



# Conducting Diagnoses of M&E Systems and Capacities



Geoffrey Shepherd

*A diagnosis of a country's monitoring and evaluation (M&E) activities is indispensable if that country is to develop projects or policy proposals to improve the impact of M&E. This note provides a guide to some of the topics that need to be considered when undertaking such a diagnosis. These topics emphasize both the institutional analysis of factors that drive demand for M&E and the technical factors that drive supply. For each of the nine topics identified, this note discusses some of the major issues analysts would need to consider. This discussion principally centers on reviewing how these issues were covered in recent M&E studies for a number of developed and developing countries.*

Monitoring and evaluation (M&E) systems are an important building block of a performance-oriented policy cycle in which policy goals are developed based on public interest and policies are designed and implemented, to the extent possible, in ways that will make them effective, efficient, and consistent. The more policy-cycle activities are based on empirical evidence and analysis (of what has or has not worked elsewhere, or is not working here), the more likely policies will be efficient, effective, and in the public interest. Demand for M&E depends on the presence of at least some elements of a performance-oriented policy cycle.

An M&E diagnosis is an analysis of what is and is not working in a country's (or a sector's, or a region's) M&E, with recommendations for improving activities and systems. The purpose of the diagnosis is to identify needs and find the combination of institutional arrangements and technical capacities that can address these needs.

The diagnosis emphasizes an institutional analysis of the factors that drive demand for and supply of M&E. Institutions are the formal and informal rules that shape the behavior of individuals and organizations in society. If there is little or no demand for better public sector performance, an M&E diagnosis will probably be of little use. But in

countries (or sectors or subnational governments) where there is demand for better performance, an M&E diagnosis can pave the way. A good M&E diagnosis provides a useful entry point into broader issues of performance management. The diagnosis also emphasizes technical capacity, because without reliable data, good analytical instruments and solid technical capacity, M&E systems will not get results, even where institutions and organizational capacity might favor good M&E practice.

The broad questions asked by analysts during any M&E diagnosis might be similar, but because country situations differ widely due to varying levels of development and local situations, the topics addressed, diagnosis presentation, and conclusions and recommendations may look very different. Mexico, for instance, has established important elements of a centralized M&E system and now one of the main goals is to ensure that the nascent performance-informed budgeting system develops properly (Castro et al. 2009). In Uganda and South Africa, the issue is more one of creating greater coherence among separate systems (Hauge 2003; Engela and Ajam 2010). In the United Kingdom, evaluation is underdeveloped compared to monitoring, and the two are relatively unconnected (Talbot 2010).

There is no single blueprint for preparing an M&E diagnosis: content and presentation depend on the specific context. This note seeks to illustrate the range of contexts by discussing a number of issues that help explain the contextual differences, drawing on, where possible, various country studies that have appeared in recent years. These studies cover six Organisation for Economic Co-operation and Development (OECD) countries—mostly studies that provide good practice lessons—and two African and four Latin American countries—a mix of diagnostic and “lessons-of-experience” studies.

## A Checklist of Topics for a Diagnosis

Table 1 presents a checklist of topics to be considered in preparing an M&E diagnosis. This checklist is broadly based on a number of M&E country diagnoses and several diagnostic how-to guides that have appeared in the last decade. These diagnoses by no means follow the same outline, but they do tend to cover a similar sets of topics. The checklist is designed for a diagnosis at the national level, although adapting a diagnosis to a particular agency or subnational government would not require any fundamental alteration of the checklist. This note

**Table 1. Checklist of Topics for Preparing M&E Diagnosis**

Block A. The national environment for M&E	
Topic A1	National policy and institutional framework: How are policies made? What role do donors play? Is political power wielded in the public interest? Do policies create a demand for M&E? How decentralized is the country? How has the relevant policy environment evolved over time?
Block B. M&E systems: If the study covers multiple M&E systems (governmentwide, cross-sector, or within ministries, agencies, audit institutions, or donor projects), then Block B applies separately to each system	
Topic B1	Historical development: How and why did the system develop? Who championed it and who opposed it? What kind of implementation strategy was adopted?
Topic B2	Objectives (announced, implicit, or revealed): These can include budget support, support to policy making and/or program improvement, or accountability.
Topic B3	Processes, tools, and products: What is produced (indicators and evaluations by type and numbers, and others)? What are the selection criteria? What is the production cycle? How is the information used (dissemination, reward, sanction, or correction)? How is the quality of the information controlled? What are the tools used to collect, manage, and analyze information and are they appropriate?
Topic B4	Relationship with other systems: How are systems interconnected, if at all? Monitoring with evaluation? M&E with the budget? Ministry or subnational monitoring systems with national systems? Monitoring with information systems? M&E with quality management systems?
Topic B5	Institutional architecture: How do the system's components fit together? How is cooperation (exchange of information, willingness to act on results) achieved within the system? How centralized is the system?
Topic B6	Organizational characteristics of public agencies that are part of the system: What is the historical reform/policy change process? The tasks of the agency? Its resources (budget, incentives, expertise, training, donor support, and so forth)? Its sources of authority (legal framework, stakeholder roles)? The obstacles it faces (information, coordination problems)?
Topic B7	Results: What are the quality, credibility, and access bility of the M&E products? What is the impact of these products? Where there are multiple objectives, are there multiple impacts?
Block C. Findings	
Topic C1	Conclusions and recommendations: What is working and not working, and why? What reforms are underway? How can things be improved?

Source: Author's compilation.

is not intended to be a how-to guide, but rather a tool to help deepen understanding of the issues.

## Diagnostic Issues

### Topic A1. National policy and institutional framework

**Placing M&E within the public policy cycle.** Perhaps the single most important point to be made about an M&E system is that it must be embedded within a broader policy framework that is conducive to M&E. This policy framework must in some sense be performance oriented—such as by setting performance expectations, monitoring progress, measuring results, and appraising, rewarding, or correcting performance. M&E cannot be considered independently of a performance-oriented policy cycle, where future actions are evidence based and strategically planned, implementation is monitored, and results are evaluated and then fed back into the planning process. Lack of clear objectives and unpredictability in future resources undermine the possibility of such a policy cycle and hence the demand for effective M&E. In short, if elements of performance management are absent, it is unlikely that M&E can flourish. Thus any M&E diagnosis has to investigate those broader elements of the policy framework that drive, or could drive, investment in M&E.

Recent M&E diagnoses make it clear that performance orientation is central to the successful development of M&E systems. Castro et al. (2009) describe how Mexico's evaluation agency, Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL), sits at the apex of a comprehensive (if not yet fully institutionalized) M&E system that is part of a performance budgeting system. Somewhat similarly, a centralized evaluation system in Chile is integrated within a broader performance-management regime (World Bank 2006). By contrast, Matsuda, Shepherd, and Wenceslau (2006) argue that the poor performance of Brazil's centralized M&E system has to be understood within the context of the country's planning and budgeting problems. A comparative diagnosis of M&E systems in Australia, Canada, the United Kingdom, and the United States (Lahey 2005) argues that, in all four countries, public sector reforms involved the devolution of managerial discretion to individual government departments and agencies, as well as a desire to link evaluation and performance information

to budget decision making and high-level policy debate. More recent studies for Canada (Lahey 2010), the United Kingdom (Talbot 2010), and Australia (Mackay 2011) tell a similar story of M&E systems embedded within performance-management regimes. In Australia, a new government elected in 1996 that favored less government partially abandoned an effective evaluation system that had been built up over the previous decade.

In Uganda, the performance driver behind M&E reforms was the poverty reduction agenda (Hauge 2003). In Ireland and Spain, evaluation requirements that accompanied structural aid from the European Union were the principal driver of M&E, but not the only driver (Boyle 2005; Feinstein and Zapico-Goñi 2010). Foreign donors have also been important as initiators of evaluation activities, but these efforts concentrated mostly on donors' own projects. In Uganda, there has been considerable donor investment in M&E; as a result, Uganda's problem is not the quantity of M&E, but the poor quality, poor coordination, and substantial duplication.

**M&E in a federal system.** M&E diagnoses have mostly covered central governments. But, particularly for federal countries, the role of M&E in linking central and subnational governments is also an important, if underdiscussed, issue. When activities are devolved to sub-national governments, it is usual for federal authorities to retain substantial overall responsibility for policies and results. But this devolution means that the federal authorities lose management authority and have greater difficulty in getting information. In other words, what were formerly outputs of the federal authorities (that is, results over which they had substantial control) become outcomes (that is, results to which they contribute, but over which they do not have substantial control). Thus the federal authorities need their own M&E systems to give them managerial control over the discretionary funds they transfer to lower-level governments and fiduciary oversight over subnational governments' use of their own funds. Additionally, federal authorities can encourage subnational governments to carry out their own M&E and thereby help improve the quality of local public spending (and national learning).

The country studies do not treat the subnational theme at great length. Talbot (2010) mentions the use of a special purpose agency in the United Kingdom, the Audit Commission, to

help oversee local governments. Engela and Ajam (2010) and World Bank (2011) note the importance of M&E in the developing federal systems of South Africa and Brazil.

### **Topic B1. Historical development**

History matters, for at least two reasons. First, current events and decisions are “path dependent”—they are conditioned by past events and decisions. Thus, national idiosyncrasies explain why M&E systems can look so different from country to country. Second, governments do not institute perfect new management systems overnight; instead, they learn by experience and adjust systems over time.

National M&E diagnoses put the development of M&E systems, as well as the broader policy framework in which they exist, within a historical narrative. One of the main history-related themes comes from studies in the United Kingdom and Canada (Talbot 2010; Lahey 2010). Even in such “mature” countries with supposedly advanced forms of performance-oriented management—and in Canada’s case, an exemplary, well-integrated M&E system—there is little sense that these countries are approaching a final, ideal M&E system. Instead, the story is of constant adjustments to improve systems. Similarly, a learning process has been at work in the development of M&E systems in Chile, Colombia, Ireland, Mexico, and South Africa. But history can also tell a more cyclical story: in some advanced countries, evaluation lost traction in the 1990s because of budget cuts (Lahey 2005); Australia abandoned its evaluation system in 1997 because of changes in political ideology (Mackay 2011); the United Kingdom apparently abandoned its successful system of Public Service Agreements when a new government took office in 2010.

### **Topic B2. Objectives**

M&E instruments can address a multiplicity of different and often incompatible objectives. It is important to understand these objectives—whether they are announced, implied, or revealed—because different objectives can require different types of measurement and analyses. Table 2 lists five sets of objectives: solving problems in program implementation; accountability within the government; provision of public information (for accountability, legitimacy, or public choice); improving program design; and prioritizing (and

possibly coordinating) across programs. The table also provides examples of national systems that have been functioning reasonably well.

In many cases, it is difficult to use the same system to respond to different objectives. A comparative study of evaluation systems in Latin America (Ospina, Cunill Grau, and Zaltsman 2004) differentiates between systems for designing better programs and policies (“planning models” in Colombia and Costa Rica) and systems for helping expenditure prioritization (“budget models” in Chile and Uruguay). Canada’s M&E system is an example of a system that can effectively answer to more than one objective (Lahey 2010). Brazil’s national M&E system failed in part because it took on too many objectives (Matsuda, Shepherd, and Wenceslau 2006). Engela and Ajam (2010) discuss differences in objectives—and resulting tensions—among different government agencies in South Africa, contrasting a more linear program logic at the Treasury (which is concerned with budget prioritization) with a more complex view of how programs work in the sectoral ministries.

### **Topic B3. Processes, tools, and products**

This topic covers the technical design of the M&E tools to collect, manage, and analyze information and the production system in which they are used, as well as the scale of the activity, for instance the number of reports produced annually or number of indicators. The topic covers:

- *Monitoring:* Criteria for selecting indicators; appropriateness, periodicity, timeliness, reliability, and quality control of indicators; and methods of collecting, reporting, and disseminating data.
- *Evaluation:* The evaluation cycle; criteria for selecting subjects for evaluation; evaluation techniques (impact, rapid, process evaluations, and so forth); training and use of evaluators; and dissemination methods.

Evaluation encompasses a broad range of tools, from the informal to the formal, and from the impressionistic to the rigorous. There is a range of formal evaluative tools (cost-benefit, cost-effectiveness, impact evaluation) and less formal tools (rapid evaluation, user surveys, and so on). Different tools respond to different needs and the availability of information. Needs differ in at least three systematic ways:

- *Different types of problems:* Problems differ from one policy area to another. In infra-

**Table 2. M&E Objectives: A Taxonomy and National Examples**

<p><b>1. Solving implementation problems:</b></p> <p>Monitoring program execution to detect and correct implementation problems. This is a managerial learning function and is usually incorporated within a management information system (MIS).</p>
<ul style="list-style-type: none"><li>• Brazil: Monitoring the Growth Acceleration Program.</li><li>• United Kingdom: Prime Minister's Delivery Unit.</li></ul>
<p><b>2. Intragovernment accountability:</b></p> <p>M&amp;E of program execution both within government and between levels of government to ensure that public agents are fulfilling their responsibilities. This is a managerial (internal) accountability function, also usually incorporated within an MIS.</p>
<ul style="list-style-type: none"><li>• Brazil: Compliance and process audits, and providing public information.</li><li>• Canada: Articulated M&amp;E information used for internal accountability and management and for external accountability, reporting to the center of government and to Parliament.</li><li>• South Africa: Program of Action (MIS).</li><li>• United Kingdom: Public Service Agreements, also with elements of public accountability.</li></ul>
<p><b>3. Providing public information:</b></p> <p>Providing information to the legislature and the public to provide: (i) external accountability; (ii) information that legitimizes a public action; or (iii) information that facilitates public choice and voice.</p>
<ul style="list-style-type: none"><li>• Brazil: Performance evaluation of government programs by the Federal Court of Audit.</li><li>• Brazil: Program evaluation by Ministry of Social Development to legitimize new programs, and also for learning related to program and policy design.</li><li>• Colombia: System of Programming and Management by Objectives and Results (SIGOB), a monitoring system for legitimization and also to inform the president of ministry performance.</li><li>• South Africa: Development indicators for accountability.</li><li>• United Kingdom: Education and health "scorecards" for consumer choice.</li></ul>
<p><b>4. Improving program design:</b></p> <p>M&amp;E allows governments to determine the efficiency and effectiveness of individual programs in order to inform decisions to extend, improve, or eliminate them.</p>
<ul style="list-style-type: none"><li>• Chile: Government program evaluations by the Budget Office, and also for budget decision making.</li><li>• Colombia: Evaluations of the National System for Evaluation of Public Sector Performance (SINERGIA), and also for policy prioritization.</li><li>• Ireland: Evaluation for projects of European Union structural funds, and for accountability.</li><li>• Mexico: Program evaluations of CONEVAL.</li><li>• Spain: Evaluations by the Spanish Evaluation Agency (AEVAL).</li></ul>
<p><b>5. Coordinating and prioritizing programs, particularly in budgeting:</b></p> <p>M&amp;E allows governments to determine efficiency and effectiveness across programs. This helps central government to coordinate among programs and/or prioritize cabinet, planning, and budget decisions.</p>
<ul style="list-style-type: none"><li>• Australia: Evaluation system of 1987–97 helped in policy and budget prioritization, and also was important for improving program performance and providing internal accountability.</li><li>• Canada: Strategic reviews, for budget.</li><li>• Chile: Comprehensive spending reviews, for budget.</li><li>• Ireland: Expenditure Review Initiative, for budget.</li><li>• United States: Program Assessment Rating Tool (PART), for budget.</li></ul>

Source: Author compilation.

structure activities, for instance, outputs and outcomes are relatively well understood and measurable, but start-up costs (including choice of technique) can strongly affect

efficiency. In social areas, start-up costs are more modest, but desired outcomes, though measurable, are harder to attain since human behavior is more difficult to engineer than

bridges. This helps explain why impact evaluation is the gold standard of policy analysis in the social sectors and cost-benefit analysis the gold standard in infrastructure sectors.

- *Different costs of analysis:* The technically best forms of analysis have costs as well as benefits—they require more information (hence cost more) and they may produce results that come too late.
- *Different objectives:* M&E activities cover a range of objectives, as discussed above.

The diagnostic studies under review differ in the amount of space they provide on production routines. Studies conducted in Chile (World Bank 2006) and Colombia (World Bank, Independent Evaluation Group 2007) are examples of diagnoses that provide a reasonable amount of detail on evaluation tools and methodologies.

#### **Topic B4. M&E system's relationship with other systems**

The idea of M&E as a self-contained system (albeit within larger systems of performance management) is a slippery one. First, monitoring and evaluation can be substantially separated from each other, even though the line between them can be blurred. Monitoring typically centers on performance—through a measured comparison (usually annual) of actual and anticipated results—and tends to target the deployment of inputs, effort, and procedures to improve efficiency. Evaluation, which is primarily *ex post*, typically compares a program's or policy's objectives to actual results, and any reform implications tend to be for program design.

Talbot (2010) differentiates the United Kingdom's monitoring regime from its evaluation regime. The former has attributes of a fairly tight system driven by an efficiency-oriented philosophy and based on organizations rather than programs. Evaluation activities are quite separate: evaluations are the product of a different professional community driven by somewhat different considerations (substantive policy more than efficiency), and evaluation tends to be more program based, less developed, and less systemic than monitoring. In Colombia and Chile, monitoring and evaluation are also largely unintegrated, even though they fall under one "system" and one organization. Canada, seemingly the exception, has been able to substantially integrate the two functions.

Second, it can be difficult to separate a specific set of activities called M&E from other policy-oriented information collecting and analytical activities systematically undertaken within the public administration. M&E, compliance and process auditing, quality management, policy analysis, and information management represent separate communities of practice that substantially overlap, but are often poorly connected to each other.

M&E crucially depends on good information and good analysis. Governments' ability to harness the power of modern information and communications technologies (ICTs) is very important in determining the sophistication of their M&E efforts. For instance, crossing physical and financial data allows program implementation to be tracked with greater accuracy and irregularities to be more easily detected. A country's public information infrastructure is influenced by its resources (software and systems development, data management, and statistical skills) and by public institutions, particularly statistics institutes, but also central ministries and line ministries. Knowledge management also requires analytical skills typically found in government think tanks, research institutes and nongovernmental organizations (NGOs), including universities, think tanks, and consultancy firms.

Few M&E diagnoses have discussed national knowledge management infrastructures. A diagnosis for Brazil (World Bank 2011) provides some examples of how innovative uses of ICT have improved monitoring capabilities in some ministries. A diagnosis for South Africa (Engela and Ajam 2010) also discusses information and data constraints. As government has become larger and more complex in OECD countries, specialized government and nongovernment agencies have come to play an increasingly important role in the analysis of public policies and programs. In effect, policy advice is becoming a more competitive business (World Bank 2010).

#### **Topic B5. Institutional architecture**

Key questions concern identification of system components (the participants and their roles) and mechanisms for securing cooperation between the components. This has been a central topic for most diagnoses. The main themes are the degree of system centralization and the incentives that induce players to part with information or to act on the

results of M&E. There are as many architectural variants as there are systems.

To start, diagnoses describe a variety of systems ranging from loose to strict. The M&E “system” characterized by Engela and Ajam (2010) for South Africa is in reality a governmentwide policy to coordinate many existing and disparate M&E initiatives. At the other extreme, World Bank (2006) describe a centralized and tightly controlled evaluation system in Chile. Canada, Colombia, Mexico, and Spain also have tightly defined national M&E systems in various stages of construction. The United Kingdom has a narrowly defined, well-established system for performance monitoring, not evaluation.

The following systems have different modes of coordination and varying degrees of centralization.

- In Chile and Mexico, a central agency carries out the evaluations. Mexico’s CONEVAL also helps set the rules for M&E practices in the line secretariats.
- In Canada (somewhat like Australia 1987–97), the central government sets the rules and provides advice, while the line departments are required to carry out the evaluations, with some degree of liberty in organizing the evaluation function. South Africa aspires to a somewhat similar arrangement.
- A peculiarity of the United Kingdom is that, in many sectors, it uses a number of specialized inspectorates, standing at arm’s length from their sponsoring departments, to carry out monitoring.
- Finally, there is a growing involvement of supreme audit institutions (SAIs), usually attached to legislatures, in performance evaluation. Brazil and the United Kingdom are good examples: their SAIs have the authority to collect information and require program changes, and they have become an important part of a broader national M&E system.

#### **Topic B6. Characteristics of organizations**

Public agencies that are part of a particular M&E system—whether a central agency that runs the system or a line agency that participates in the system—have particular organizational characteristics that determine their own incentives to supply information and to use M&E results. Different agencies in the same government can have quite different capabilities and quite different

incentives. These differences are shaped by the history of the agency, what it does, the resources it can command, and the external factors and pressures that influence its behavior (legal framework, stakeholders, and so on). This topic overlaps with the topics on the national policy and institutional framework and on institutional architecture because all three address different aspects of incentives (box 1).

The organizational framework is generally not systematically treated in the diagnoses and studies under review, but there is a substantial, if episodic, treatment of related themes.

- Studies on Brazil (World Bank 2011) and South Africa (Engela and Ajam 2010) discuss features that explain differences among ministries in their cooperation with M&E systems. Talbot (2010) notes the variability of departmental M&E capacities in the United Kingdom.
- Castro et al. (2009) describe the organizational characteristics of political independence and technical capacity that have made CONEVAL a credible and influential evaluator of social programs in Mexico.
- Lahey (2005) emphasizes the importance of formal, legislated policies in reinforcing the use of evaluation in individual government departments and agencies in Australia, Canada, the United Kingdom, and the United States.
- The obstacles to coordination are discussed in several diagnoses. Cunill Grau (2010) suggests that poor information flow and lack of cooperation among public agencies has forced a centralized approach to M&E systems in Latin America, in contrast to the more decentralized approach in OECD countries. Engela and Ajam (2010) describe the culture of information hoarding in South Africa, as well as other differences among public agencies that impede cooperation.

#### **Topic B7. Results**

Clearly, there is not much point to an M&E system, however well-designed and functioning, if it does not get results in terms of goals fulfilled, programs and policies improved, budgets rationalized, or a public better informed. Unfortunately, it is quite difficult to measure such improvements and attribute them to specific causes.

Generally speaking, evidence on results is weak or missing from existing M&E diagnoses.

### Box 1. Looking at Incentives in an M&E Diagnosis

Incentives are the factors that motivate a particular course of action by individuals or organizations. At the macrolevel, incentives are woven into the institutions that determine governance—do the formal and informal rules impel public officials to act in the public interest or in private interests? At the microlevel, incentives take the form of sticks (sanctions for poor performance) or carrots (for instance material rewards) that affect the effort that individuals and organizations make, at the margin, to perform better.

Incentives figure prominently in three of the checklist's topics:

- The national policy and institutional framework: The public interest nature of national policies, the clarity and the binding nature of national policy directives, and the nature of national management rules (especially budget rules) determine the core incentives for countries to pursue M&E. It is thus vital for M&E diagnoses to identify these characteristics. But since they are driven largely by politics, it is often difficult to propose solutions in the context of an M&E diagnosis.
- The M&E system—the institutional architecture: A specific architecture helps determine a narrower set of incentives for agencies to cooperate—exchange information or take corrective actions based on M&E results—within a circumscribed system. Different systems make different trade-offs between the use of hierarchical authority (in centralized systems) and the appeal to the self-interest of agencies in pursuing M&E (in decentralized systems).
- M&E system—the organizational characteristics: At the agency/sectoral level, similar incentive issues arise as for the national policy and institutional framework. At the agency level, politics can also be a driving force shaping incentives, but so too are specific features of the sector, as well the sticks and carrots of microincentives. Agency-level incentives may thus be more effective than national incentives.

Source: Author's compilation.

The study on Chile (World Bank 2006) was able, in broad terms, to trace the effects of evaluations on budget decisions. Some studies cited individual examples of evaluations that were influential. Other studies made judgments about the quality of evaluation reports as an implicit proxy for their expected impact: World Bank 2006 had experts judge the quality of a sample of program evaluations in Chile, and Matsuda, Shepherd, and Wenceslau (2006) compared the quality of program evaluations carried out independently by the SAI and by ministries. For the most part, studies had to be content with author inferences based on, at best, knowledge of the system and interviews and focus groups with participants.

#### Topic C1. Conclusions and recommendations

The conclusions to an M&E diagnosis should present the overall strengths and weaknesses of M&E systems and activities, as well as indicate current or planned reforms. Because existing diagnoses provide different models and elucidate the topics discussed here, they can help frame these conclusions.

Existing diagnoses can also provide a useful guide to possible reform options. On the other

hand, because of the large number of national variables—different historical starting points and national institutional environments, the range of objectives and architectures, and the differences between public agencies—no one model will fit all reform situations. Indeed, realistic recommendations must start from local realities. Public sector reform is subject to bounded rationality—there is not enough information to know how to get it all right the first time—and path dependence. Thus, one has to build on what is in place and what, under local circumstances, is more likely to work. Guidelines, taxonomies, and checklists can provide only limited help.

#### Organizing the Diagnosis

In addition to the valuable information contained in several existing diagnoses, several guides provide useful advice on diagnostic issues, content, and organization (box 2).

A few final points on organizing an M&E diagnosis:

- An initial decision has to be made about whether an M&E diagnosis will be useful. In countries where there is no drive for improved public sector performance, there will be no



## Box 2. Published Guides for M&E Country Diagnosis and Capacity Building

- Boyle, Richard, and D. Lemaire, eds. 1999. *Building Effective Evaluation Capacity: Lessons from Practice*. New Brunswick, NJ: Transaction Publishers.
- Cunill Grau, Nuria, and Sonia Ospina Bozzi, eds. 2003. "Guide for Case Studies" (in Spanish). In *Evaluación de resultados para una gestión pública moderna y democrática: experiencias latinoamericanas*, appendix. Caracas: CLAD (Centro Latinoamericano de Administración para el Desarrollo).
- Görgens, Marelize, and Jody Zall Kusek. 2009. *Making Monitoring and Evaluation Systems Work: A Capacity Development Toolkit*. World Bank.
- Kusek, Jody Zall, and Ray C. Rist. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System*. Washington, DC: World Bank.
- Mackay, Keith. 1998. "Evaluation Capacity Development: A Diagnostic Guide and Action Framework." World Bank, Evaluation Capacity Development Working Paper Series No. 6, Washington, DC.
- ———. 2007. "The Importance of Country Diagnosis." In *How to Build M&E Systems to Support Better Government*, chapter 12. Washington, DC: World Bank, Independent Evaluation Group.
- World Bank, Independent Evaluation Group. 2007. "Annex: Terms of Reference for an In-Depth Diagnosis of SINERGIA." In "A Diagnosis of Colombia's National M&E System, SINERGIA." ECD Working Paper Series 17, Washington, DC.

Source: Compiled by author.

motivation for M&E. In such cases, a more basic examination of institutional conditions for public sector reform is required.

- As with any diagnosis of public sector issues, it makes sense to involve the government as much as possible in the design and implementation of a study, both to ensure better access to information and to make it more likely that the study's findings will be taken seriously (topic B5).
- The scope of the study will need to be defined. Will the study concentrate on a particular system, or will it look at M&E activities more generally? Where local governments are important partners in implementing national priorities, it will be important for M&E diag-

noses to link central and subnational government M&E activities.

- How large will the study be? Mackay 2007 makes the point that a feasible diagnosis can range from a small team spending one or two weeks in the field to a large team undertaking—in effect, an impact evaluation. No single diagnosis is likely to cover, in any detail at least, all the topics detailed in this note. Thus, the smaller the study, the more choices that have to be made.
- The better prepared an investigative team is by the time it sets foot in-country, the better the end product is likely to be.
- Finding out about an M&E system's results is likely to be the hardest challenge in the field. Interviews and focus groups must be designed to listen to multiple perspectives—producers and users of M&E, participants and observers, and supporters and critics.
- There are different ways to tell a story, and the particular context must dictate the particular way chosen for any diagnosis. A historical narrative has proven important in many diagnoses. On the other hand, where comparability is required with other studies, a more instrumental approach might make more sense.

## Acknowledgments

For their comments, the author thanks Nick Manning (Adviser, PRMPS), Philipp Krause (Consultant, PRMPR), Keith Mackay (Consultant, PRMPR), and Gladys Lopez-Acevedo (Senior Economist, PRMPR). The views expressed in this note are those of the author.

## About the Author

*Geoffrey Shepherd* is an economist with a particular interest in institutions and public sector reform. He received his D.Phil. from the *University of Sussex*, where he also worked from 1978 to 1986 (with appointments in the *Sussex European Research Centre*, then in the *Science Policy Research Unit*). He worked for the *World Bank* in the 1970s, then again from 1986 to 2001, on private sector development and trade issues in many parts of the world and, later, on public sector reform and institutional development issues in Latin America. He now consults for the *World Bank*.

## References

- Boyle, Richard. 2005. "Evaluation Capacity Development in the Republic of Ireland." Independent Evaluation Group, ECD Working Paper Series 14, World Bank, Washington, DC.
- Castro, Manuel Fernando, Gladys Lopez-Acevedo, Gita Beker Busjeet, and Ximena Fernandez Ordonez. 2009. "Mexico's M&E System: Scaling Up from the Sectoral to the National Level." Independent Evaluation Group, ECD Working Paper Series 20, World Bank, Washington, DC.
- Cunill Grau, Nuria. 2010. "Latin American Monitoring and Evaluation Systems and Their Governance." In *Challenges in Monitoring and Evaluation: An Opportunity to Institutionalize M&E Systems*, ed. Gladys Lopez-Acevedo, Katia Rivera, Lycia Lima, and Helena Hwang. Fifth Conference of the Latin America and the Caribbean M&E Network, World Bank.
- Engela, Ronette, and Tania Ajam. 2010. "Implementing a Governmentwide Monitoring and Evaluation System in South Africa." Independent Evaluation Group, ECD Working Paper Series 21, World Bank, Washington, DC.
- Feinstein, Osvaldo, and Eduardo Zapico-Goñi. 2010. "Evaluation of Government Performance and Public Policies in Spain." Independent Evaluation Group, ECD Working Paper Series 22, World Bank, Washington, DC.
- Hauge, Arild, O. 2003. "The Development of Monitoring and Evaluation Capacities to Improve Government Performance in Uganda." Operations Evaluation Department, ECD Working Paper Series 10, World Bank, Washington, DC.
- Lahey, Robert. 2005. "A Comparative Analysis of Monitoring and Evaluation in Four Selected Countries: Canada, United States, Australia, and United Kingdom." August, prepared for the World Bank, Washington, DC.
- . 2010. "The Canadian M&E System: Lessons Learned from 30 Years of Development." Independent Evaluation Group, ECD Working Paper Series 23, World Bank, Washington, DC.
- Mackay, Keith. 2007. *How to Build M&E Systems to Support Better Government*. Washington, DC: World Bank, Independent Evaluation Group.
- . 2011. "The Australian Government's Performance Framework." Independent Evaluation Group, ECD Working Paper Series 25, World Bank, Washington, DC.
- Matsuda, Yasuhiko, Geoffrey Shepherd, and Juliana Wenceslau. 2006. "Management and Evaluation within the Plano Plurianual: Institutionalization without Impact?" LCSPS, unpublished, World Bank, Washington, DC.
- Ospina, Sonia, Nuria Cunill Grau, and Ariel Zaltsman. 2004. "Performance Evaluation, Public Management Improvement and Democratic Accountability: Some Lessons from Latin America." *Public Management Review* 6(2): 230–51.
- Talbot, Colin. 2010. "Performance in Government: The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." IEG Working Paper Series 24, World Bank, Washington, DC.
- World Bank. 2006. *Chile: Study of Evaluation Program, Impact Evaluations and Evaluations of Government Programs*. Report No.34589-CL, World Bank, Washington, DC.
- . 2010. *The Practice of Policy-Making in the OECD: Ideas for Latin America*. Washington, DC: Public Sector and Governance Unit, Poverty Reduction and Economic Management Department, Latin America and the Caribbean, World Bank.
- . 2011. "Policy Note: Monitoring and Evaluation in Brazil's Federal Government." LCSP, World Bank, Washington, DC.
- World Bank, Independent Evaluation Group. 2007. "A Diagnosis of Colombia's National M&E System, SINERGIA." ECD Working Paper Series 17. World Bank, Washington, DC.