1

Introduction

A growing number of governments are working to improve their performance by creating systems to measure and help them understand their performance. These systems for monitoring and evaluation (M&E) are used to measure the quantity, quality, and targeting of the goods and services—the outputs—that the state provides and to measure the outcomes and impacts resulting from these outputs. These systems are also a vehicle to facilitate understanding of the causes of good and poor performance.

There are many reasons for the increasing efforts to strengthen government M&E systems. Fiscal pressures and ever-rising expectations from ordinary citizens provide a continuing impetus for governments to provide more government services and with higher standards of quality. These pressures are also reason enough to find more cost-effective ways of operating so that governments can do more with less. Countries in the developing world often look to the richest countries—the members of the Organisation for Economic Co-operation and Development (OECD)—and adopt the public sector management tools that these countries typically employ, such as M&E and performance budgeting. Civil society and parliaments are also putting accountability pressures on governments to publicly report and explain their performances. And international donors are being pressed to demonstrate the results of the large volumes of aid spending for which they are responsible; they in turn are working to persuade and support developing countries to strengthen their own M&E systems.

In recent years, donor support has particularly focused on poorer countries, that is, those that prepare poverty-reduction strategies as part of debt-relief initiatives. These countries are trying to achieve the Millennium Development Goals (MDGs). Donors are also starting to appreciate that country M&E systems can play a role in anti-corruption efforts; these systems help identify "leakages" in government funding, as well as some of the possible consequences of corruption—such as when government spending is not reflected in the physical quality of infrastructure or in the volume and quality of services provided.

M&E, and systems for M&E, are often viewed in narrow, technocratic terms, akin to developing a financial management or a procurement system. There are indeed technical aspects of M&E and M&E systems that need to be managed carefully. But a technocratic emphasis is highly inadequate if it ignores the factors that determine the extent to which M&E information is actually used. Where an M&E system is underutilized, this not only constitutes a waste of resources, but it is also likely to seriously undermine the quality of the information the system produces. It also throws into question the sustainability of the system itself.

Evaluation specialists often argue that M&E and M&E systems are a "good thing" and have intrinsic merit. Management improvement experts often appear to argue that results-based management or other ways of using M&E information offer some sort of panacea for improving government performance. Weak advocacy arguments of this kind are unconvincing to skeptical or overstressed governments in the developing world. However, highly convincing examples do exist of governments that have devoted the necessary effort to building an M&E system. These governments heavily utilize the M&E information the systems produce and have used this information to significantly improve the performance of their policies, programs, and projects.

The purpose of this volume is to help governments in their efforts to build, strengthen, and fully institutionalize their M&E systems, not as an end in itself but in support of improved government performance—in other words, sound governance. The volume brings together the considerable experience accumulated by the World Bank's Independent Evaluation Group (IEG) in its longstanding program of support for governments and Bank staff in their efforts to build M&E systems. It also draws on the growing body of literature from other sources on this issue.

There is currently a great deal of emphasis on *results* and on being able to demonstrate performance. Part I focuses on exactly what M&E has to offer governments; it endeavors to clarify the substance underlying the rhetoric. Chapter 2 provides a brief primer on monitoring and evaluation. Annex E is a continuation of that discussion, presenting the glossary of key terms in evaluation and results-based management developed by the Development Assistance Committee (DAC) of the OECD. To some senior officials and donor staff, M&E can appear to be a highly technical topic with techniques that are difficult for nonspecialists to understand. This chapter endeavors to demystify M&E by providing an eagle's-eye view of a range of different types and methods of M&E.

Chapter 3 follows with a discussion of the specific ways M&E can and has been used to improve government performance. The focus here is on ways governments have used M&E, such as in support of budget decision making. Examples are provided to show civil society's use of M&E to measure and prompt governments to improve their performance. A number of examples of influential evaluations are also provided, and these confirm that M&E can be highly cost-effective for governments. Key trends that are influencing countries to build M&E systems are considered briefly in chapter 4.

Part II focuses on the experience of several countries that have succeeded in building a well-functioning M&E system. Chapter 5 outlines what a "successful" government M&E system looks like; three dimensions of success are proposed. A common question is whether there are actually any countries that have successfully created a national M&E system. The answer to this question is an unambiguous "yes." Many developed and a small but growing number of developing countries have succeeded in building well-performing, whole-of-government M&E systems. Case studies of three such countries—Chile, Colombia, and Australia—are presented in chapters 6, 7, and 8.

The special case of Africa is discussed in chapter 9. Africa is clearly a high priority for the international community, and it has been the focus of a considerable amount of development assistance, including capacity building. Efforts to strengthen M&E have been made in the context of preparation of poverty-reduction strategies. This chapter considers some of the difficulties and opportunities faced by poor countries in these efforts.

Many developed and developing countries have accumulated substantive experience in building monitoring and evaluation systems. As with any form of capacity building, there are a number of hard-won lessons about what works best and what does not, and these are discussed in Part III.

The consistency of these lessons across different countries and Regions is not surprising; the experiences are in accord with international experience with other types of capacity building in the area of public sector management, such as budget systems or civil service reform (chapter 10). As noted earlier, the issue of utilization of M&E information is central to the performance and sustainability of an M&E system. Utilization depends on the nature and strength of demand for M&E information—which depends, in turn, on the incentives to use M&E. Countries whose demand for M&E is low or only lukewarm may be perceived as facing an insuperable barrier to building M&E systems. But this perspective is far too pessimistic; there are ways to increase demand by strengthening incentives, and these are discussed in chapter 11, which focuses on carrots, sticks, and sermons to ensure utilization of M&E information.

Ways that a government M&E system can be strengthened are presented in Part IV. One of the lessons for building an M&E system is the importance of conducting a country diagnosis (chapter 12). That diagnosis provides a sound understanding of the current M&E efforts, the public sector environment, and opportunities for strengthening M&E systems and using M&E information. The information can be used for core government functions such as budget decision making and ongoing management of programs and projects. Perhaps equally important, a diagnosis helps focus key stakeholders within government and the donor community on the strengths and weaknesses of current M&E arrangements. Such a diagnosis can also ensure that these stakeholders share a common awareness of the issues. A diagnosis naturally leads to an action plan that identifies the main options for strengthening a government M&E system (chapter 13).

This volume cannot and does not attempt to address all the issues that can arise when seeking to build a government M&E system or to strengthen an existing system. Part V maps out some issues about which international experience is not well understood or documented. These frontier, cutting-edge issues are topics that are important for the institutionalization of M&E but where current knowledge appears to be insufficient (chapter 14). They include good practice models of M&E at the subnational and sectoral levels; ways governments can effectively and credibly work to support civil society on issues of M&E of government performance; and the cost-effectiveness of alternative models of donor support for the institutionalization of M&E.

Some concluding remarks are provided in chapter 15.

Part VI provides answers to a number of questions that commonly arise at national and international conferences on the topic of this volume. The frequency with which similar questions are raised helps us identify key issues that have to be addressed when seeking to institutionalize an M&E system.

One challenge faced by government and donor evaluation offices alike is to ensure that the evaluations they produce are used intensively. It is now well understood that it is not enough to complete an evaluation report, make it publicly available, and assume that utilization will somehow take care of itself. Instead, individual evaluators and their evaluation offices need to be highly proactive in implementing a detailed strategy for dissemination of not just the evaluation report but its findings as well—to encourage acceptance of the evaluation's findings and implementation of its recommendations. Various lessons and tips for ensuring evaluations are influential are presented in annex A.

Annex B includes an example of a country diagnosis for Colombia. Illustrative terms of reference (ToRs) for a more in-depth diagnosis of Colombia's M&E system are presented in annex C. The findings of a self-evaluation of IEG's support for institutionalizing M&E systems are shown in annex D. This evaluation provides one possible model for donors and governments seeking to evaluate their own efforts to build or strengthen a country's M&E system. Finally, a glossary of key M&E terms is provided in annex E.