

# OPS4

## PROGRESS TOWARD IMPACT



FOURTH OVERALL PERFORMANCE STUDY OF THE GEF

## COMMENTS OF THE SENIOR INDEPENDENT EVALUATION ADVISORS

METHODOLOGICAL PAPER #21

*The development of this report was conducted as part of the Fourth Overall Performance Study of the Global Environment Facility (GEF). The views expressed herein are the responsibility of the authors and do not necessarily reflect the views of the GEF Evaluation Office. The GEF Evaluation Office assumes no liability for the contents or use thereof.*



**Global Environment Fund  
Fourth Overall Performance Study**

**Comments of the Senior Independent Evaluation Advisors  
Robert Picciotto and Shekhar Singh**

***Abstract***

*The OPS4 terms of reference were extraordinarily demanding and the time and resources allocated to the review were necessarily limited. Hence it is not surprising that OPS4 could not fulfill all its intended objectives. But the final OPS4 report is highly relevant to the replenishment process. It provides a sobering account of environmental financing trends. Its treatment of GEF's focal area performance record is instructive. It breaks new ground through an innovative evaluation methodology (ROtI) focused on the likelihood of achieving global environmental benefits. Its detailed review of financial management practices and its independent review of GEF's governance arrangements include useful recommendations. OPS4 also identifies communication gaps that hinder GEF's interaction with the conventions and stresses the importance of a portfolio approach to maximize global, national and local environmental benefit. Equally, OPS4 underlines the need for a sharper focus on social and gender issues. Finally, it proposes more organizational learning through knowledge management. These findings and the associated recommendations deserve serious consideration in the context of the GEF replenishment. But given that OPS4 was not mandated to evaluate alternative delivery mechanisms, its advocacy regarding GEF replenishment levels was not backed up by adequate evidence. Equally, beyond reiterating past recommendations, OPS4 did not ascertain how GEF's management of its administrative, human resources and project oriented business processes could be transformed to further improve the efficiency and environmental impact of its operations. Nor did OPS4 face up to the limitations and leniency of outcome and sustainability ratings currently used to track GEF performance under GEF replenishment undertakings - or the need to further improve GEF evaluation practices along the lines we recommended at the interim report stage.*

**Introduction**

We were tasked to (i) provide an independent perspective on the findings, conclusions and recommendations of OPS4; (ii) verify the extent to which it has met the terms of reference (TORs) approved by the GEF council on September 5, 2008; and (iii) attest to the quality of major OPS4 products. This note assesses the final OPS4 report (GEF/R.5/18; September 25, 2009) from this perspective. It complements the comments we offered on the interim report (GEF/R.5/Inf.12, June 12, 2009).

To tackle the ambitious TORs of OPS4, the Evaluation Office (EO) relied on a remarkably diverse range of evaluation products – project-level evaluations, country-level assessments and process reviews. Extensive resort to interviews and stakeholder consultations complemented this evidence base and additional country reviews were

undertaken over and above sample field verifications of terminal evaluations and “progress towards impact” reports.

Yet, as highlighted by prior EO reports, the quality and relevance of the evaluation building blocks used for OPS4 was mixed given the limited “*evaluability*” of original project designs; the weakness of their built-in monitoring and evaluation arrangements; the scarcity of verifiable impact indicators; and the limited technical and scientific content of terminal evaluations. Thus, the observations that follow highlight the need for further improvements in the monitoring and evaluation practices used by GEF and its partners along the lines we recommended at the interim report stage.

## **Role of GEF**

The first cluster of the OPS4 terms of reference called for an examination of the *operating context* – global environmental trends, international architecture and value added of GEF support.

### *Implications of environmental trends*

OPS4 provides a sobering account of environmental trends especially with respect to climate change but it neglects to draw the implications of the changed international context created by the growing public recognition of climate change as a massive and urgent existential threat.

### *Adequacy of financial resources*

OPS4 puts forward convincing evidence in support of more funding to tackle global environmental problems (recommendation 1). But its advocacy regarding the level of the GEF replenishment (recommendation 2) is not backed up by adequate evidence since OPS4 did not (and indeed was not mandated to) evaluate alternative channels of assistance.

### *Catalytic role*

Similarly, the evidence linking the effectiveness of GEF’s catalytic role to its funding level (recommendation 4) is not fully convincing. To be sure, GEF has a long track record in catalyzing global environmental initiatives. But strictly speaking, only foundation and demonstration activities are catalytic so that the value added of retaining large investment oriented activities within the GEF tool kit is not self-evident given potential alternatives<sup>1</sup>.

### *Donor funding performance*

As acknowledged by the Evaluation Office, the methodology sketched by OPS4 to assess donor funding performance needs refinement since it does not take into account the differentiated responsibilities of individual countries implied by their respective legacies of environmental damage.

### *Interface with national priorities*

In order to facilitate the “greening” of national development policies, we fully support the need to strengthen the social and gender dimensions of GEF interventions and the need

for more adaptive project implementation practices connected to changes in the operating environment (recommendation 7).

OPS4 also recommends GEF support for the creation of GEF National Committees and GEF business plans. But unless such actions are meant to replace GEF's project approval process by a program funding approach (still untested) the proposed approach could increase transaction costs and add yet another bureaucratic hurdle for recipient countries.

Improved environmental programming at country level is a legitimate goal but capacity building for environmental programming would best address the full gamut of environmental activities at country level.

## Results

The second cluster of the TORs aimed to help ascertain the ***concrete, measurable and verifiable results*** achieved by GEF, i.e. the global environmental benefits of its interventions.

### *Need for scientific evidence*

GEF lacks the scientific and empirical evidence that would demonstrate that up-scaled action geared to the generation of global environment benefits is actually catalyzed by its interventions. This is because GEF project designs rarely include the tracking instruments needed to monitor and measure such effects at project end. Nor, beyond anecdotal evidence, is the "greening" of projects sponsored by the implementing agencies systematically and rigorously traced to GEF. Consequently, OPS4 could not demonstrate conclusively the extent to which GEF's catalytic interventions have been successful. This should be remedied, starting with GEF5.

### *Focal area assessments*

Within the above limitations, Chapters 3.2-3.6 are very informative. They identify plausible drivers of project success and provide judicious lessons of experience<sup>ii</sup>. Accordingly, they make a distinctive contribution to corporate learning and provide interesting analyses of GEF's linkages to conventions. They are refreshingly forthright and provide ample evidence that the road from project level outcomes to global environmental impacts is long, hazardous and poorly marked.

### *Bridging the ROtI- completion ratings disconnect*

In particular, the focal area assessments make ingenious use of a new and promising methodology (the Road from Outcomes to Impact or ROtI). This evaluative instrument was introduced by EO to help identify the distinctive accountabilities of GEF partners in the achievement of agreed global environmental benefits. While still experimental<sup>iii</sup> ROtI reveals that only 39% of the projects display solid progress towards impact. By contrast, GEF's annual performance reports show a share of satisfactory outcome ratings of 80% and a share of sustainability ratings of 58% for the OPS4 reporting period.

Given what completion ratings are in principle expected to measure<sup>iv</sup> we do not endorse the OPS4 assertion that implementing agencies only aim at evaluating the impact of their

operations through counterfactual methods. Indeed, we hold the view that, taken together, outcome and sustainability ratings were always meant to estimate the likelihood of achieving positive global environmental impacts – but that terminal evaluations have minimized post-implementation risks. Looking ahead, the current disconnect between ROtI estimates of “solid progress towards impact” and annual performance review ratings should be tackled through methodological harmonization and more rigorous GEF quality control using ROtI<sup>v</sup>.

#### *Tracking GEF’s impact*

It follows that we fully support the proposed integration of the new methodology in the Results Based Management system for GEF-5<sup>vi</sup>. In the meantime we recommend that the performance targets included in future GEF replenishment agreements should focus on “solid progress towards impact” measures carried out by EO.

#### *Emulating good evaluation practice*

Still looking forward to GEF5, we recommend that EO should emulate the good practices of other evaluation units including (i) providing explicit ratings for the creativity, innovation and up-scaling potential of innovative projects and distinguishing between agency and partners’ performance ratings and outcome ratings<sup>vii</sup>; and (ii) assessing results at the higher plane of country environmental strategies in consultation with its partners<sup>viii</sup>. Such evaluative practices would generate improved incentives to achieve results and in conjunction with ROtI encourage closer and more effective partnerships by attributing results to the actions (or the inaction) of individual partners.

#### *Towards country led and joint evaluations*

Finally, we reiterate our interim report recommendation that GEF should join hands with its partners to implement country-led and joint evaluations of country environmental strategies in order to improve the coherence of global, national and local environmental action and to reduce the administrative load of evaluations as prescribed by the Paris Declaration and the Accra Agenda for Action.

### **Relevance of the GEF**

Under the third cluster, the terms of reference tasked OPS4 to assess the ***relationships between GEF, the conventions and recipient governments’ policies.***

#### *Link to conventions*

Chapter 2.3 notes that the guidance is voluminous, cumulative, often ambiguous and sometimes directed to the Parties rather than to GEF. Nevertheless, it concludes that the general policy directions embedded in convention guidance are relatively clear. Yet one third of respondents to a survey consider that GEF has not been responsive to convention guidance.

These perceptions are attributed to the tensions among interest groups; inadequate dissemination of information about GEF’s mandate and resources; basic policy differences (e.g. “incrementality”, co-financing, RAF); lack of congruence between the

replenishment cycles and the convention calendars; and inadequate funding allocations for communications to the conventions of some countries<sup>ix</sup>.

The recommendations under this heading are sensible (e.g. better reporting to the conventions; direct feedback from the conventions to the Council; etc.) but the prioritization of guidance recommendations at national level is not adequately documented and it is not entirely clear how COP guidance should be treated in future project completion reports or what clarification of roles is needed between STAP and the convention secretariats.

#### *Link to recipient governments*

As to the relevance of GEF to national policies, it is treated lightly. Yet, country ownership is of critical importance to the sustainability and replication of GEF operations and there are inevitable tensions between national and global environmental priorities.

### **Performance issues affecting GEF results**

#### *Governance*

Under the fourth cluster, OPS4 was expected to ascertain whether GEF's **governance** system is "*adequate and up to international standards*". The recommendations of the senior independent consultant are sensible and suitably tailored to the evidence. One notable finding is that the lack of broad based understanding of operating policies continues to undermine the quality of GEF partnerships: many Parties consider co-financing as conditionality to access GEF funding and the Resource Allocation Framework as well as the incremental cost principle continue to be divisive and hard to implement equitably.

#### *Other performance issues*

The TORs required OPS4 to include an update regarding GEF's **resource allocation framework (RAF)**. In this connection, as noted in Section 1, OPS4 stood by the findings of the RAF mid-term review. As for the review of GEF's **efficiency and cost effectiveness** it was not carried out by focal areas, agency and modality as envisaged in the TORs and the cost comparisons with other agencies that were attempted proved partial and inconclusive. Nor were systematic efforts made to relate cost comparisons to the comparative advantage of implementing agencies or to estimate the impact of project cycle and co-funding reforms on GEF efficiency as required by the TORs.

The extent to which GEF's *composition, structure and divisions of roles and responsibilities* help to meet its mandate, operations and partnerships were not evaluated as prescribed by the TORs. The TORs had also inquired as to the extent to which GEF succeeded as a **learning organization** including state of the art science and technology. In this connection, the findings of Chapter 4.2 are sound and its recommendations unexceptionable albeit rather general (improved knowledge management; a more strategic STAP, etc.) and the chapter fails to pinpoint the responsibility for slow utilization of past corporate evaluations with respect to business processes, partnership quality and resource allocation protocols. .

## Resource mobilization and financial management

The final and fifth cluster of the TORs asked how effective has GEF been in ***mobilizing and managing its human, financial and administrative resources***. Chapter 4.3 is largely descriptive with respect to human resources and administrative aspects but it does add significant value through its detailed assessment of trust fund management, fiduciary standards and fee structures.

---

<sup>i</sup> Funds currently used by GEF for investment (about 25% for national projects - OPS4: p.56) could in principle be managed directly by the World Bank and/or regional development banks.

<sup>ii</sup> Cross cutting themes could have been identified more explicitly and more explicit attention to the interface among focal areas would have illuminated the comparative advantage of multi-focal and regional projects. We also missed an evaluative assessment of the balance struck by GEF among the focal areas, regions and countries.

<sup>iii</sup> The validity of ROI estimates can only be rigorously ascertained from field testing of project impacts many years after terminal evaluation. Such scientific investigations have yet to be funded as an integral part of GEF project designs.

<sup>iv</sup> Taken together, outcome and sustainability ratings are supposed to capture operational impacts. A satisfactory outcome means that relevant project objectives are expected to be achieved with no, minor or moderate shortcomings at the time of evaluation. A likely sustainability rating means that the project is considered likely to generate continued benefits after project implementation - with no or only moderate risks.

<sup>v</sup> Our spot checks confirmed that the performance ratings produced by the implementing agencies are lenient and not always consistent or of high quality across evaluation units. They do not capture effectively cross border impacts or interactions among focal areas and do not systematically rate the creativity, innovation or up-scaling potential of projects or the efficiency of GEF processes. More rigorous verification of ratings by EO is needed through ROI methods and increased resort to field performance audits.

<sup>vi</sup> The distinction that OPS4 draws between attribution and contribution is useful since it highlights the primary responsibility of governments and implementing agencies for achieving results. But the report does not rate the relevance, efficacy and efficiency of the GEF contribution. Since outcomes and impacts result from the actions of many actors their distinctive accountabilities for results ought to be rated separately.

<sup>vii</sup> International Fund for Agricultural Development, Office of Evaluation, *Evaluation Manual: Methodology and Processes*, Rome, 2009 (p.11).

<sup>viii</sup> In development evaluation the gap between evaluation ratings at project and country levels (the “micro-macro paradox”) has induced a shift to a higher evaluative plane: country assistance strategies have replaced projects as privileged units of account. See World Bank, Independent Evaluation Group, *Annual Review of Development Effectiveness: Shared Global Challenges, 2008*, World Bank, Washington DC, 2008 (p.17). For GEF, on the other hand, country portfolio evaluations have a restricted focus and do not evaluate the performance of implementing agencies.

<sup>ix</sup> The well known dissatisfaction of the UNFCCC with GEF could have been analyzed in greater depth since it goes well beyond a problem of inadequate communication.





[WWW.GEFEO.ORG](http://WWW.GEFEO.ORG)