



10

Building Country M&E Systems— Lessons from Experience

The growing literature on experience with strengthening government M&E systems suggests that there is broad agreement among experts on a number of key lessons (box 10.1) (see African Development Bank and World Bank 1998; Boyle 2005; Compton, Baizerman, and Stockdill 2002; DAC 2006; Development Bank of Southern Africa, African Development Bank, and World Bank 2000; Mackay 1998d, 2004; May and others 2006; OECD 1997a, 1997b, 1998a, 2004; Sciavo-Campo 2005; UNDP 2000).

The first and foremost lesson is that **substantive demand from the government is a prerequisite to successful institutionalization**. That is, an M&E system must produce monitoring information and evaluation findings that are judged valuable by key stakeholders, that are used to improve government performance, and that respond to a sufficient demand for the M&E function to ensure its funding and sustainability for the foreseeable future.

Efforts to build an M&E system will fail unless real demand exists or can be intentionally created, especially by ensuring that powerful incentives are in place to conduct and use M&E. It is not enough to issue an edict that M&E is important and should be done; this is likely to produce only lip service and is certainly unlikely to produce good quality monitoring information and evaluations. Such efforts at top-down compliance can, unless accompanied by a range of other actions, easily lead to ritual compliance or even active resistance.

Achieving substantive demand for M&E is not easy. And a barrier to demand is lack of knowledge about what M&E actually encompasses, particularly where the buy-in of key stakeholders such as government ministers or finance ministries is necessary before substantive effort will be put into creating and funding an M&E function. So there is frequently a chicken-and-egg problem: There is a lack of government demand for M&E because of the lack of understanding of M&E and what it can provide; there is a lack of understanding because of the lack of experience with it; and there is a lack of experience because of weak demand.

The way around this conundrum is to increase awareness of M&E—its range of tools, methods, and techniques—and its potential uses. Demand can be increased once key stakeholders in a government begin to understand it better, when they are exposed to examples of highly cost-effective monitoring systems and evaluation reports, and

Box 10.1: Elements of Successful Country M&E Systems

- Substantive government demand
- Strong role for incentives
- Diagnosis of existing M&E as first step
- Key role of a powerful “champion”
- Centrally driven by capable ministry
- No overengineering the system
- Reliable ministry data systems
- Utilization as the measure of success
- Training in M&E and in using M&E
- Limitations of relying on government laws, decrees, and regulations
- Structural arrangements ensure M&E objectivity and quality
- A long-term effort, requiring persistence
- Development in a nonlinear, less predictable manner
- Regular evaluation of M&E system itself.

when they are made aware of other governments that have set up M&E systems that they value highly. It can be persuasive to point to the growing evidence of very high returns on investment in M&E (Bamberger, Mackay, and Ooi 2004).

The supply side is also important—provision of M&E training, manuals, and procedures and identification of good M&E consultants, and so forth; M&E expertise is certainly necessary if reliable M&E information is to be produced. Those who view M&E in technocratic terms as a stand-alone technical activity tend to focus only on these issues.

But the supply side of producing M&E information is less important than demand. If demand for M&E is strong, then improving supply in response can be relatively straightforward, but the converse does not hold. Demand from governments is crucial for the utilization of M&E information and for the ongoing funding of M&E activities. Thus, it is necessary to secure the buy-in of the key stakeholders such as government ministers or finance ministries if substantive effort is to be put into creating, strengthening, or funding M&E func-

tions. Continuing reliance on donor funding and on donor M&E activities is not sustainable.

Incentives are an important part of the demand side. There need to be strong incentives for M&E to be done well, and in particular for monitoring information and evaluation findings to be actually used. In other words, strong incentives are necessary if the M&E function is to be successfully institutionalized (see box 10.2 on Ireland’s experience, for example). This observation is also consistent with the extensive literature on achieving any type of institutional reform, particularly in the context of public sector management and sound governance (for extensive literature reviews see World Bank 1997a; OECD 2004, 2005).

Simply having M&E information available does not guarantee that it will actually be used, whether by program managers in their day-to-day work, by budget officials responsible for advising on spending options, or by a Congress or Parliament responsible for accountability oversight. This underscores both the dangers of a technocratic view, which sees M&E as a set of tools with inherent merit, and the fallacy that simply making M&E information available would ensure its utilization.

No governments build M&E systems because they have intrinsic merit. Governments build M&E systems because (1) those systems directly support core government activities, such as the budget process; national planning; development of policies and programs; and the management of ministries, agencies, programs, and activities, or (2) provide information in support of accountability relationships. Thus, M&E systems are often linked to public sector reforms such as performance-based budgeting (PBB), evidence-based policy making, results-based management, and the like; such initiatives share a number of common elements (see chapter 3).

In working to build or strengthen a government M&E system, it helps to **start with a diagnosis of what M&E functions currently exist** and their strengths and weaknesses—both on the demand

Box 10.2: Lessons from Ireland

Ireland developed its government evaluation system in response to the formal requirements for accession to the European Union. Since its accession, Ireland's system has been strengthened for internal reasons related to the government's wish to improve the value for money obtained from all areas of public expenditure. This has been reflected in the government's Expenditure Review Initiative.

For developing countries, Ireland provides a number of lessons about both success factors in and impediments to developing an M&E system. One lesson is that strong external pressures, linked to the availability of significant resources, can be a key catalyst in initiating an M&E system. (An analogy for poor countries is the requirement to prepare poverty-reduction strategies, with related M&E systems, in the context of debt relief under the Heavily Indebted Poor Country Initiative.) Once in existence, an M&E system can be used for additional, nationally driven purposes. Of course, country demand and incentives to utilize M&E information are very important for the institutionalization and sustainability of such a system. The case of Ireland again underlines the difficulty of ensuring a direct link between M&E information and budget decision making and other resource-allocation processes. Formal procedures and practices may be necessary to establish direct links.

Source: Boyle 2005.

Another lesson is the merit of periodically reviewing progress in developing such a system and reorienting the system—sometimes substantially—as a result. Ireland is continuing to pilot further improvements to its evaluation system.

Ireland has a small pool of evaluators, and this has been a constraint on the system, although it has enabled Ireland to develop rapidly its understanding of the requirements of the system. The small skills pool has underlined the importance of using this resource carefully; it has implications on both the demand and supply sides.

On the demand side, it suggests the importance of not trying to develop an overly complex or demanding system. Instead, it is better to focus on the most cost-effective M&E activities; these are determined by the likely utilization of the M&E information produced. On the supply side, there would be merit in working to expand the limited capacities in a planned manner—for example, through targeted training, curriculum development, on-the-job skills development, secondments, networking support, regular review of M&E quality, or period contracts with consulting companies. Where there is reliance on civil servants to undertake evaluations, it is particularly important to ensure that they are sufficiently trained and are provided with adequate guidelines and other support to enable them to function effectively.

and supply sides. The extent of actual (as distinct from the hoped-for) extent of utilization of M&E information must be clear, as well as the particular ways in which it is being used. Such diagnoses are themselves a form of evaluation, and they are useful not just for the information and insights they provide, but also because they can be a vehicle for raising the awareness of stakeholders in government, civil society, and the donor community about the importance of M&E and the need to build a new system or strengthen existing systems.

Another dimension to the demand side, and another success factor, is having a **powerful champion**—a powerful minister or senior official who is able to lead the push to institutionalize M&E, to persuade colleagues about its priority, and to

devote significant resources to creating a whole-of-government M&E system. A champion needs to have some understanding of M&E, in terms of tools and methods, and an appreciation of its potential usefulness for government—for one or more of the four main uses of M&E information (outlined in chapter 3).

Government champions have played important roles in the creation of some of the more successful government M&E systems, such as those of Chile, Colombia, and Australia (discussed for each country, respectively, in May and others 2006; IEG 2003b [chapter 11]; Mackay 2004). However, powerful champions constitute a success factor; they do not provide a guarantee of success. There are examples, such as Egypt, where the support of a

group of key ministers for M&E has been substantially frustrated by skeptical mid-level officials (IEG 2004a; Schiavo-Campo 2005).

Creating a whole-of-government M&E system—whether focused solely on a system of performance indicators or encompassing various types of evaluation and review—requires a significant effort. It involves recruiting and training staff to conduct or manage M&E and use their findings; creating the bureaucratic infrastructure to decide which government programs should be evaluated and what issues should be addressed in each evaluation; and creating data systems and procedures for collecting, sharing, and reporting M&E information.

It also requires active oversight of the M&E system by senior—and usually very busy—officials. Like other systems, in areas such as financial management or procurement, it takes sustained effort over a period of years to make an M&E system operate efficiently. The OECD has concluded that—

It takes time to develop a performance measurement system and to integrate it into a management system. No OECD member country considers developing a performance measurement system as easy. On the contrary, it is perceived as an exhausting exercise which needs constant monitoring and controlling (OECD 1997a, p. 19).

Thus, another feature of successful government M&E systems is the **stewardship of this process by a capable ministry** that can design, develop, and manage an M&E system. In many developed and upper middle-income countries (for example, Australia, Canada, Chile, and the United States) this has meant the finance ministry. It certainly helps to have the institutional lead of an M&E system close to the center of government (for example, a president's office) or the budget process (Bedi and others 2006). The need to have clarity concerning the roles and responsibilities of key stakeholders is also reflected in the Latin American experience (box 10.3).

In some countries, capable sector ministries have set up strong M&E systems. Perhaps the most

notable example is in Mexico, where the Secretariat for Social Development (SEDESOL), a capable and respected ministry, manages an M&E system that emphasizes both qualitative and impact evaluations. The ministry is also working to strengthen its system of performance indicators to better support the evaluations it conducts (Hernandez 2006).

The genesis for this sector ministry effort was a law passed by the Congress mandating the evaluation of social programs; Congress was concerned that the executive government might use its social programs to buy votes, and it wanted solid evidence of program performance before it would agree to fund programs. This law was influenced at least in part by the series of rigorous impact evaluations of the *Progres*a program.

Although these are among some of the most expensive impact evaluations ever done, costing millions of dollars, their quality has also been widely acknowledged, as has their enormous impact on the government. They were instrumental in persuading the government to retain the *Progres*a program and to expand significantly when it morphed into the *Oportunidades* program—by 2005, the government was spending about \$6 billion on this program, which covers some 21 million beneficiaries, or about one-fifth of the Mexican population.

These evaluations, although expensive, can be viewed as having been very cost-effective. Governments in other countries find examples of highly influential evaluations to be quite persuasive in relation to the potential usefulness of evaluation and the merits of setting up a sound M&E system.

As noted in chapter 3, the success of M&E in SEDESOL has helped persuade the powerful finance ministry and the comptroller's office to join the national evaluation council to create a whole-of-government M&E system. This indicates the powerful demonstration effect a successful sector agency can have.

One point to note in passing: it is rarely if ever the case that a ministry that decides to create a strong

Box 10.3: Latin American Experience with M&E Systems

A World Bank–Inter-American Development Bank (IADB) Regional conference held in June 2005 focused on the experience of five leading or promising countries in Latin America: Brazil, Chile, Colombia, Mexico, and Peru. Several of these countries have achieved considerable success in creating M&E systems and institutionalizing them.

It was evident from the country experiences that there is no single “destination” for countries in terms of what a well-performing M&E system looks like. Some countries stress a system of performance indicators, and others focus on carrying out evaluations (program reviews or rigorous impact evaluations). And although most countries have created a whole-of-government approach driven by finance or planning ministries, some are more focused on sector M&E systems. One key characteristic of most of the systems in the Region is that they reflect country-led, rather than donor-driven, efforts to institutionalize M&E.

The shared experience of these countries has led to some collective wisdom about the development of solid M&E systems:

- There is a need to clearly define the roles and responsibilities of the main actors—the planning and finance ministries, the president’s office, sector ministries, and Congress. It is also critical to create the right incentives to encourage these stakeholders to assume a greater role in M&E.
- Strengthening M&E systems is not only, even principally, a supply-side issue requiring a “technical fix.” For an M&E system to be successful and sustainable, the information and findings of M&E have to be utilized intensively by all stakeholders, including sector ministries and, depending on how the system has been devised, civil society.
- Conservative ministries and staff may resist efforts to implement M&E systems and use M&E as a management and budget tool. It is essential to have a high level of commitment and ongoing support from powerful champions at the ministerial and senior official levels.
- There is an implicit debate on how to prioritize evaluations—to focus on problem programs, pilot projects, high-expenditure or high-visibility programs, or on systematic research to respond to questions of program effectiveness.

Source: May and others 2006.

M&E system has to start from scratch. Even in the poorest African countries there is usually a range of performance indicators available, and there will also be a number of evaluations conducted by donors. The problem is more the poor quality and partial coverage of performance information and its substantial underutilization.

A common mistake is to overengineer an M&E system. This is more readily evident with performance indicators—for example, Colombia’s M&E system, SINERGIA, had accumulated 940 performance indicators by 2002; for Colombia, this number was unwieldy for its accountability uses of the information, and it has subsequently reduced the number to around 500 (Castro 2006a, 2006b; Mackay and others 2007; annex B of this volume). In contrast, Chile’s M&E system includes 1,550 performance indicators. This is a very large number, but the highly capable finance ministry is able to use the information effectively. It would be best to regard Chile as the exception that proves the more general rule, that less is more.

The appropriate number of performance indicators also depends on the number of government programs and services and on the type of performance indicator. Indicators can focus on the various parts of the “results chain” for a program (defined in annex E): inputs, activities, outputs, outcomes, and impacts. Senior officials would tend to make use of high-level strategic indicators such as outputs and outcomes. Line managers and their staff, in contrast, would tend to focus on a larger number of operational indicators that target processes and services.

The regular production of this kind of detailed performance information is only a first step. Senior officials and managers would usually not have the time to scrutinize raw data. It can assist their work considerably if capable analysts review the data and provide summaries and analyses for senior officials.

Ministry data systems are often—and perhaps typically—uncoordinated, so a single ministry can

possess several such systems, each with its own data definitions, data sources, periodicity of collection, and quality assurance mechanisms (if it has any). Mexico's SEDESOL, for example, had eight separate management information systems in 2004 (World Bank 2004c). In Uganda, one problem is the number of uncoordinated M&E systems—as many as 16 separate sector and subsector systems, which the government is now working to coordinate through a new national integrated M&E strategy (Hauge 2003; Government of Uganda 2006).

A problem in African countries, and perhaps in some other Regions, is that although sector ministries collect a range of performance information, the quality of data is often poor. This is partly because the burden of data collection falls on overworked officials at the facility level, who must provide the data for other officials in district offices and the capital but who rarely receive any feedback on how the data are actually being used, if at all.

This leads to another chicken-and-egg problem: Data are poor partly because they aren't being used; and they're not used partly because their quality is poor. In such countries there is too much data, not enough information. Thus, another lesson for the institutionalization of a government M&E system is the **need to build reliable ministry data systems**—to help provide the raw data on which M&E systems depend.¹ An audit of data systems and a diagnosis of data capacities can be helpful in this situation. It would provide the starting point for any necessary rationalization of data collections or improvements in their quality. It would give the data the credibility necessary to be used.

Data verification and credibility is partly a technical issue of accuracy, procedures, and quality control. Related to this issue of technical quality is the need for data to be potentially useful—for information to be available on a timely basis, easy to understand, consistent over time, and so forth. There is also an issue of honesty and objectivity; as performance information becomes important—particularly when it is used for accountability purposes—there will arise incentives to falsify the

data.² This is clearly of concern and requires specific measures to verify the data, such as through independent data audits. Verification could also involve accreditation of an agency's processes for data verification. Of course, efforts to falsify data can also be taken as a sign of success that the M&E system is starting to have an impact—that it is having real “bite.”

Financial data on program spending comprise a fundamental type of information; quality and ready availability must be ensured. This information is supplied by financial management information systems. Sound performance budgeting, program management, and government accountability all require that information on the costs and results of government programs be linked. Ironically, many evaluations then fail to discuss the cost of the programs being evaluated; this makes it harder to gauge the cost-effectiveness of the program.

The objective of government M&E systems is never to produce large volumes of performance information or a large number of high-quality evaluations per se; this would reflect a supply-driven approach to an M&E system. Rather, the objective is to achieve intensive utilization of whatever M&E information exists to ensure that the M&E system is cost-effective—utilization in support of core government functions, as noted earlier. **Utilization is the yardstick of success of an M&E system**; conversely, it would be hard to convince a skeptical finance ministry that it should continue to fund an M&E system whose outputs are not being utilized. Such systems would deservedly be regarded as useless.

For an M&E system to perform well, it is necessary to have well-trained officials or consultants who are highly skilled in M&E. For this reason, most capacity-building plans place considerable emphasis on **provision of training in a range of M&E tools, methods, approaches, and concepts**. Those governments that contract out their evaluations need to ensure that their officials possess the skills and experience necessary to oversee and manage evaluations—this requires a broader set of competencies than the ability to simply conduct an evaluation. They also need to

understand the strengths and limitations—the relative cost-effectiveness—of various types of M&E.

Introductory training in M&E can also raise awareness of and demand for M&E information. Training should also extend to the *use* of M&E information. Budget 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 of evaluations; that is, they must be able to tell when an evaluation’s methodology is flawed and its findings unreliable. Finally, parliamentarians and their staff need to be able to understand M&E information.

Another lesson is that there are **limitations when a country relies on a law, decree, cabinet decision**, or other high-level pronouncement to create an M&E system. In Latin American and francophone countries—those with the Napoleonic system of law—there is a tradition of relying on such legal instruments to create and legitimize M&E systems.³ Thus, countries such as Colombia have a series of laws and decrees mandating evaluation; these were even enshrined in the Constitution in 1991. Yet in the intervening years, the fortunes of the government’s evaluation system have waxed and waned, and it was only after a change in government in 2002 that the system started to perform strongly (Castro 2006a, 2006b; Mackay and others 2007).

The point here is not that a law or decree mandating M&E is irrelevant: on the contrary, these can be useful vehicles for legitimizing M&E, particularly in those countries where the presence of such a legal instrument is viewed as necessary for any government reform to be perceived as worthwhile and taken seriously.⁴ 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.

The structural arrangements of an M&E system are important from a number of perspectives. One is the need to ensure the objectivity, credibility, and rigor of the M&E information the system produces. On the data side, some governments (for example, Chile) rely on external

audit committees to perform this function, some rely on the national audit office (for example, Canada) (Mayne and Wilkins 2005), and some rely principally on internal ministry audit units (for example, Australia). Some rely on the central ministry checking data provided by sector ministries (for example, Colombia), and others have no audit strategy (for example, Argentina) (Zaltsman 2006a).

On the evaluation side, issues of objectivity and credibility are particularly important. As noted in box 3.1, Chile (and most other Latin American countries) deals with this by contracting out evaluations to external bodies such as academic institutions and consulting firms; moreover, the evaluations are commissioned and managed by the finance ministry rather than by sector ministries, and the process of seeking bids and awarding contracts to conduct the evaluations is entirely transparent.⁵

The downside of this approach is a lack of ownership of these evaluation findings by the sector ministries, which do not make much use of the evaluations commissioned by the finance ministry. That may not be so great a problem in Chile, however, where the powerful finance ministry is able to use evaluation findings not only to support budget decision making, but also to impose management and program changes on the sector ministries (Rojas and others 2005). This centrally imposed system is unique.

Most OECD governments rely on sector ministries to conduct evaluations themselves (Curstine 2005), although this raises questions about the reliability of self-evaluations. In the United States, the OMB (the finance ministry) rates the performance of government programs and marks those programs with either no M&E information about their performance or with unreliable M&E information—see box 3.1.

Countries that have built a government M&E system have found that it is a **long-haul effort, requiring patience and persistence** (OECD 1997a; Mackay 1998b; Lahey 2005; May and others 2006). It takes time to create or strengthen data systems; to train or recruit qualified staff; to plan, manage, and conduct evaluations; to build systems

for sharing M&E information among relevant ministries; and to train staff to use M&E information in their day-to-day work, whether that involves program operations or policy analysis and advice. Australia and Chile were able to create well-functioning evaluation systems (in terms of the quality and number and utilization of the evaluations) within four or five years; but in Colombia's case, it has taken more than a decade.

This is not to say that a slow and measured approach to building an M&E system is appropriate, however. Government champions will eventually depart, and the window of opportunity—indeed, the priority a government gives to any type of public sector reform—can close as quickly as it opened. This suggests that an approach of working in a focused, purposeful, and even intense manner to build various components of the M&E system is necessary, and that governments need to institutionalize them as quickly as possible.

It appears that **most countries with well-performing M&E systems have not 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. Instead, when one examines the experience of countries such as Australia (Mackay 2004), Canada (Lahey 2005), Chile (Zaltsman 2006a; see also chapter 7), Colombia (Mackay and others 2007), Ireland (Boyle 2005), and the United States (Joyce 2004; Lahey 2005), it is evident that these countries' M&E systems have been developed incrementally and even in a piecemeal manner, with some false starts and blind alleys along the way.

This would appear to be caused partly by the different amounts of time it takes to build particular M&E functions—a system of performance indicators relative to the conduct of program reviews or rigorous impact evaluations. It would also appear to be caused by a number of mid-course corrections made as the progress, or lack of progress, with particular M&E initiatives becomes evident. There is also the important influence of external factors, such as a change of government, which can not only alter the direc-

tion of an M&E system but can lead to it being significantly strengthened—such as in Colombia after 2002 and in the United States after 2000. A change in government can also result in an M&E system being substantially run down or even abandoned—such as in Australia after 1997 (chapter 8) and the United States after 1980 (GAO 1987).

There appears to be a rather worrying asymmetry with government M&E systems; they are slow to build up but can be run down quickly. For governments that have largely abandoned their M&E system, this would appear to reflect an ideological preference for “small government” rather than a considered decision about the cost-effectiveness of the M&E system; the negative effects on the M&E system thus appear simply to be collateral damage.

The frequency of mid-course corrections as M&E systems are being built indicates another lesson from experience: the **value of regularly evaluating an M&E system** itself, with the unsurprising objective of finding out what is working, what is not, and why. It is valuable to identify the nature of the roadblocks encountered and to indicate possible changes in direction. Such evaluations provide the opportunity to review both the demand and the supply sides of the equation and to clarify the extent of actual—as distinct from the hoped-for—extent of utilization of M&E information, as well as the particular ways in which it is being used.

The Chilean finance ministry's careful stewardship of that country's M&E system is exemplified by the review it commissioned the World Bank to conduct into the two principal evaluation components of the system (Rojas and others 2005). It commissioned this review partly to support the ongoing management and improvement of the M&E system and partly to apply the same standards of performance accountability to itself as it applies to sector ministries and the programs they manage—the finance ministry has, as a matter of course, reported the World Bank's evaluation findings to Chile's Congress. There are a number of diagnostic guides available to support such evaluations of government M&E systems (see chapter 12).⁶