



Frontier Issues

Knowledge about building government systems for monitoring and evaluation is a work in progress. Considerable experience has been accumulated and analyzed in OECD countries (chapter 3), and the literature on the experience of developing countries has started to grow over the last 10 years in particular.

But there are many issues that are either not well understood or not well documented (box 14.1). Nine frontier issues are suggested here. Most if not all of these issues deserve additional, in-depth investigation through a long-term research program.

Cost-Effective Approaches

As flagged in chapter 13, one important issue is **which approaches to strengthening government M&E systems are most effective, and in which types of country.** A long list of different types of actions has been presented in this volume, but the relative cost-effectiveness of each is unknown. One yardstick is provided by the evaluation criteria that can be applied to any type of capacity-building effort, such as its efficiency, effectiveness, sustainability, and so forth.¹

Donors are becoming increasingly active in helping governments strengthen their M&E systems, and we can expect a number of donor evaluations of those capacity-building projects in coming years. Until now, most donor support that has been provided through lending projects has been as part of a larger reform effort, and thus when

the projects are evaluated, relatively little attention is devoted to issues such as the cost-effectiveness of individual actions—such as those listed in table 13.1.

The IEG, which has had an M&E capacity-building program for more than 20 years, conducted a self-evaluation of its work (IEG 2004a); this evaluation is discussed in annex D. More evaluations of donor and government efforts to build M&E systems are needed to deepen our understanding of what works best in different situations. The World Bank is currently preparing its first project devoted solely to strengthening a government M&E system, for Colombia. The future monitoring and evaluation of this loan should provide invaluable lessons.

The growing literature on government M&E systems focuses on good practice or promising practice countries (chapter 5). This is understandable for at least two reasons. First, donors working in this topic and governments interested in strengthening their M&E systems want to replicate success; thus, they want to understand the key capacity-

Box 14.1: Strengthening Government M&E Systems—Some Frontier Issues

- Cost-effectiveness of alternative approaches to building or strengthening government M&E systems
- Examples of countries that have tried but failed to strengthen their M&E systems
- Cost-effectiveness of different types of M&E systems
- How much evaluation is enough?
- Sector systems for M&E
- Subnational M&E systems
- Donor harmonization and country alignment
- Civil society involvement in monitoring and evaluating government performance
- Formal standards for good practice M&E systems.

building actions and underlying success factors. Second, donors that advocate government M&E systems as a desirable means for improving sound governance need to be able to identify credible examples of well-functioning M&E systems.

What is missing from this approach are **examples of countries that have tried but failed to strengthen their M&E systems**. We know as evaluators that we often learn as much from failure as from success. But there are few documented examples.² Once more donor evaluations of capacity-building work in this area have been conducted, it will be important to focus specifically on these “failures.”

Another question of interest is **which types of M&E system**—each with its own specific purposes, such as for performance budgeting or for better program management—**are most cost-effective**. A yardstick here is provided by the extent of utilization of M&E information in different ways and the value of that utilization—a consistent argument in this volume is that utilization is the bottom-line measure of success of an M&E system. An example here is Chile; its M&E system costs the government about \$0.75 million per year, and the M&E information it produces is used intensively by the finance ministry for its budget decision making and for imposing man-

agement improvements on sector ministries and agencies (chapter 6).

The value of M&E information is an interesting question, but it may be not be possible to provide a clear answer; much depends on the circumstances and level of demand (or commitment) of each individual government. It is likely that some types of M&E systems—for example, those relying on an agenda of rigorous impact evaluations—might be too demanding and costly for the poorest governments. In such a situation, it might, however, be possible to rely on donors to fund and manage such evaluations.

Related to this issue is the question of **how much evaluation is enough**. How much should governments be prepared to spend annually on their M&E system, once it has been created? The evaluation literature occasionally makes passing reference to this issue, and there has been some suggestion that it would be appropriate to allocate around one percent of an organization’s total spending on evaluation. But this number has no logical or empirical basis. The issue remains important, however. Chile’s annual government budget is about \$20 billion, and the finance ministry spends only \$0.75 million annually on the M&E system. Is this enough, or is it too much?

Chile’s finance ministry funds fewer than 20 evaluations per annum. Compare this to Australia, where, in the mid-1990s, some 160 major evaluations were under way at any one time (chapter 8). As in Chile, Australia’s evaluations were used intensively in the budget process. Which country produced the optimum number of evaluations? This is not an easy question to answer. The parsimonious approach of Chile’s finance ministry would seem to imply that it would never finance a series of expensive impact evaluations on the scale of Mexico’s *Progres*a evaluations (discussed in chapter 6). Yet the Chilean government utilizes the M&E information intensively and considers its system highly cost-effective; the returns it has derived from its investment in M&E are highly positive. Thus, it should spend more—and arguably much more—on its M&E system.

If we accept that utilization of M&E information is the basic yardstick of success of a system, this suggests a simple decision rule as to how much to spend on M&E: Continue to spend on the M&E system if the benefits from using the M&E information are judged to be high; and if the potential benefits from a proposed increase in spending on M&E are also judged to be high, then spend that additional amount—using either the government’s budget funds or donor loans. This puts the onus on those who manage the M&E system to demonstrate credibly that M&E information is indeed being heavily used and that the system is cost-effective.

This again underlines the value of undertaking regular M&E of the system itself. In those situations where M&E information is not being used intensively, clearly an analysis of the reasons needs to be conducted. If there are perceived problems with the reliability of the M&E information, then more money might have to be spent to improve it. And if there are problems of weak government demand, even for reliable information, then steps would need to be taken to strengthen demand. If those steps do not work and it is not possible to increase demand, then one can argue that the M&E system might need to be scaled back, or perhaps even abolished.

Most of the discussion in this volume is focused on national, whole-of-government M&E systems. This is typically the domain of finance or planning ministries or the president’s office. On the donor side, the staff who work on such issues are often public sector management or public expenditure management experts whose work is largely concerned with systemic issues of sound governance; these staff also tend to have a focus on the central ministries of a government.

Sector Systems

Another frontier issue is **sector systems for M&E**, such as often occur in health and education ministries. Donor staff who work at this level are often sectoral specialists. National and sectoral systems are clearly related; indeed, a national system usually has to rely on sectoral systems for

much of the monitoring information it requires. That information may come from administrative records, or possibly from special surveys and censuses conducted either by sector entities or by the national statistical office.

Sector systems for M&E are clearly important in their own right and also because of the role they play in a national system. Yet the literature on government M&E systems, and the literature on government statistical systems,³ appears to concentrate largely on national systems. This gap needs to be filled. There are some sector ministries and agencies in developing countries that appear to have well-performing M&E systems, especially in the health, education, and social welfare sectors. But we don’t know enough about these possible islands of good practice, the factors that led to their existence, or their sustainability.

There are at least some documented examples of high-quality work on sectoral M&E systems, and these in themselves merit greater attention. One example is the Health Metrics Network (HMN), a global partnership supported by the World Health Organization, which promotes the development of health sector information systems.⁴ The Network has prepared a detailed diagnostic tool for assessing the performance of health information systems, identifying critical weaknesses, and monitoring the performance of efforts to address those weaknesses. This tool can be used to assign ratings to a large number of dimensions of a health information system, and these ratings are assigned to five categories of performance, ranging from “highly adequate” to “not functioning.”⁵

This volume takes the approach that the lessons from institutionalizing national M&E systems—those discussed in chapter 10—also apply to sectoral M&E systems. There is some evidence to support this belief, such as the analytical work conducted in relation to Mexico’s social development agency, SEDESOL (World Bank 2004c; Hernandez 2006). But more research is warranted into this proposition that sectoral M&E issues are essentially a microcosm of the issues facing a national M&E system. At the very least, there are likely

to be some important differences in emphasis. Collectively, we need to understand what these are.

Subnational M&E Systems

A comparable issue to sector M&E systems is **subnational M&E systems**, especially where there is not a unitary system of government. Many governments have a federal system with several layers, with varying degrees of devolution and decentralization of functions.

The relationships between these layers of government can be complex, including accountability relationships and formal requirements for provision of information. Few case studies have been prepared on M&E at the subnational level. Yet it is clearly important because of the complexity of the relationships.

Another reason this is relevant is the reality that typically most monitoring information is collected at the facility level—thus, the nature of the relationship between individual facilities and sector ministries/agencies on the one hand, and subnational levels on the other, may help determine the quality of the information provided, as well as the extent of use of that information by the facilities themselves. As noted in the case of Uganda (chapter 9), the information workload at the facility and district levels can be onerous, with highly negative implications for the quality of data.

Donor Harmonization

A perennial issue in the donor evaluation community is **donor harmonization**, or the lack of it. This has multiplied the workload imposed on donor countries because of their need to comply with the different donor requirements for M&E. In the 2005 Paris Declaration, donors pledged to endeavor to harmonize these requirements and to align themselves as much as possible with country systems and approaches to “results frameworks and monitoring”—that is, **country alignment** (High Level Forum on Harmonisation, Alignment, Results 2005).

Donors and countries have also committed to working jointly to strengthen country capacities

and demand for results-based management. It will be interesting to learn to what extent, if any, these objectives have been met; a series of evaluations of the results of the Paris Declaration are under way, and they will be reported over the 2008–10 period.⁶

One bright spot is the evidence of a growing number of donor/government impact evaluations. The World Bank, for example, is currently planning or conducting around 30 rigorous impact evaluations jointly with governments in Latin America, typically as part of Bank projects with governments; it appears that these are being conducted in a highly collaborative manner. It is likely that there are some lessons here for the conduct of other types of evaluation and for joint donor/government evaluations in other Regions.

Examples of **civil society involvement in monitoring and evaluating government performance** are presented in chapter 3—the Citizen Report Card Initiative, which was first conducted in Bangalore, and the comparable *Bogotá Cómo Vamos* approach. These initiatives have successfully pressured municipal governments to improve their performance. The CRC has been replicated in a number of other cities in India and other countries.

The challenge for the donor community is how to encourage the replication of this kind of civil society engagement in other countries. This will depend in part on each country’s circumstances and the space that civil society has in the country.

The World Bank and some other donors face an additional difficulty because their primary interaction is with governments. Nevertheless, it is possible for the Bank to use its convening power to help ensure that the voice of civil society is heard. The Bank can and has showcased to civil society groups in various countries a range of examples of civil society engagement in monitoring and evaluating government performance. The hope is that this will stimulate civil society’s interest and demand for involving itself in this issue (see, for example, Mackay and Gariba 2000). It is important that the donor community, and civil so-

ciety itself, identify additional ways of fostering this type of engagement.

Formal Standards

One final frontier issue is suggested here: whether there is a need to establish **formal standards for good practice M&E systems**. Similar standards are applied by the international community—especially by donors—for national and some sector statistical systems and for government financial management systems.⁷ It would be possible to develop standards for government M&E systems based on the checklist criteria developed as part of the various existing diagnostic guides (chapter 12; Mackay 1998b).

One argument in favor of such formal standards is that this would provide clearer guidance to governments as to what standards they could, and perhaps should, aspire to. A second argument is that it would facilitate periodic monitoring and evaluation of the M&E systems themselves; as argued in chapters 12 and 13, it is important that efforts to strengthen a government M&E system are monitored and evaluated, and the assessment of an M&E system against explicit criteria would facilitate this. One argument against a standardized list of criteria, however, is that there is no single good practice model—or even handful of models—for M&E systems governments should aspire to. Rather, there is a very wide array of possible models.