



# Global Environment Facility

GEF/R.5/Inf.12  
June 12, 2009

---

Second Meeting for the Fifth Replenishment of the GEF Trust Fund  
June 25-26, 2009  
Washington, D.C.

## **FOURTH OVERALL PERFORMANCE STUDY OF THE GEF (OPS4)**

### **COMMENTS OF THE SENIOR INDEPENDENT EVALUATION ADVISORS ON THE INTERIM REPORT**

**(PREPARED BY THE GEF EVALUATION OFFICE)**

**GLOBAL ENVIRONMENT FUND  
FOURTH OVERALL PERFORMANCE STUDY  
INTERIM REPORT**

**COMMENTS OF THE SENIOR INDEPENDENT EVALUATION ADVISORS  
ROBERT PICCIOTTO AND SHEKHAR SINGH**

1. The interim report offers a valuable preview of Fourth Overall Performance Study (OPS4) results. The tentative findings are lucidly presented against the context of emerging global environmental trends. But given space constraints the interim report lacks methodological details. Accordingly, this note attests to the adequacy of evaluation evidence and methods with a view to informing final OPS4 report deliberations.

**GEF's Unique Role**

2. From a methodological perspective, the technical challenge posed by OPS4 is connected to GEF's limited mandate and resources. The organization's role is largely indirect: protecting the global environment is the collective responsibility of member countries. Thus, ever since its pilot phase, GEF has been perceived as one of the many actors accountable for the success of global environmental initiatives.

3. GEF operates mostly through voluntary partnerships. Its funding of incremental costs of environmental action provides an incentive for the generation of undersupplied global environmental goods. GEF is also tasked to help beneficiary countries meet their international commitments under the global environmental conventions to which they are signatories.

4. Specifically, by providing grant finance, technical advice and support, GEF helps to incubate innovations, motivate participants and build capacities that enable beneficiary countries to protect the global commons. In fulfilling its catalytic role, GEF is expected to be selective and judicious in its resource allocation, choice of instruments and implementing agency partners.

**Adequacy of Evidence**

5. This strictly limited and catalytic mandate has had major consequences for the structure and methods of the GEF independent evaluation function. Whereas the Evaluation Office has been asked to ascertain the impact of GEF interventions its remit does not extend to the national governments and the implementing agencies that are primarily accountable for the generation of global environmental benefits. Accordingly, GEF has chosen to rely on the evaluation units of the implementing agencies for its primary evaluation products and to complement these by its own independent assessments.

6. In particular, as a meta-evaluation, OPS4 has tapped into a remarkably wide range of evaluation products – project-level evaluations, country-level assessments and process reviews. The extensive resort to interviews and stakeholder consultations was a welcome complement to the evidence base. The sample verifications of terminal evaluations and progress towards impact reports were also useful.

7. The OPS4 team has been diligent in its interpretation and synthesis of this diverse and voluminous material. But it has had to contend with the constraints that the Evaluation Office has repeatedly highlighted: the limited scientific content of primary evaluations, the inadequate “evaluability” of project designs, the weakness of monitoring and evaluation systems and the scarcity of impact indicators.

### **Adequacy of Methods**

8. The performance ratings produced by the Implementing Agencies should be interpreted with caution not only because they are not always consistent or of high quality across evaluation units but also because their “objectives based” evaluation methods are not well adapted to GEF’s catalytic role. In particular, they do not capture effectively cross border impacts and the interactions among focal areas. Nor do they systematically rate the creativity, innovation or up-scaling potential of projects or the efficiency of GEF processes.

9. Fundamentally, using projects as privileged units of account is problematic since performance at the project level does not automatically translate into impact at the country level - let alone the global level. In development evaluation the gap between evaluation ratings at various levels (micro-macro paradox) has induced a shift to a higher evaluative plane: country assistance strategies have become the privileged units of account.

10. For GEF, on the other hand, country portfolio evaluations have a restricted focus and do not evaluate the performance of implementing agencies. The non-existence of GEF country strategies or programs underlies this state of affairs. Special studies are needed to probe country and regional contexts, horizontal (e.g. inter-sector) influences and country capacity constraints.

### **Contribution and Attribution**

11. The interim report confirms that GEF cannot be held accountable for actions taken by other agencies or by countries within which GEF operates. Instead, it should be assessed for its distinctive contribution to the overall process. But how specifically is this catalytic role to be evaluated under OPS4? This question needs full and explicit treatment in the final OPS4 report.

12. In principle “theory of change” assessments should track the distinctive accountability of partners including GEF – “who is supposed to do what”. But as noted above the primary evaluation products are not conceived in this fashion. This limits their

usefulness as building blocks for the assessment of GEF project selection, design and sequencing of operations at country and global levels.

13. The ROTI approach was a highly innovative approach meant to fill this evaluation gap but the scope of ROTI studies is not comprehensive and the validity of ROTI estimates only comes from field testing of project impacts many years after terminal evaluation and the linkages with project level sustainability ratings are hard to track.

14. The Evaluation Office is well aware of these limitations and we trust that it will exercise the utmost care in drawing its final evaluative judgments. The key performance tests should deal with the effectiveness of GEF's operational practices, its resource allocation decisions and its partnership processes. OPS4 is tasked to ascertain whether these processes were well designed and efficiently implemented in order to pilot innovative approaches and to motivate and enable partners to take on the up-scaling challenge.

## **Conclusions**

15. Given the above considerations, the final OPS4 report should take full account of the limits imposed by available evidence in reaching its evaluation judgments. It should describe in full the triangulation methods the Evaluation Office has used to secure consistent judgments about GEF performance. It should also lay out steps that will improve the quality of future GEF independent evaluations.

16. In particular, OPS4 should probe the extent to which country evaluations and joint evaluations should become central to the evaluation of GEF effectiveness, whether ROTI concepts should be applied in primary evaluations; how best to verify that incremental funding is truly incremental; how to measure whether GEF is succeeding in integrating environmental concerns into governments and Implementing Agencies policies and practices; etc.

17. Thus, we recommend that the OPS4 report should:

- Articulate the limits of current project level ratings and their consistency across implementing agencies.
- Interrogate the limits of project ratings as currently designed and implemented.
- Draw the implications of the “country program evaluation gap” for the validity of OPS4 findings.
- Explain the scope, potential and limits of the innovative ROTI methodology.
- Define the links between sustainability ratings and ROTI assessments and their relationships to impact measures.
- Consider the need to move towards joint evaluations at country and focal levels designed to improve the relevance, coherence of current evaluation practices and reduce their heavy administrative load.

June 11, 2009