘snapshot’ GEF IEO will receive may be very different/ not representative.	The revised approach paper keeps the number of projects to be sampled flexible and bases the number on the number of different processes that are applied for self-evaluation of the GEF activities.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
NA	On sharing of the protocols and evaluation schedule.	8. We are certain you are already addressing this point in the revised approach paper, but mention it here for completeness' sake: given the new context, could you please include information in the approach paper about any changes you envision in the timelines, interview formats, evaluation activities, workshop formats etc. This would be most helpful to us in getting a head-start on the internal coordination/organization that this evaluation will require (e.g. defining members of the evaluation system team, management, IADB's independent evaluation office, etc.), especially in these unusual times and working arrangements.	A revised schedule has been included in the revised approach paper. The formats and other instruments are under development. The pandemic has thrown the evaluation schedule off gear. We will share more information as it becomes available.
Comments by Independent Office of Evaluation (IOE) of IFAD (Sent by Fabrizio Felloni)			
	On coverage of the Agency evaluation offices by the evaluation.	IOE IFAD: 1) Our understanding is that the evaluation will review the <u>entire</u> self-evaluation systems of the agencies, including IFAD, with special focus as it pertains to GEF. In the case of IFAD, we understand that this will relate to self-evaluation under Management's responsibility. It would not include IOE's independent evaluation. Is this correct?	The focus of the evaluation of the Agency self-evaluation system. However, the evaluation will touch upon the linkages of the self-evaluation system with the independent evaluation system.
	On whether operation units are informed about the evaluation.	IOE IFAD: Your office has probably been in touch with the Operational Policy and Results Division of IFAD (responsible to coordinate self-evaluation) and they will be the main stakeholders. Grateful for confirming it. We can also provide you with their contacts if needed.	The draft approach paper was shared with the official contacts of the Agency operational units and Agency evaluation offices.

GEF IEO Evaluation of the Agency Self-Evaluation Systems
Paper trail of feedback and action taken

Page	Context in which comment is made	Comments	Response by GEF IEO Evaluation Team
	On work done by the IOE of IFAD on assessment of the performance of the IFAD's self-evaluation system.	IOE IFAD: IOE remain, nonetheless, available for any inputs that you may require. IOE in the past conducted a quality assessments on the evaluation reports commissioned by IFAD on their GEF-funded projects. We have not received any request to do so since the second part of 2019. While IOE did not conduct a dedicated evaluation of the IFAD self-evaluation systems, this topic is discussed in the Annual Report on Results and Impact of IFAD Operations (ARRI). We also analyze selected aspects of self-evaluation in our project-level and country-level evaluations.	Thanks, this will be a useful input.
	On relevant scholarly paper that may be relevant for the evaluation.	IOE IFAD: Among the references in the paper, you may also consider Mayne, J. (2010). Building an Evaluative Culture: The Key to Effective Evaluation and Results Management. <i>The Canadian Journal of Program Evaluation</i> Vol. 24 No. 2. From a substantive point of view, this contribution highlights the importance of leadership drive in an organization, the existence of explicit feedback mechanisms from findings to design of new programmes and strategies and steering of the on-going ones. It also highlights the importance of going beyond 'formalism', recognizing that opportunities always exist for 'gaming the system'.	Thanks for suggesting this paper. We contacted the author and acquired a more updated version of this paper. The paper is cited in the revised approach paper.