56364

The World Bank

PREMnotes

AUGUST
2010
NUMBER 1

Special Series on

The Nuts & Bolts of Government M&E Systems

Conceptual Framework for Monitoring and Evaluation

Keith Mackay

Series Introduction

This is the first note in a monthly series on government monitoring and evaluation (M&E) systems led by the PREM Poverty Reduction and Equity Group under the guidance of Jaime Saavedra, Gladys Lopez-Acevedo, and Keith Mackay, with contributions from several World Bank colleagues. The main purpose of this series is to synthesize existing knowledge about M&E systems and to document new knowledge on M&E systems that may not yet be well understood. The series targets World Bank other donor staff who are working to support client governments in strengthening their M&E systems, as well as government officials interested in learning about the uses and benefits of M&E and in adopting a more systematic approach toward M&E in their governments.

A growing number of governments in developing countries around the world are working to improve their performance by creating systems to measure and help them understand the performance of their services and programs. This trend is influenced by Organisation for Economic Co-operation and Development (OECD) countries, most of which place a high priority on the four main uses of M&E findings:

1. policy development
2. evidence-based policy making and budgeting
3. management performance
4. accountability.

The priority for measuring and better managing government performance in middle- and low-income countries is intensified by continuing fiscal and macroeconomic pressures affecting all countries, and by ever-rising expectations from ordinary citizens. It is also influenced by the need for citizens, governments, and the international community to make state actions more effective in increasing welfare, reducing poverty, and improving opportunities for all. An additional impetus to focus on performance is the strong expectations of international donors.

There is an increasing body of literature on the experience of countries in building and strengthening their M&E systems, and on the M&E tools and techniques that they use. This literature comprises conference proceedings, research papers on country case studies, academic reviews, and numerous Web sites of donors, governments, and evaluation associations. Accessing and digesting the information in its various formats can be difficult, especially for senior officials who are new to M&E.

Why M&E Systems Improve Government Performance

This note[1] outlines the main ways in which M&E findings can be used throughout the policy cycle to improve the performance of government decision making and of government services and programs, including the

use of M&E for evidence-based policy making, budgeting, management, and accountability. There are many different types of M&E tools and approaches, each with advantages and limitations. This note presents four examples of successful government systems for M&E—in both developed and developing countries—and discusses some of their hard-earned lessons for building M&E systems. These lessons are evidence of what works and what doesn't in the development and sustainment of successful M&E systems.

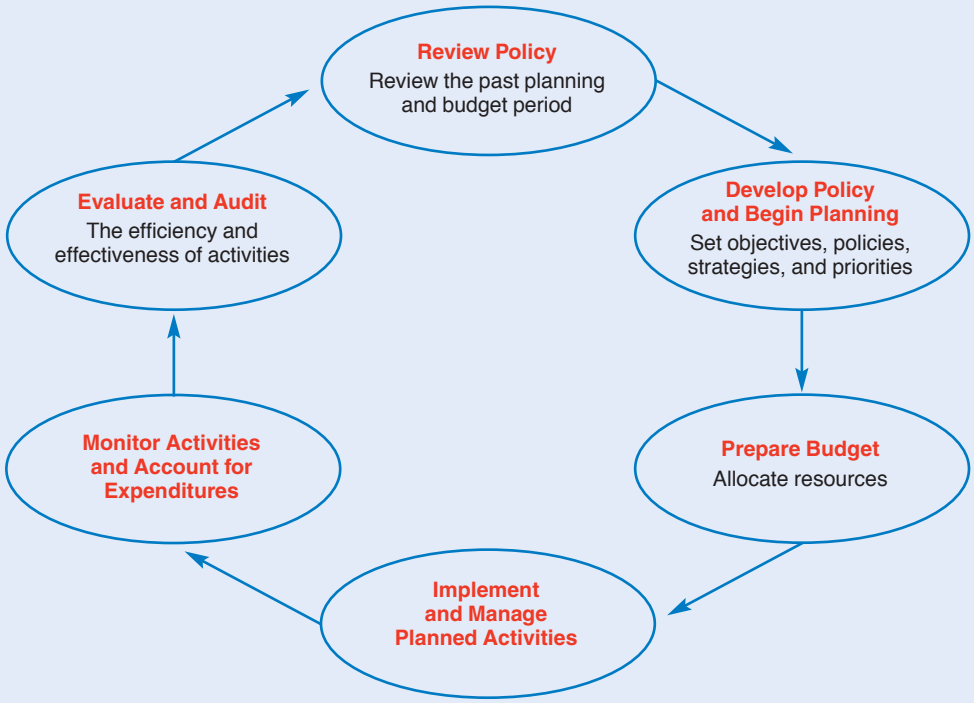
Government systems for M&E focus on measuring the results produced by government—its outputs, outcomes, and impacts. The M&E system may exist at the level of an individual agency, entire sector, or the government as a whole. M&E can provide unique information about the performance of government policies, programs, and projects—at the national, sector, and subnational levels. It can identify what works, what does-not, and the reasons why. M&E also provides information about the performance of a government, of individual ministries and agencies, and of managers and their staff. Highlighting examples of good practice and poor practice can help improve performance.

Three defining characteristics of successful M&E systems are

1. intensive utilization of the M&E information in one or more stages of the policy cycle;
2. information that meets standards for data quality and evaluation reliability; and
3. sustainability, by which the system will survive a change in administration, government ministers, or top officials.

Perhaps the best way to understand the potential contribution of M&E to sound government is to view it at different parts of the policy cycle (figure 1). The early stages of the

Figure 1. The Policy Cycle: Linking Policy, Planning, Budgeting, Management, and M&E



Source: Adapted from World Bank (1998).

policy process—the analysis and development of government policy and the planning of priorities and strategies, including how consistent are they with the objective of poverty reduction—all benefit from evidence of what has or has not worked in the past; in other words, evidence-based policy making. It is an important discipline for governments to consider carefully what they are trying to achieve from their policies, and to plan for them accordingly. Thus it helps to clarify, for each possible program, what success would look like and how the government will know if it has (or has not) been achieved. It also helps to clarify what are the potential poverty and distributional effects of policies and programs. Setting performance targets, and measuring the extent of progress toward achieving them, is thus an important part of government planning and policy review.

Information on the performance of existing government programs, and on the expected performance of new programs, is important for the next stage of the policy cycle: budget decision making—the allocation of resources in the budget. M&E information, and especially evaluation findings that explain past performance, helps to guide government decisions so that the most cost-effective collection of policies and programs can be adopted in the annual budget.

At the next stage in the policy cycle—the implementation and management of activities funded by the budget—M&E helps managers to monitor their activities, including government service delivery and staff management, so that they learn quickly what is working and what is not, for example, in terms of expected outputs, expected outcomes, or even higher level objectives such as increasing welfare. Performance indicators can be used to make cost and performance comparisons—performance benchmarking—among different administrative units, regions, and districts. Ongoing monitoring of these activities—including spending, processes, outputs, outcomes, and impacts—is particularly important. Comparisons made over time can help identify good, bad, and promising practices. Evaluations or reviews can identify the reasons for this good or bad performance. This is the learning function of M&E, often referred to as results-based management.

The final stages of the policy cycle include accountability relationships. M&E reveals the extent to which the government has achieved its objectives and thus provides the evidence needed to ensure strong accountability, such as of government to the congress or parliament, to civil society, and to donors. M&E also supports accountability relationships within government, such as between sector ministries and central ministries, among agencies and their sector ministry, and among ministers, managers, and staff. Strong accountability can provide the incentives necessary to improve performance.

M&E can also play a role in anticorruption efforts. It can help identify "leakages" in government funding, as well as some of the possible manifestations of corruption—such as when government spending is not reflected in the physical quality of infrastructure or in the volume and quality of government services provided.

What Is M&E?

To some senior officials and donor staff, M&E can appear to be a highly technical topic with techniques that at first glance can be difficult for nonspecialists to understand. To make matters worse, M&E contains a lot of buzz words and terminology that are confusing to those not in the field. Sometimes the phrase "M&E" is identified with specific tools, and means different things to different people. Technical and semantic debates are unimportant to the readers of this series; what is important is the appropriate use of the key M&E tools and techniques that can measure the performance of government programs in different ways. Some commonly used tools in-

clude performance indicators, rapid evaluations, impact evaluations, and comprehensive spending reviews. A forthcoming note in this series will discuss M&E tools.

What Does a Government System for M&E Look Like?

Most governments have data systems for measuring their spending, processes, and outputs. But this is not the same as a system to monitor and evaluate the performance of all of its programs. A much smaller number of governments possess such systems. These systems involve the regular, systematic collection and use of M&E information; an M&E system may exist at the level of an individual agency, or an entire sector, or for the government as a whole. The focus of these systems is on measuring the results produced by government—its outputs, outcomes, and impacts.

Fortunately, there are a number of convincing, well-documented examples of governments that have devoted the effort necessary to build high-performing M&E systems. Examples of four countries with well-documented and analyzed M&E systems are Australia, Chile, Colombia, and the United States (see box 1). Each of the four governments utilizes its M&E information intensively. Other common features include a powerful central ministry with leading role in the M&E system (such as the finance or planning ministry), and—as a consequence—an emphasis on using M&E information to support the budget process. What is less well known and documented, however, is how effective these systems have been in supporting the use of M&E information for ongoing program management. Of particular note are the differences between these four M&E examples. Different countries emphasize different M&E tools and techniques. Some place a heavy reliance on monitoring data, others on evaluations and reviews of various kinds. Some countries have extensive (and expensive) M&E systems; others have more streamlined, low-cost systems.

There is no simple answer to the common question: what does a good M&E system look like? Rather, it all depends on country-specific factors, such as government demand for M&E information; the uses to which the information will be put; the availability and quality of existing government data; the abilities of officials and consultants to conduct evaluations and to analyze M&E information; and the amount the government is prepared to spend on M&E.

The Three Defining Characteristics of Successful M&E Systems

A "successful" M&E system has three defining characteristics. **The first is intensive utilization of the M&E information provided by the system in one or more of the stages of the policy cycle** discussed above. It may seem trite to argue that M&E information should only be collected if it is going to be used, but most evaluators in governments (and in donor agencies) have a surprisingly poor understanding of the extent to which the M&E information they produce is actually used by others. If M&E information is not being used then it is important to discover the reasons why. Is it because the M&E information is regarded as being of poor quality, or not timely, or because evaluations have not addressed the most relevant questions concerning program performance? Or is it because the intended users within the government—such as the finance or planning ministries—have neither the skills nor interest in using this information in their work?

Reliable, quality information is another feature of successful M&E systems. There are various standards of what constitutes quality monitoring data and evaluations, and these standards can be used to assess the reliability of the information that any M&E system produces. Most government evaluation offices have some sort of quality control mechanism in place. Most, however, do not

Box 1. Four Examples of Successful Government M&E Systems

Australia: The government evaluation system was managed by the Department of Finance (DoF), and it required ministries to evaluate every program every three to five years. The evaluations were conducted by the line ministries themselves, but they were overviewed by the DoF and other central departments. By 1994, almost 80 percent of new spending proposals in the budget process relied on evaluation findings, usually to a significant degree. About two-thirds of savings options also relied on evaluation findings. DoF officials, who attended the Cabinet meetings that considered these budget proposals, judged that this information was highly influential on the Cabinet's budget decision making. The Australian National Audit Office found that line departments also used this information intensively, particularly to help improve their operational efficiency.

Chile: The Ministry of Finance (MoF) commissions the evaluations externally to academics and consulting firms, and it uses standardized terms of reference and methodologies for each type of evaluation it uses. MoF officials use the M&E findings intensively in their budget analysis of the performance of each ministry and agency. The ministry also uses the information to set performance targets for each agency and to impose management improvements on both ministries and agencies. The MoF carefully oversees the extent to which each ministry implements these management improvements.

Colombia: The government's M&E system, SINERGIA, is managed by the Department of National Planning. The system includes information for 500 performance indicators, as well as a number of rapid and impact evaluations. The president has used the information provided by SINERGIA intensively in his monthly management control meetings with each minister and in his weekly town hall meetings in municipalities around the country.

United States: In 2002, the government created the Program Assessment Rating Tool (PART), building on earlier efforts to measure government performance. The performance of all 1,000 government programs have been rated using the PART methodology, and PART ratings are required to be used by departments in their annual budget funding requests to OMB. The requests must highlight the PART ratings, the recommendations for improvements in program performance, and performance targets. OMB, in turn, also uses the PART ratings when it prepares the administration's funding requests to the Congress, and to impose performance improvement requirements on departments.

Source: Mackay 2007

appear to conduct or commission formal reviews of the quality of their work. Three of the four countries highlighted in box 1 have conducted such reviews: Australia, Chile, and Colombia.

The third characteristic of a successful M&E system is sustainability. This relates to the likelihood that the M&E system will survive a change in administration or in government ministers or top officials. When the utilization of M&E information is firmly embedded—that is, mainstreamed—in core government processes, such as the budget cycle, it can be said to be institutionalized and thus is likely to be sustained over time. Conversely, when M&E has only a handful of key supporters or is little used, or if it is largely funded by donors rather than by the government itself, then sustainability is less likely.

Building a Government M&E System—What to Do and What Not to Do

Many developed and developing countries have accumulated substantial experience in building M&E systems. As with any form of capacity building, there are a number of hard-earned lessons about what works best and what does not (discussed more fully in Mackay [2007]). Eight key lessons are discussed below.

Lesson 1: First and foremost is the need for substantive government demand for M&E information. Such demand is necessary

if a serious effort to build an M&E system is to be started and sustained. A significant effort is required to build an M&E system, including the creation or upgrading of data systems such as decisions about types of data to be collected, data collection methods, storage, quality control, and transmission. Equally important components of the M&E system include the training of statistical analysts; choice of evaluation tools and techniques, and their adaptation to local circumstances and priorities; training of evaluators and development of national evaluation consultants; creation of M&E offices inside a lead ministry and probably in some or all sector ministries; training of the users of M&E information—mid-level analysts, senior officials in central and sector ministries, and possibly their ministers; and the creation of a bureaucratic infrastructure to decide which government programs should be evaluated and what issues should be addressed in each evaluation. Frankly, this effort is not worthwhile unless the resulting M&E information is likely to be used intensively.

Lesson 2: Incentives are a key part of the demand side. Strong incentives are needed for M&E to be conducted, and for the information to be used. M&E experts often make a basic mistake by asserting that M&E information is intrinsically "a good thing" and that if the information is made available, then it will automatically be used. This technocratic view that M&E has inherent merit is naïve; M&E information has value only if it is reliable and if it is used intensively.

Utilization does not usually—and does not regularly—happen by chance. There need to be incentives for M&E information to be used by program managers in their day-to-day work, by budget and planning officials responsible for advising on policy options, or by a congress or parliament responsible for accountability oversight.

There are three types of incentives: carrots, sticks, and sermons.[2] An example of a carrot is the provision of greater autonomy to managers who can demonstrate (through reliable M&E information) that their programs are performing well. An example of a stick is to set challenging (but realistic) performance targets that each ministry and program manager is required to meet. An example of a sermon is a high-level statement of support for M&E, such as from a president or influential minister. Many of these incentives have been applied successfully in building M&E systems in developed and developing countries.

Lesson 3: It helps to start with a diagnosis of what M&E functions already exist in the country—in the government, academia, and the consulting community. A diagnosis should identify the strengths and weaknesses of what exists on both the demand and supply sides. A diagnosis is really a type of evaluation, and the very process of conducting it provides an opportunity for key stakeholders within the government to become more familiar with M&E and its potential benefits to the government. A diagnosis naturally leads to an action plan to strengthen M&E. This can facilitate a coalition of support from interested sector ministries and the donor community.

Lesson 4: Another dimension of the demand side is the need for a powerful champion, an influential minister or senior official who is able to lead the push to institutionalize M&E, to persuade colleagues about its importance, and to allocate significant resources to creating a whole-of-government M&E system. Government champions have played pivotal roles in some of the most successful M&E systems.

What has been a lot less successful is reliance on a law, decree, or cabinet decision to institutionalize an M&E system. Such an approach can help to legitimize an M&E system, particularly in those countries where the presence of a legal instrument is viewed as necessary for any government reform to

be perceived as worthwhile and to be taken seriously. But a law or decree on its own does not ensure that the considerable efforts required to build an M&E system will be undertaken and maintained.

Lesson 5: Another common feature of successful M&E systems is the stewardship by a capable ministry that can design, develop, and manage the system. Thus it helps to have the institutional lead of the M&E system close to the center of a government, such as in the president's office or in the finance or planning ministries.

One role of this institutional leader is to continually review the extent of progress in developing the M&E system and make any necessary adjustments to its action plan. Difficulties and roadblocks are inevitable, so it is important to identify what is working, what is not, and why. Regular reviews of progress provide the opportunity to analyze both the demand and supply sides of the M&E system. In fact, most countries with well-performing M&E systems have not actually developed them in a linear manner—that is, starting with a clear understanding of what the system would look like once fully mature and then progressively achieving this vision. Rather, countries' M&E systems are more commonly developed incrementally and even in a piecemeal and opportunistic manner, with some false starts and blind alleys along the way.[3]

Lesson 6: A common mistake once M&E has been embraced enthusiastically is to overengineer the M&E system. This is often evident in the large number of performance indicators that are collected. Over-engineering can also result in the proliferation of ministry data systems. These are often uncoordinated even within each ministry. The problem is multiplied if there are several whole-of-government data systems, which may be managed by different central ministries and may well require related (but different) information to be provided by sector ministries and agencies. In Mexico, for example, the social development agency (SEDESOL) had eight different, uncoordinated management information systems. In Uganda in recent years, there were as many as 16 separate sector and subsector systems that the government had to try to coordinate.

That said, there is real value in building reliable ministry data systems: these provide the raw data on which a whole-of-government M&E system depends. An audit of data systems and a diagnosis of data capacities can be helpful in this situation because they could provide a basis for rationalizing existing data collections and improving their quality.

Lesson 7: This is, unsurprisingly, that building an M&E system usually includes training for a range of M&E tools, methods, approaches, and concepts. It is necessary to have well-trained officials or consultants who are highly skilled in M&E. Training should provide more than competencies in M&E, however. Senior officials need to understand the strengths and limitations—the relative cost-effectiveness—of various types of M&E tools and techniques. Introductory training can also raise awareness of and demand for M&E information. Training should also extend to the use of M&E findings. Budget analysts, poverty analysts, and program managers need to be able to interpret monitoring data to understand trends, data definitions, breaks in data time series, and so forth. They also need to be discriminating consumers; they must be able to tell when an evaluation is reliable or when its methodology or findings are questionable.

Lesson 8: A final lesson is that it requires a long-term effort, with patience and determination, to build an effective M&E system. It takes time to create or strengthen data systems; to recruit and train qualified staff; to plan, conduct, and manage evaluations; and to train staff to use M&E in their day-to-day work, whether that involves program opera-

tions or policy analysis and advice. Australia and Chile were able to create well-functioning M&E systems—in terms of the quality, number, and utilization of evaluations—within four or five years; but in Colombia's case, it has taken more than a decade.

Conclusions

A growing number of developing countries are successfully building government M&E systems. They look to the examples of developed countries—especially members of the OECD—but increasingly they are also looking to their peers: countries such as Chile, Colombia, and Mexico. These countries have created well-performing M&E systems where the M&E findings that are produced are used intensively. These countries have demonstrated not only that it is feasible to build a government M&E system, but that the systems are also valued highly by the governments.

Hopefully, this short note and the future notes in this series will engage the interest of senior officials in developing countries and prompt them to fully investigate whether their government should devote the time and effort to building such a system. The donor community—including the World Bank—stands ready to support them in this work.

About the Author

Keith Mackay *retired from the World Bank in 2008 and now works as a consultant. He was a lead evaluation officer in the Bank's Independent Evaluation Group, providing technical advice and other support to countries working to strengthen their M&E systems as an integral part of sound governance. Before joining the Bank in 1997, he was the manager of Australia's whole-of-government evaluation system. He has written over 60 publications on M&E systems.*

Endnotes

1. This note has benefited from comments by Gladys Lopez-Acevedo, Nidhi Khattri, Jaime Saavedra, and Helena Hwang. The views expressed here are those of the author.

2. A long list of M&E incentives is provided by Mackay (2007, chapter 11).

3. Reasons for this are discussed more fully by Mackay (2007).

References

IEG (Independent Evaluation Group). 2004a. "Influential Evaluations: Evaluations That Improved Performance and Impacts of Development Programs." World Bank, Washington, DC.

———. 2004b. "Monitoring and Evaluation: Some Tools, Methods and Approaches." 2nd edition. World Bank, Washington, DC.

Lopez-Acevedo, Gladys, Katia Rivera, Lycia Lima, and Helena Hwang, eds. 2010. "Challenges in Monitoring and Evaluation: An Opportunity to Institutionalize M&E Systems." World Bank and Inter-American Development Bank, Washington, DC.

Mackay, Keith. 2007. *How to Build M&E Systems to Support Better Government*. Washington, DC: World Bank.

World Bank. 1998. *Public Expenditure Management Handbook*. Washington, DC: World Bank.

World Bank Web Sites

www.worldbank.org/poverty

www.worldbank.org/lacmonitoringandevaluation

www.worldbank.org/ieg/ecd

PREM network
Poverty Reduction and Economic Management

This note series is intended to summarize good practices and key policy findings on PREM-related topics. The views expressed in the notes are those of the authors and do not necessarily reflect those of the World Bank. PREMnotes are widely distributed to Bank staff and are also available on the PREM Web site (http://www.worldbank.org/prem). If you are interested in writing a PREM*note*, email your idea to Madjiguene Seck at mseck@worldbank.org. For additional copies of this PREM*note* please contact the PREM Advisory Service at x87736. PREM*notes* are laid out by UpperCase Publication Services, Ltd.

Prepared for World Bank staff