

# Independent Evaluation: Principles, Guidelines And Good Practice

The World Bank  
Development Grant Facility (DGF) Technical Note

November, 2003

## Introduction

Global and regional partnership programs (GPPs) are comprised of multiple partners that operate across borders to tackle issues that cannot be addressed at the country level alone. They are increasingly important in achieving the Millennium Development Goals (MDGs) and in complementing and reinforcing the World Bank's country activities.

As GPPs become mainstreamed into the Bank's business,<sup>1</sup> and as the Bank puts new emphasis on learning and on demonstrating accountability for results, independent evaluation is taking on new importance. Independent evaluation is also used to inform the Bank's funding decisions for GPPs through the DGF.

---

***Evaluation provides critical evidence of the outcomes of global programs and their roles in enhancing the development impacts of partner and country efforts on the ground.***

---

This note is intended for GPP managers and development partners. It summarizes the principles of good evaluation (Box 1), reviews guidelines (Box 2),<sup>2</sup> and discusses good practices in the conduct and design of independent evaluations.

---

<sup>1</sup> See *The World Bank's Approach to Global Programs: An Independent Evaluation—Phase 1*, OED, World Bank, May 8, 2002.

<sup>2</sup> The guidelines in this note (and in Box 2) are for DGF-supported programs but are broadly applicable to GPPs. The DGF approach to monitoring and evaluation is described in *The Development Grant Facility: FY99 DGF Annual Review and Proposed FY00 Budget*, July 26, 1999. In addition to independent evaluation, it requires programs to undertake the following:

**Progress/Grant Completion Reporting:** The annual DGF application includes a progress report on activities and accomplishments. Bank task team leaders are also required

## Box 1: Principles of Independent Evaluation

**Usefulness:** For evaluation to affect decision making, findings must be perceived as useful and as geared to current operational concerns.

**Independence:** For evaluation to be impartial, it must be free from bias in findings, analysis and conclusions. In turn, this means independence from line management.

**Credibility:** Evaluation must be perceived as objective, rigorous and impartial. Credibility rests on the professional quality of evaluators and the rigor of methods.

**Transparency:** Credibility and usefulness will also depend on the transparency of the evaluation – the ready availability of findings to all stakeholders.

Drawn from *Lessons & Practices*, Number 10, OED, World Bank, July 1997.

The DGF is the World Bank's grant mechanism for global and regional programs. It provides \$175 million annually to some 50 GPPs.

This note was prepared by Randall Purcell of the DGF Secretariat. The author is grateful to OED for drawing on its work on global programs.  
[rpurcell@worldbank.org](mailto:rpurcell@worldbank.org), tel: 202-473-5571

to report on outputs and impacts of grant programs within six months of the closing date of the grant.

**Financial Reporting/Audit:** The DGF requires from an external grant recipient the provision of an audited financial statement together with a statement detailing the specific use of DGF funds. DGF-supported programs may be audited by the Bank's Internal Audit Department. The results are reported to the Bank's Senior Management and the Audit Committee of the Board.

## Box 2: Evaluation Guidelines<sup>3</sup>

Evaluations should:

- ❑ **review measurable indicators of performance** – including outputs, outcomes and impacts – and assist programs to develop improved indicators (see Box 4).
- ❑ **be conducted every three to five years** (for programs that receive total DGF funding of \$300,000 or more).
- ❑ **be owned and utilized by all partners** and not be driven solely by the needs of the Bank. Therefore, timing and coverage should take into account partner preferences.
- ❑ **be overseen by the program’s governing body** and conducted with the assistance of program management. Evaluators should report to the governing body, not to management.
- ❑ **be conducted by individuals or companies that have been at “arm’s-length”** from the initiative.
- ❑ **report on outputs, outcomes and impacts.**
- ❑ **address the performance of the governing body** as well as of management.
- ❑ **make recommendations for financial sustainability** as well as program delivery.
- ❑ **be discussed by the relevant Bank Sector Board** and used by them as the basis for considering DGF funding requests.
- ❑ **have terms of reference and selection of consultants cleared with the DGF Secretariat.** The DGF grant can cover a part of evaluation costs which should also be borne by other partners.

## The Purpose of Independent Evaluation

An independent evaluation should serve both as a learning tool and as a means to demonstrate accountability.

**Learning.** An independent evaluation is an effective means for learning how GPPs can be strengthened and how resources can be used more effectively. It provides an understanding of the following:

- 1) how well a program has articulated its vision and is achieving its mission;
- 2) the quality of activities and how useful they are in meeting clients’ needs;
- 3) how capacities such as financial and executive management, governance and country-based resources affect quality.

Evaluation that is driven by the need to make better program decisions should clarify with some specificity “who needs to learn what.” This frames the exercise and gives it greater legitimacy and focus. Thus, a primary audience for the evaluation should be agents for change within programs, including task team leaders and managers. Funders will also need assurance that they are supporting program learning.

---

***It is important that the evaluation go beyond asking, ‘ is this being done well’ to ask, ‘how can this be done better’.***

---

**Accountability.** Because development resources are limited and the time frame to achieve the MDGs is short, donors are increasingly focusing on results. GPPs are under particular scrutiny because grant resources are especially scarce and because these programs are largely independently governed.

Independent evaluation is crucial for demonstrating how resources have been spent and what results have been achieved. At the same time, it is important that the evaluation go beyond results and their attribution to ask, What would have happened if the program did not exist? What would happen if the program ceased to exist?

---

<sup>3</sup> Requirements for DGF-supported programs. Examples of good practice for evaluation terms of reference and evaluation reports can be found on the DGF website.

## Conduct and Form of the Evaluation

Independent evaluations should be carried out in ways that support the principles of good evaluation shown in Box 1. Important questions that need to be addressed are the following:

**Who governs the evaluation process?** The governing body, in collaboration with management, should formulate key questions in consultation with partners. Final decisions on terms of reference and selection of consultants should be made by the governing body; the evaluators should report directly to that body.

**Who pays for the evaluation?** Costs should be borne equitably by the various donors. No one donor should pay for or drive the process.

**How much should the evaluation cost?** Costs will vary depending mostly on the size of the program and the need for field work. Evaluations of newer programs that do not require extensive field investigation may typically cost between \$75,000 and \$150,000. Older programs may have more development impacts to demonstrate and would be more costly to evaluate.

To contain costs (and to focus attention), evaluations should generally be limited to three to six months depending on the size and complexity of the program.

### How should evaluators be selected?

Evaluators should be selected on the basis of their expertise in the field, the diverse points of view they bring to a team and their objectivity. Evaluators should have knowledge of evaluation methodology and be familiar with development issues and international organizations.

To avoid conflict of interest, evaluators should have an “arm’s-length” relationship with the program: They should have had no substantive prior involvement in its establishment or operation. Competitive bidding for consultants is encouraged but not required.

### How should the evaluation be carried out?

The evaluation will typically involve a desk review of documents, interviews with stakeholders and, where applicable, collection of data and evidence to determine development impacts. Because of the importance of assessing results at the country level (see below), the involvement of country counterparts and country clients is often crucial (see Box 3).

### Box 3: Approaches for Integrating Country Partners in Evaluations

- Invite recipient-country representatives to participate on the evaluation team.
- Form an advisory panel comprised of recipient-country nationals to advise the evaluation team and/or review its findings.
- Involve recipient-country nationals as key informants, focus-group members, researchers, interviewers, country/regional experts, etc.

Instead of a full-fledged evaluation, programs in the early stages of implementation might first undertake a *process review* in order to take stock of operating procedures and the adequacy of controls. Such an exercise is often limited to a study of documents and to interviews. As implementation progresses, subsequent evaluations would focus more on outcomes and impacts through assessments in the field.

### What should be the scope of the evaluation?

An independent evaluation should normally be done every three to five years. It should examine past performance focusing on the period after the last evaluation. Based on past performance, it should develop lessons that can improve performance going forward.

---

***The evaluation should not be conducted simply to determine the extent to which a program worked but should also develop lessons that can improve the way it and future programs will function.***

---

**What should be evaluated?** The evaluation should examine how activities have been carried out and how they could be done differently. The focus should be on the following:

- program performance, including an assessment of its mission, objectives, governance, policies, management, operational modalities and impacts;
- the performance of other partners, and coordination needed among them;<sup>4</sup>

---

<sup>4</sup> For DGF-supported programs, TOR and consultants should be cleared with the DGF Secretariat.

---

<sup>4</sup> The success of most GPPs will ultimately be measured by progress towards achievement of the MDGs. It is vital, therefore, that the evaluation examines the coordination needed among partners to achieve results at the country level.

- the perspectives of other stakeholders, including beneficiaries;
- the key exogenous or contextual factors that affect implementation of the program.

Programs that principally involve networking or advocacy should pay particular attention to how successfully the program has facilitated the participation and commitment of partners.

Evaluations should examine progress against performance indicators established by the program (see Box 4). The World Bank's Operations Evaluation Department has developed a set of core evaluation questions designed to assess performance against such indicators (Box 5). Evaluations should focus on responding to these questions (or some version of them) and should provide evidence to substantiate findings and conclusions.

**How are evaluation findings reported and made accessible?** Evaluation reports typically include the following sections:

- summary and conclusions;
- description of program objectives, components and activities evaluated;
- methodology used to develop findings;<sup>5</sup>
- implementation experience;
- responses to core questions;
- summary of key findings, including strategy for financial sustainability or disengagement;
- conclusions and lessons learned;
- recommendations.

A completed evaluation should be sent first to the program's governing body. Key partners should have an opportunity to provide feedback before the evaluation is released to the public. With the approval of the governing body, the evaluation should be disseminated widely, including on the Internet.

World Bank task team leaders should ensure that evaluations are discussed at Sector Board meetings and that a record of that discussion is made available to the DGF.

---

<sup>5</sup> The evaluation should include information on who commissioned it, what process was followed, how evaluators were selected and to whom they reported, who was consulted and what documents were reviewed.

#### Box 4: Developing and Using Performance Indicators<sup>6</sup>

##### **What are performance indicators?**

Performance indicators are measures of program *impacts, outcomes, outputs, and inputs*. They are monitored throughout implementation to assess progress towards program objectives.

##### **How should indicators be developed?**

Performance indicators should be based on an underlying *logical framework* that links program objectives with program components and their respective inputs, activities and outputs at different stages of implementation.

##### **What do performance indicators look like?**

Examples of performance indicators are given below for a hypothetical research program.<sup>7</sup>

- Inputs: funding received, number of staff hired, equipment purchased
- Outputs: studies conducted, training courses organized
- Outcomes: government policies improved, donor investments committed, technologies developed
- Impacts: improved access to goods and services (e.g. vaccines, water, energy), health and nutrition status improved

It is useful to think of the results chain of indicators as a continuum mirroring the logical means-end relationship of the program: inputs yield outputs that contribute to outcomes and impacts.

**How should indicators be utilized?** The development of performance indicators should be an integral part of a program's strategic plan. Indicators should be utilized periodically in progress reports<sup>8</sup> and should be used to measure success in an independent evaluation.

---

<sup>6</sup> Drawn from *Performance Monitoring Indicators: A Handbook for Task Managers*, OED, World Bank, 1996.

<sup>7</sup> These are "direct" indicators. "Indirect" measures of performance might include indicators for risk, efficiency, effectiveness and sustainability (see Box 5)

<sup>8</sup> The annual progress report to DGF should describe progress against previously developed indicators.

## Box 5: Key Evaluation Issues and Questions<sup>9</sup>

Global Relevance
<p>1. <b>International consensus:</b> To what extent is there an international consensus, either formal or informal, concerning global challenges and concerns in the sector and that global collective action is required to address these challenges and concerns?</p>
<p>2. <b>Relevance:</b> To what extent is the program</p> <ul style="list-style-type: none"> <li>• addressing global challenges and concerns in the sector,</li> <li>• consistent with client countries' current development priorities,</li> <li>• consistent with the missions and strategies of partners?</li> </ul>
<p>3. <b>Eligibility criteria:</b> To what extent is the program</p> <ul style="list-style-type: none"> <li>• providing global and regional public goods,</li> <li>• supporting international advocacy to improve policies at the national level,</li> <li>• producing and delivering cross-country, relevant lessons to client countries,</li> <li>• mobilizing substantial incremental resources?</li> </ul>
<p>4. <b>Compatibility:</b> To what extent do the activities of the program complement, duplicate or compete with other development instruments?</p>
Outcomes, Impacts and their Sustainability
<p>5. <b>Efficacy:</b> To what extent has the program achieved or is it expected to achieve its stated objectives?</p>
<p>6. <b>Value added:</b> To what extent is the program adding value to</p> <ul style="list-style-type: none"> <li>• what partners are doing in the sector,</li> <li>• what developing countries are doing in the sector in accordance with their own priorities?</li> </ul>
<p>7. <b>Sustainability of outcomes and impacts:</b> How resilient are the outcomes and impacts of the program?</p>
<p>8. <b>Monitoring and evaluation:</b> To what extent does the program have</p> <ul style="list-style-type: none"> <li>• clear program and component objectives with verifiable indicators,</li> <li>• a structured set of quantitative or qualitative indicators,</li> <li>• systematic and regular processes for data collection and management,</li> <li>• independence of program-level evaluations,</li> <li>• effective feed feedback from monitoring and evaluation?</li> </ul>

(continued on p. 6)

<sup>9</sup> Adapted from OED, *Global Public Policies and Programs Study Background Paper*, Annex A. October 8, 2003.

Box 5 (continued)

Organization, Management and Financing	
<b>Governance and management</b>	<p><b>9. <i>Efficiency</i>:</b> To what extent has the program achieved or is it expected to achieve</p> <ul style="list-style-type: none"> <li>• efficient allocation of resources,</li> <li>• benefits that are more cost effective than those that could be achieved by providing the same service on a country-by-country basis,</li> <li>• Benefits that are more cost effective than those that could be achieved if individual contributors to the program acted alone?</li> </ul>
	<p><b>10. <i>Legitimacy</i>:</b> To what extent is the authority exercised by the global program effectively derived from those with a legitimate interest in the program (including donors, developing and transition countries, clients and other stakeholders) taking into account their relative importance?</p>
	<p><b>11. <i>Governance and management</i>:</b> To what extent is the governance and management of the program</p> <ul style="list-style-type: none"> <li>• transparent in providing information about the program,</li> <li>• clear with respect to roles and responsibilities,</li> <li>• fair to clients,</li> <li>• accountable to donors, clients, scientists/professionals and other stakeholders?</li> </ul>
<b>Partnerships and participation</b>	<p><b>12. <i>Partnerships and participation</i>:</b> To what extent do developing and transition country partners, clients and beneficiaries participate and exercise effective voice in the various aspects of program design, governance, implementation, monitoring and evaluation?</p>
<b>Financing</b>	<p><b>13. <i>Financing</i>:</b> To what extent is funding positively or negatively affecting</p> <ul style="list-style-type: none"> <li>• the strategic focus of the program,</li> <li>• the governance and management of the program,</li> <li>• the sustainability of the program?</li> </ul>
	<p><b>14. <i>Bank's presence</i>:</b> To what extent has the Bank's presence as a partner catalyzed non-Bank resources for the program?</p>
<b>Risks and risk management</b>	<p><b>15. <i>Risks and risk management</i>:</b> To what extent have the risks associated with the program been identified and effectively managed?</p>
Partner Performance (examples given are for the World Bank)	
<p><b>16. <i>Comparative advantage</i>:</b> To what extent are partners maximizing their comparative advantages in support of the program at the <i>global level</i> (global mandate and reach, convening power, mobilizing resources), and at the <i>country level</i> (multi-sector capacity, analytical expertise, country-level knowledge)?</p>	
<p><b>17. <i>Linkages to country operations</i>:</b> To what extent are there effective and complementary linkages, where needed, between global program activities and partner country activities to the mutual benefit of each?</p>	