

Draft Report of Evaluation of the GEF Agency Self Evaluation Systems – feedback and action taken

June 10th 2021

The draft report of the Evaluation of the GEF Agency Self Evaluation Systems was shared with the key stakeholders – including the GEF Secretariat and GEF Agencies – in May 2021. This document presents an account of the feedback received from the key stakeholders and how it was addressed in finalization of the report.

The draft report was shared with the key stakeholders as a standalone document and revised based on the feedback. The feedback was generally received in form of email messages and as attached documents. In this document only substantive elements of the feedback from the Agencies are presented. The feedback from the GEF Secretariat was communicated to the evaluation team during a meeting – the feedback did not call for any change in the report and is, therefore, not tracked in this document.

The revised report was then integrated as Part A of a combined report that presented two pieces of work related to results-based management, i.e. evaluations of GEF Agency Self Evaluation Systems (Part A) and GEF Portal (Part B). The combined report will be presented to the GEF Council as a Working Document at its June 2021 meeting¹.

Feedback and Evaluation Ton Draft Report of Evaluation

S. No.	Feedback from key stakeholders	GEF IEO evaluation team response
1.	Amelie Solal-Celigny, Evaluation Officer, FAO Office of Evaluation	
1.1	While FAO's "stakeholder approach" and "highly consultative process" are highlighted as good practices, in table 2 FAO's stakeholder consultation is rated as non-compliant. Could you please explain?	Change For greater clarity the relevant category has been split into two. This makes it clear that the non-compliance is related to lack of sharing with the GEF Operational Focal Point. The terminal evaluations are generally compliant as far as sharing with other stakeholders is concerned.
1.2	Para 30 says FAO Evaluation Unit is not "structurally independent of management", but is "functionally independent".	No Change

¹ https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.E_C60_07_RBM_SES_Portal_Combined_Report_FINAL.pdf

	We suggest to remove the wording “not structurally independent of management” or explain further. OED is a unit independent from any line department and the Director of evaluation reports primarily to the programme committee. The reporting line to the DG is indeed there but the special recruitment process of the director and the reporting to our GB ensures a fair degree of independence.	We believe that the reporting is accurate. The term functionally independent is well understood and consistent with the explanation provided in the feedback. Therefore, there is no substantive disagreement. The additional explanation specifically for FAO is not necessary as the situation is not unique to FAO.
1.3	Para 44 says that FAO conducts impact evaluations for a sample of projects several years after their implementation. This is not exact: OED does not conduct impact evaluations after years of implementation for projects. The OED note on Impact Evaluations explains that “OED conducts impact evaluations, assessing changes by FAO interventions on people’s well-being and livelihoods as well as on their environment. The focus of OED impact evaluations is at the thematic/programmatic level, covering areas with substantial volume of FAO work that also support organizational learning, rather than at individual projects. This is important given the nature of its core functions, where in general it is difficult to attribute impact to a single intervention/project. Geographically the evaluations are aimed to cover countries where FAO has a large portfolio of work. The overall intent is to determine whether the organization has contributed to change and impact in a meaningful line of causality”.	Change The language of the relevant paragraph has been revised to make it more accurate and consistent with the current practice at FAO.
1.4	Para 45 is not completely correct: FAO OED used to manage directly MTRs until 2018, after which they have been decentralized and OED has no role. Between 2018 and 2020 OED managed some MTRs upon exceptional request.	Change The narrative has been made consistent with FAO’s current practice.
1.5	Box 3, page 13 (dated 2015...?) mentions Reference Groups as part of the QA process. Reference groups are a consultative mechanism and not part of the QA process, which is done by OED.	Change The text has been revised.

2.	Margarita Arguelles, Results and Knowledge Specialist, UNDP	
2.1	Paragraph 9 states that, “The GEF core indicators are specified in the guidance provided by CAF, FAO and FECO.” Please note that GEF Core Indicators are also specified multiple times in the Guidance on Conducting Terminal Evaluations for UNDP-supported GEF-financed Projects and also in the TE TOR template for UNDP-supported GEF -financed projects.	Revised UNDP has also been added as an example in the relevant para.
2.2	Typo: “It is still to early to...” Change “to” to “too”	Changed
2.3	Paragraph 39 states, “GEF Agencies use a variety of quality assurance practices. GEF coordination units support all self-evaluation procedures and provide support for quality assurance. All GEF Agencies, except for UNDP and IFAD, have set up such a coordinating unit.” Please note that there is a GEF coordination unit - led by Executive Coordinator, Pradeep Kurukulasuriya - within the Nature, Climate and Energy (NCE) team in UNDP’s Bureau for Policy and Programme Support (BPPS). This unit issues guidance and quality assurance support on the entire GEF project cycle, including on self-evaluation systems. For example, this unit coordinates the annual PIR process (detailed guidance, training, hiring of an external consultant to QA submitted PIR reports). NCE region-based Technical Advisors provide quality assurance support to COs and project teams for each PIR. A dedicated RBM team within the GEF coordination unit (internally referred to as the NCE Vertical Fund Unit) also issues guidance, templates and provides support for the GEF MTR/TE processes. The RBM team quality assures draft GEF Core Indicators and, upon request, reviews the quality of draft MTR and TE reports. This coordination unit also manages the Project Information Management System (PIMS+) platform which is used by the NCE team to oversee GEF project milestones and monitor GEF project performance. The PIMS+ Risk Dashboard identifies and displays projects with implementation delays or challenges as monitored and reported in GEF PIRs, through spot checks and through independent MTRs and allows senior managers to identify and manage projects that require enhanced oversight to get back on track. Serious	Changed. Thanks for sharing detailed information on this specific point.

	performance issues can be escalated to NCE and bureau-level senior management, Regional Bureau and the Country Office. The PIMS+ platform is programmed to send out automatic notifications to NCE staff and Country Office staff well in advance of MTR and TE deadlines for example and also when MTRs/TEs are overdue.	
2.4	Paragraph 40 states, “World Bank’s coordination unit does not make similar arrangements for validation of terminal evaluations of projects below the funding threshold used by their evaluation unit for validation. In IFAD the evaluation unit provides feedback on the drafts of a few terminal evaluations and validations are conducted by another unit that is not independent of the management. UNDP’s evaluation unit also validates terminal evaluations, but does not apply a project funding threshold for validation.” It may be useful for the reader if the text clarified that the UNDP IEO validates all GEF terminal evaluations (which is why there is no project funding threshold for validation).	Changed Relevant text has been revised for greater clarity.
2.5	Paragraph 41 states, “Evaluation units in UNEP, UNIDO and FAO, manage the evaluation process including providing quality assurance. Of these, the GEF IEO accepts the project performance ratings provided by UNEP’s evaluation unit – for UNIDO and FAO the GEF IEO conducts terminal evaluation validations.” It might be useful to also mention that the GEF IEO accepts the validated project performance ratings from the UNDP IEO.	No change. The information is already provided in the paragraph that discusses evaluation offices that are structurally independent and are not involved in conduct of terminal evaluation.
2.6	The evaluation states that, although Agencies use the same rating criteria, there are differences in what is addressed within each criterion and how it is applied. Given this statement, it would be useful for the GEF IEO to issue guidance on the required evaluation criteria to promote consistency in TEs across Agencies.	No change Thanks for the input. This is an important issue for the future. We will see how we could address it through better guidance.
2.7	UNDP is aligned with the GEF requirements except for relevance. “ Please refer to the following UNDP-issued documents which show that Relevance is rated on a 6-point scale:	Changed.

	- Table 9 on page 36 in the Guidance on Conducting Terminal Evaluations for UNDP-supported GEF-financed Projects - TOR Annex F in the TE TOR template for UNDP-supported GEF - financed projects	
2.8	On the topic of investing in training project management staff, we would like to point out that the UNDP IEO developed an evaluation certified training course which is mandatory for all UNDP staff whose work involves M&E.	Changed UNDP is now included as an example.
2.9	On the topic of evaluation awards, please note that the UNDP IEO recently announced the winners of the 2020 inaugural Evaluation Excellence Awards. Additional information can be found here: http://web.undp.org/evaluation/award/index.shtml	Changed Provided as an example.
2.10	Table 1. This table shows that the 'role of GEF core indicators in evaluation' is missing for UNDP. Please note that GEF Core Indicators (as well as Tracking Tools) are specified multiple times in the Guidance on Conducting Terminal Evaluations for UNDP-supported GEF-financed Projects and also in the TE TOR template for UNDP-supported GEF -financed projects. A dedicated RBM team within UNDP's GEF Coordination Unit QAs draft Core Indicator files.	Changed.
2.11	<u>Table 2. 'Quality control by GEF Unit or Agencies' program/portfolio managers'</u> Please note that region-based Technical Advisors provide quality assurance support to UNDP Country Office staff. For self-evaluation systems, a dedicated RBM team that is part of the GEF coordination unit in UNDP, issues guidance and training on PIRs, MTRs and TEs; QAs draft Core Indicators files; and reviews draft MTR/TE reports upon request to ensure quality and compliance with GEF and UNDP requirements.	Changed. Change made in the relevant table.
2.12	<u>Table 2. 'Internal peer review process'</u>	Changed.

	This table indicates that UNDP TEs and PIRs include an internal peer review process but MTRs do not. Please note that the UNDP Guidance for MTRs of GEF-financed projects places emphasis on an independent, participatory and collaborative approach to MTRs (similar to that of TEs). The MTR process involves: project beneficiaries, stakeholders, UNDP regional staff, UNDP country office staff, GEF OFP, project team, Implementing Partners, partner organizations, etc. All draft MTR reports are circulated for comments by stakeholders.	Change made in the relevant table.
2.13	<u>Table 2. 'Updating ToC at regular intervals'</u> Please note that the project document template for UNDP-supported GEF-financed projects includes “periodic appraisal of the Project’s Theory of Change and Results Framework with reference to actual and potential project progress and results;” as part of the project M&E Officer’s list of responsibilities. (Example: See page 136 of the project document for GEF ID 10073)	Changed. Change made in the relevant table.
2.14	Table 4. This table shows that UNDP’s rating scale for Relevance is not aligned with the GEF rating scale. Please refer to the following UNDP-issued documents which show that Relevance is rated on a 6-point scale: - Table 9 on page 36 in the Guidance on Conducting Terminal Evaluations for UNDP-supported GEF-financed Projects - TOR Annex F in the TE TOR template for UNDP-supported GEF - financed projects	Changed. Change made in the relevant table.
3	Shaanti Kapila, Senior Operations Officer, World Bank	
3.1	Para 49 (a) on Relevance includes the statement, “Further, ADB, WBG and EBRD do not address the project’s relevance to the GEF.” We disagree with this statement and request that it be revised. Most Bank ICRs comment on the project’s relevance to the GEF, making explicit reference to a project’s relevance to the country’s NDC or NBSAP and/or other relevant sectoral policies and strategies related to the environment and climate change.	Changed. Correction has been made.

	I just quickly reviewed 5 randomly selected ICRs from recently closed GEF-5 and 6 projects and they all comment on the project's relevance to national CC and environment policies, strategies and action plans, and some make explicit reference to GEF FA strategies.	
3.2	Para 60 includes the statement, <i>"Except for World Bank, the terminal evaluations for full size projects and medium size projects are of similar quality. In the World Bank the quality of terminal evaluations for medium-size projects is relatively lower because the quality assurance arrangements for medium size projects are not robust."</i> We disagree with the statement that QA arrangements for TEs/ICRs of World Bank MSPs are not robust and request that it be revised. It is true that the World Bank applies a simplified process for the ICRs of small recipient-executed trust funds, which include GEF MSPs. These projects can use a simplified ICR template and there are three main differences between processing simplified ICRs vs. regular ICRs: simplified ICR does not include TOC or efficiency sections, ICR doesn't go to Board, and ICR doesn't go to IEG for evaluation. However, approval processes are similar for both (PM clears and CD/RVP concurs). This approach is consistent with the Bank's overall approach to apply simplified procedures for smaller projects. To say that they are not robust in an absolute sense implies that they are worse than the QA arrangements of other agencies (which is not the case for most agencies) and that there should be the same arrangements for projects regardless of size (which for the Bank is not reasonable). Moreover, it is not clear how IEO is determining the quality of the Bank's MSP ICRs (Table 3 p. 18) as ratings for these ICRs have not been provided by IEG and the GEF IEO TE validation criteria is not explained in the text.	Partial revision. Some change in the text to moderate the inference. The reviews of all the terminal evaluations prepared by the World Bank for MSPs was conducted by the GEF IEO. GEF IEO also conducted reviews of a sample of terminal evaluations for World Bank's full-size projects. This provides a basis for comparison of quality.
4.	Stephen Hutton, Senior Evaluation Officer, Independent Evaluation Group	
4.1	The evaluation seems to come with a presumption that more evaluation is always better, and the existing GEF rules are not questioned. For example, the World Bank is listed as out of compliance for not conducting an independent validation on every GEF-financed activity, even medium-sized GEF projects that may be only \$1 million. Yet, it is hard to imagine that it would be good value for	No change Our assessment is based on the information provided in the terminal evaluation reports submitted by the World Bank. There are

	money to incur the costs of full self-evaluation and independent validation for all such activities.	information gaps in the terminal evaluations both for full size and medium size projects. For full size projects terminal evaluations submitted by the World Bank are at par with those submitted by other Agencies. However, for medium size projects we have found more gaps for the World Bank than for other Agencies. We have stated the factual position.
4.2	That evaluation systems are relatively weak on the question of whether the right things were done. This certainly resonates; relevance has long been one of the weakest parts of many project evaluation systems, in particular when it tends to default to alignment with strategies. However, this point could be interrogated further. Deciding whether the right thing was done, in a particular place and context, is a challenging thing to do. It relies implicitly on some form of counterfactual, which is difficult to generate in a project evaluation system that relies on monitoring data and comparisons to targets. It would be useful to unpack whether it is realistic to use a mandatory systematic project evaluation system with accountability purposes to generate adequately rigorous	Change Agree with the feedback. Some context has been added to highlight experiences of the Agencies in using other evaluations – other than the mandatory evaluations – to generate systematic information on impacts of interventions.
4.3	Evaluation systems do not support cross-agency learning. This is undoubtedly true. But again, it would be useful to explore this in the context of evaluation systems that often struggle to support effective intra-agency learning. It is probably even more difficult to promote learning across agencies than it is internally. There could be a public good value in the GEF-EO effectively subsidizing cross-agency learning and creating a platform, but it would be useful to have more thinking on what this could look like and how it might work, and how it would address existing learning barriers faced already by operational staff and decision-makers.	Thanks for the suggestion – totally agree. This idea is reflected in one of the recommendations of the evaluation.
4.4	Finally, the report discusses the changes made so far to IEG's Project Performance Assessment Reports, including the relegation of ratings to an appendix. It might be useful for GEF-EO to know that this product line is undergoing further reform, and may end up dropping ratings completely and prioritizing learning goals, though this has not yet been finalized.	No change. Thanks for sharing the information. We have focused on the present practice. We take note of action plans, but do not discuss

		different ideas that are being considered by Agencies.
5.	Michael Spilsbury, Director, Evaluation Office, UNEP	
5.1	I still remain unconvinced by the Title of the SES evaluation. The report highlights the variety of arrangements for evaluation across GEF agencies and they are seldom, as this report shows, exclusively self-evaluation. Even within the WB there is a pivotal role played by independent evaluation in the internal GEF evaluation system. It's a pity as I think it sends an incorrect message to Council.	No change. There are different terms being used across the GEF Partnership – each of these terms has its advantages and drawbacks. We have defined what the term means within the context of this evaluation. The evaluation does look at the relationship between self-evaluation and evaluation unit. We aim to describe the situation as accurately and fairly as possible.
5.2	Para 14..... The self-evaluation system of a GEF Agency – as it relates to the GEF supported activities – is the primary unit of analysis.” I think the unit of analysis throughout the report was actually broader – it was the agency evaluation arrangements for GEF projects - not just the self-evaluation component (where it exists) ref WB example above and description paras 28-30. The arrangements for the agencies mentioned in Para 30 don't fit the self-evaluation definitions given, it is true these agencies don't have full structural independence – but that does not make their work self-evaluation - as their Peer Reviews indicate.	Indeed, some of the aspects covered by the evaluation are broader than an agency's self-evaluation system. For some others, such as desk review of terminal evaluations to assess compliance with the GEF IEO guidelines, project is the unit of analysis. However, broadly, the focus of the evaluation is self-evaluation system of an Agency. As explained, this does not preclude our using other units for analysis where this may be useful.
5.3	“In the evaluations conducted by the evaluation units of the GEF Agencies, there is a recognition that the self-evaluation systems are not being used for learning to the extent that they should.” I suggest the following to improve the clarity of the first sentence. “In evaluations of internal self-evaluation systems conducted by the evaluation units of the GEF Agencies, there is a recognition that such systems are not being used for learning to the extent that they should.”	No change. The report will be edited for such errors.

5.4	Table 2 on Page 17 shows UNEP as only partially compliant with the requirements for Social and Environmental Safeguards in TEs. I attach our GEF ToR template that clearly indicates how this aspect is included in the evaluations we undertake. Cross-checking the TE Guidance document and the Safeguards policy, I am not sure why our approach is deemed only partially compliant. If it is based on report content then recently completed TEs are often of projects designed some time ago when less attention was afforded to the issue. As with the incorporation of gender issues over time, we can expect these aspects (gender / safeguards) to feature more strongly in evaluations as the cohorts of GEF projects that had more design and implementation attention on these issues reach completion. That the findings in reports may be lean, should not affect compliance if the ToRs clearly set out that the issue is to be examined. It is challenging to evaluate the absence of something. Similarly, Table 2 also lists UNEP as partially compliant with the requirements for stakeholder consultation. A quick perusal of our standard ToRs (attached) shows that comprehensive attention is paid to this aspect, and indeed consideration of this factor is integrated in the analysis throughout UNEP TE reports. (ctrl-F “stakeholders” in our standard TOR template).	Change. The review on compliance is based on the actual reporting in the terminal evaluation reports. The reports covered in the review were prepared from October 2017 onwards. On this we will report what we found. However, we have noted the recent changes made by UNEP in its terminal evaluation guidance documents to adequately address social and environmental safeguards elsewhere in the narrative of the report.
6.	Orissa Samaroo, Senior Director, Conservation International	
6.1	Para 29 notes that “CI has no evaluation function but is currently developing new management structures for that”. This is misleading. CI does not have an independent evaluation unit (as is pointed out in the first sentence of the paragraph) but we do have an evaluation function and we are able to conduct MTRs/TEs independently. To ensure segregation of duties from divisions within CI that may implement or execute GEF projects, the evaluation function currently sits within the General Counsel’s Office, specifically with the Senior Director for Risk and Compliance. We are currently analyzing if the evaluation function should be moved to another division.	Change The relevant text has been revised to reflect the CI situation accurately (the line in question has been removed).
6.2	Para 33 notes that fewer agencies have guidelines for conducting MTRs, it does not list CI as one of these agencies. We have developed a MTR TOR/RFP that	Change

	explicitly states what is required for an independent evaluator to conduct MTRs.	CI has been added to the list of Agencies cited as examples of Agencies with guidance for mid-term reviews.
--	--	---