

INDEPENDENT EVALUATION GROUP
& THE THEMATIC GROUP FOR POVERTY ANALYSIS, MONITORING AND IMPACT
EVALUATION
THE WORLD BANK



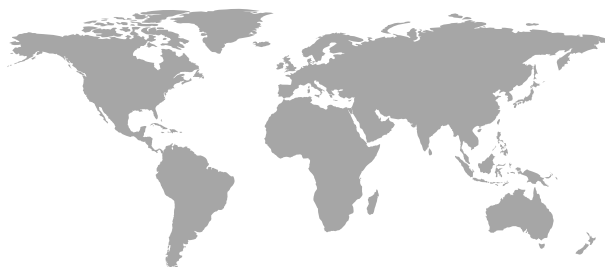
EVALUATION CAPACITY DEVELOPMENT

ECD WORKING PAPER SERIES - NO. 15: JANUARY 2006

Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management

Keith Mackay





Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management

ECD Working Paper Series ♦ 15

Keith Mackay

A growing number of countries are pursuing a results orientation by building or strengthening their government monitoring and evaluation (M&E) systems. This paper provides an overview of the increasingly rich body of experience with these efforts. The dimensions of a 'successful' government M&E system are considered, using Chile as an example. Success factors and mistakes to avoid are examined. Finally, the special case of Africa is outlined.



Copyright 2006
Independent Evaluation Group
Knowledge Programs & Evaluation Capacity Development
Email: eline@worldbank.org
Telephone: 202-473-4497
Facsimile: 202-522-3125

Evaluation Capacity Development (ECD) helps build sound governance in countries—improving transparency, and building a performance culture within governments to support better management and policymaking, and to strengthen accountability relationships—through support for the creation or strengthening of national/sectoral monitoring and evaluation systems. A related area of focus is civil society, which can play a catalytic role through provision of assessments of government performance. IEG aims to identify and help develop good-practice approaches in countries, and to share the growing body of experience with such work.

The IEG Working Paper series disseminates the findings of work in progress to encourage the exchange of ideas about enhancing development effectiveness through evaluation. An objective of the series is to get the findings out quickly, even if the presentations are somewhat informal.

The findings, interpretations, opinions, and conclusions expressed in this paper are entirely those of the author. They do not necessarily represent the views of the Independent Evaluation Group or any other unit of the World Bank, its Executive Directors, or the countries they represent.

CONTENTS

Foreword.....	i
1. INTRODUCTION.....	1
2. WHAT DOES ‘SUCCESS’ LOOK LIKE? — THE CASE OF CHILE.....	2
3. COUNTRY CAPACITIES — FOR WHAT.....	4
4. LESSONS FROM EXPERIENCE — SUCCESS FACTORS FOR BUILDING COUNTRY M&E SYSTEMS.....	5
5. KEY TRENDS INFLUENCING COUNTRY REALITIES: A DONOR PERSPECTIVE.....	11
6. THE SPECIAL CASE OF AFRICA.....	13
7. CONCLUSIONS, AND CHALLENGES FOR THE FUTURE.....	15
REFERENCES.....	16

FOREWORD

The Independent Evaluation Group (IEG) — formerly known as the Operations Evaluation Department (OED) — of the World Bank has a long-standing program of support to strengthen monitoring and evaluation (M&E) systems and capacities in developing countries, as an important part of sound governance. As part of this support, IEG has prepared a collection of resource material including case studies of countries which can be viewed as representing good-practice or promising-practice. This resource material is available electronically at:

<http://www.worldbank.org/ieg/ecd/>

This paper also comprises part of the ongoing efforts of the World Bank's thematic group for poverty analysis, monitoring and impact evaluation to provide operationally relevant guidance on strengthening poverty monitoring systems and improving the quality and use of impact evaluations. This resource material is available at: <http://www.worldbank.org/povertymonitoring> and <http://www.worldbank.org/impacetevaluation>

The purpose of this paper is to draw together the extensive and growing body of experience with the institutionalization of government M&E systems, particularly those in developing countries. Using Chile as an example, the paper outlines what ‘success’ looks like; but the point is made that it is dangerous to look for best-practice country examples. Each country is unique, in terms of its starting point and also in terms of the destination to which it aspires — much depends on the particular uses of M&E information for which the system is being designed. That said, a number of lessons and success factors can be identified, as well as a number of mistakes to avoid.

The paper also considers a number of international trends and influences on country efforts to institutionalize M&E, such as the demonstration effect of rich countries, pressures on governments to provide more services in a climate of fiscal constraints, and the greater emphasis of international donors on the achievement of measurable results. Finally, the paper discusses the special case of Africa and the types of M&E issues which they face.

The views expressed in this paper are solely those of the author, and do not necessarily represent the views of the World Bank.

Klaus Tilmes
Manager
Knowledge Programs & Evaluation Capacity Development

1. INTRODUCTION

There is a growing appreciation within the development community that an important aspect of public sector management is the existence of a results or performance orientation in government. Such an orientation — in effect, an ‘evaluation culture’ — is considered to be one avenue for improving the performance of a government, in terms of the quality, quantity and targeting of the goods and services which the state produces. In support of this objective, a number of countries are working to ensure a results orientation through building or strengthening their monitoring and evaluation (M&E) systems.

The focus here is on governments, although civil society organizations — such as national evaluation societies, universities, and non-government organizations (NGOs) — also have a role to play. International donors are key stakeholders in country efforts to institutionalize evaluation; these donors support such efforts partly for altruistic purposes and partly to support their own, increasing emphasis on measuring and managing for results.

There exists a growing literature on the topic of country efforts to strengthen M&E capacities and systems. A considerable part of this literature, written by evaluation specialists, has a strong advocacy flavor: that M&E and M&E systems are a ‘good thing’ and have intrinsic merit. Unsurprisingly, this kind of argument is a hard sell to skeptical or over-stressed governments in the developing world. And so, this paper starts with a desirable end-point of what ‘success’ can look like. This is followed by a consideration of various success factors which have become evident in recent years — ‘success’ in terms of the institutionalization of M&E — and in the context of key trends which are influencing country realities. The special case of Africa, as both the neediest continent and the weakest in terms of evaluation or other sophisticated skills needed for good governance, is then considered. African experience is relevant to countries in other regions, such as poor countries which are preparing poverty reduction strategies. It also has lessons for how to build M&E capacities incrementally, especially when there is the possibility of intensive donor assistance. The paper concludes by outlining a number of challenges and some options for influencing the future.

2. WHAT DOES ‘SUCCESS’ LOOK LIKE? — THE CASE OF CHILE

For those of us whose career revolves around helping countries build their monitoring and evaluation (M&E) systems, Chile can seem like the promised land. It has a well-performing M&E system and one which is home-grown. Thus while the government has drawn on the lessons from international experience, including via study tours to learn from other countries, it developed its system for its own purposes and not to satisfy donor conditionalities for M&E. The system is run by the Finance ministry, a highly capable, well-respected (and feared) organization. This ministry has developed the M&E system progressively, over the past decade in particular, and with a focus on the annual budget cycle and its information needs (Box 1).

Box 1: Chile’s Whole-of-Government M&E System

Chile’s M&E system, which is managed by the powerful Finance ministry, has developed progressively over time, partly in response to fiscal pressures, partly in response to the changing landscape of public sector reforms, and partly in an opportunistic manner (see Guzman, 2003, World Bank, 2005, and Zaltsman, 2006). Major milestones include the following:

- Ex ante cost-benefit analysis required for all government projects (1974).
- Performance indicators collected for all government programs (1994). Regular information is collected on about 1600 indicators. These are used in the formal reports prepared for the Congress, and to provide key data for the various types of evaluation which are conducted.
- *Comprehensive Spending Reports* (1996).
- *Government Program Evaluations* (1996). These are in the nature of program reviews, and about 160 have been conducted so far. They comprise clarification and agreement of detailed program objectives, preparation of a logframe analysis, desk review, and analysis of existing data. Their average cost is some US\$11,000, and they usually take from 4 to 6 months to complete.
- Rigorous impact evaluations (2001). These entail primary data collection, and often the use of control groups and difference-in-differences. Fourteen have been completed to date, at an average cost of \$88,000 and taking up to 18 months to finish. About 60% of the government’s budget has been evaluated by means of either these impact evaluations or the *government program evaluations*.
- *Comprehensive Spending Reviews* (2002). These review all programs within a particular functional area, and look at issues of inefficiency and duplication of programs. They include desk reviews, and have cost \$48,000 on average. Five have been completed so far.

What is impressive about the Chilean case is not the rich list of types of monitoring and evaluation which the government undertakes; this architecture does not make it a success story. Rather, it is the generally high quality of the M&E work which is undertaken,¹ and in particular it is the intensive utilization of the monitoring information and evaluation findings which the M&E system produces. A recently-completed World Bank evaluation of Chile's M&E system found that the government's evaluations (which are outsourced to consultants and to academia) are used by the Finance ministry for its resource allocation decisions within the budget process, and to impose management and efficiency improvements on sector ministries in the programs for which they are responsible (World Bank, 2005). The Finance ministry also ensures that this M&E information is reported fully to the Congress, which in turn is highly appreciative of it.

The important role of the Finance ministry and its powerful position in the government — much more prominent than finance ministries in most countries — augur well for the sustainability of the Chilean government's M&E system. However, an unfortunate side-effect of the forceful use of M&E information by the Finance ministry has been the low level of ownership and utilization of it by sector ministries and their agencies. There exists an unexploited opportunity for them to use this information for their own strategic planning, policy development, and ongoing management and control.

¹ Some weaknesses are evident. For example, the *government program evaluations* and the impact evaluations do not always report their methodology, nor pay sufficient attention to the program logic (the 'logframe'). And some of the impact evaluations have quality problems such as lack of a control group or baseline data. These evaluations also tended to stress quantitative methods while under-utilizing qualitative information such as level of beneficiary satisfaction.

3. COUNTRY CAPACITIES — FOR WHAT?

If we accept that Chile is a success story — albeit not nirvana — does it represent a best-practice model which other countries should emulate? The answer to this is ‘yes and no’. As a general proposition, it is dangerous to look for best-practice, and perhaps even good-practice, countries. Each is unique, with particular circumstances and realities, and each government has developed its M&E functions in particular directions, for particular purposes. Chile is an upper-middle-income country, with a very capable and respected civil service. Its very centralized government system and highly capable Finance ministry are not typical even of Latin American countries.

Countries such as Brazil have stressed a whole-of-government approach to the setting of program objectives and the creation of a system of performance indicators (May et al, 2006). Others such as Colombia have combined this with an agenda of rigorous impact evaluations. Yet others, such as Australia, the United States and the United Kingdom, have stressed a broader suite of M&E tools and methods: including performance indicators, rapid reviews, impact evaluations and performance audits (Lahey, 2005). Some countries have succeeded in building a whole-of-government M&E system, while others such as Uganda comprise an as yet uncoordinated and disparate collection of about 16 separate sectoral monitoring systems (Hauge, 2003). And the poorest countries — those which are required by multilateral donors to prepare *Poverty Reduction Strategies* — stress the regular collection of performance indicators to measure the millennium development goals (MDGs).

These country approaches are highly diverse. This tells us that not only are the starting points faced by each country different, but so are the destinations to which they aspire. There is no single, ‘best approach’ to a national or sectoral M&E system. Instead, it all depends on the actual or intended uses of the information which such a system will produce: whether to assist resource-allocation decisions in the budget process; to help in preparation of national and sectoral planning; to aid ongoing management and delivery of government services; or to underpin accountability relationships. The more ambitious government systems endeavor to achieve two or more of these desired uses.

4. LESSONS FROM EXPERIENCE — SUCCESS FACTORS FOR BUILDING COUNTRY M&E SYSTEMS

The growing literature on experience with strengthening government M&E systems suggests a broad agreement on a number of key lessons (Box 2).² First and foremost is that substantive demand from the government is a prerequisite to successful institutionalization — and by *institutionalization* I mean the creation of an M&E system which produces monitoring information and evaluation findings which are judged valuable by key stakeholders, which are used in the pursuit of good governance, and where there is sufficient demand for the M&E function to ensure its funding and its sustainability for the foreseeable future.³

Box 2: Lessons from Building Country M&E Systems

- Substantive government demand is a prerequisite for successful institutionalization
- Role of incentives
- Key role of a powerful ‘champion’
- Start with a diagnosis of existing M&E
- Centrally-driven, by capable ministry
- Build reliable ministry data systems
- Danger of over-engineering the system
- Utilization is the measure of ‘success’
- Limitations of relying on government laws, decrees and regulations
- Role of structural arrangements to ensure M&E objectivity and quality
- A long-haul effort, requiring patience

Achieving substantive demand for M&E is easier said than done. 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 an M&E function and funding it. So, there is frequently a chicken and egg problem: a lack of government demand for M&E because of lack of understanding of M&E and what it can provide; lack of understanding because of lack of experience with it; and lack of experience because of weak demand. The way around this conundrum is to try to increase awareness of M&E — its range of tools, methods and techniques — and of 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 M&E activities, and when they are made aware of other governments which have set up M&E systems and which value them

² In addition to the references cited above, see also, for example, African Development Bank and World Bank, 1998; Boyle, 2005; Compton, Baizerman and Stockdill, 2002; Development Bank of Southern Africa, African Development Bank and World Bank, 2000; Mackay, 1998, 2004; May et al, 2006; OECD, 1997a, b, 1998a, 2004; Schiavo-Campo, 2005; and UNDP, 2000.

³ See World Bank, 1997, especially chapter 9, “The challenge of initiating and sustaining reforms”.

highly.⁴ It can be persuasive to point to the growing evidence of the high returns to investment in M&E (Bamberger, Mackay and Ooi 2004). The supply side is also important — provision of M&E training, manuals, procedures, etc. This tends to be emphasized by those who view M&E in technocratic terms, as a stand-alone technical activity. However, while M&E expertise is certainly necessary if reliable M&E information is to be produced, a supply-side emphasis on its own does not address the demand side of the equation.

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. 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.⁵ 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, or by budget officials responsible for advising on spending options, or by a congress or parliament responsible for accountability oversight. This underscores the dangers of a technocratic view of 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, but because they directly support core government activities, such as the budget process, national planning, the management of ministries, agencies and programs, or to provide information in support of accountability relationships. Thus M&E systems are often linked to public sector reforms such as results-based management, performance budgeting, evidence-based policy-making, and the like; such initiatives share a number of common elements (May et al, 2006).

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 about the need to devote significant resources to create a whole-of-government M&E system. Government champions have played important roles in the creation of some of the more successful government M&E systems, such as those of Chile, Australia and Colombia.⁶ However, powerful champions 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.⁷

Creating a whole-of-government M&E system — whether focused solely on a system of performance indicators, or whether encompassing various types of evaluation and review — is not a minor effort. It involves the recruitment and training of staff to conduct or manage M&E and to use their findings; creation of the bureaucratic infrastructure to decide which government programs should be evaluated, and what issues should be addressed in the evaluations; creation of data systems and procedures for sharing information; and procedures for reporting M&E findings; etc. Like other systems, in areas such as financial management, or procurement, it takes sustained

⁴ For a number of years the Independent Evaluation Group within the World Bank has had a program of support to governments to help them institutionalize M&E systems. IEG has assembled evidence on good-practice government systems and on highly influential and cost-effective evaluations and reviews. See <http://www.worldbank.org/ieg/ecd/>

⁵ An excellent review is in World Bank, 1997. See also Levy and Kpundeh, 2004.

⁶ May et al, 2006; Mackay, 2004; and OED, 2003a (chapter 11).

⁷ OED, 2004a (annex E).

effort over a period of years to make an M&E system operate efficiently. The Organisation for Economic Co-operation and Development (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, p19)

Thus another feature of the successful government M&E systems listed earlier is the stewardship of this process by a capable ministry; in many developed and upper middle-income countries (e.g., Australia, Canada and Chile) 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 (e.g., a President's Office) or the budget process.⁸

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 which emphasizes 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 if it was to agree to fund them. This law was influenced, at least in part, by the series of rigorous impact evaluations of the *Progres*a program — while these are among some of the most expensive impact evaluations ever done, they have also been widely acknowledged as being of high quality and as having had an enormous impact on the government, persuading it to retain the *Progres*a program and to expand its size significantly when it morphed into the *Oportunidades* program. Governments in other countries find such examples of highly influential evaluations to be quite persuasive vis-à-vis the potential usefulness of evaluation and the merits of setting up a sound M&E system.

One point to note in passing: it is rarely if ever the case that a ministry which decides to create a strong M&E system has to start from scratch. Even in the poorest African countries there is usually a range of performance indicators available, and often qualitative program reviews are also undertaken with donor support. As we shall see below, the problem is more the poor quality and partial coverage of performance information, and its substantial underutilization.

A common mistake is to over-engineer 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 viewed as unwieldy for its accountability uses of the information, and it has subsequently reduced the number to around 300 (Castro, 2006). In Uganda, as already noted, one problem is the number of uncoordinated M&E systems — as many as 16 separate sector and sub-sector systems, which the government is now working to coordinate through a new *national integrated M&E strategy* (NIMES) (Hauge, 2003, Government of Uganda, 2005).

A problem in African countries, and perhaps in some other regions, is that while 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 over-worked officials at the facility level, who are tasked with providing the data for other officials in district offices and the capital, but who rarely

⁸ Bedi et al., 2006.

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.⁹

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 findings exist, to ensure 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*.

Another lesson is the limitations to relying 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 to legitimize an M&E system.¹⁰ Thus countries such as Colombia have a series of laws and decrees mandating evaluation, which was even enshrined in the Constitution in 1991. Yet in the intervening years, the fortunes of the government's evaluation system, SINERGIA, have waxed and waned, and it was only after a change in government in 2002 that the system has started to perform strongly (Castro, 2006). The point here is not that a law or decree mandating M&E is irrelevant: to the contrary, they can be a useful vehicle for legitimizing M&E, particularly in those countries where the presence of such a legal instrument is viewed as a necessary precondition if any government reform is to be perceived as worthwhile. 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 that the system produces. On the data side, some governments (e.g., Chile) rely on external audit committees to perform this function, some rely on the national audit office (e.g., Canada — see Mayne and Wilkins, 2005), while some rely principally on internal ministry audit units (e.g., Australia); some rely on central ministry checking of data provided by sector ministries (e.g., Colombia), while others have no audit strategy (e.g., Argentina — see Zaltsman, 2006).

On the evaluation side, issues of objectivity and credibility are particularly important. Chile 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

⁹ Note also the importance of data collected by national statistics offices, such as population censuses and household surveys. A good example of a review of existing data collected by ministries and the national statistics office in Uganda is provided by Kiryegyera, Nuwagaba and Ochen, 2005.

¹⁰ An interesting, if perhaps controversial, analysis of the nature and impact of the application of Napoleonic law in developing countries is provided by Beck, Demirguc-Kunt and Levine, 2002. By contrast, countries with a legal system based on English common law tend to interpret laws in a more pragmatic manner, stressing their adaptation as circumstances evolve.

to conduct the evaluations is entirely transparent.¹¹ The downside of this approach, however, 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 to impose management and program changes on the sector ministries (World Bank, 2005). This centrally-imposed system is quite unique. Most governments appear to rely on sector ministries to conduct evaluations themselves, although this raises questions about the reliability of self-evaluations. In the United States, the Office of Management and Budget (the finance ministry) rates the performance of government programs, and marks down those programs with either no M&E information about their performance, or with unreliable information (GAO 2004 — the OMB's procedure is called the *Program Assessment Rating Tool*, PART).

Countries which have built a government M&E system have found that it is a long-haul effort, requiring patience and persistence (for example, OECD, 1997a; Mackay, 1998; Lahey, 2005; and May et al, 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 be able 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 a well-functioning evaluation system (in terms of the quality, number and utilization of the evaluations) within four or five years; but in Colombia's case, it has taken 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 an approach of working in a focused, purposeful, and even intense, manner to build various components of the M&E system, and to seek to institutionalize them as quickly as possible.

It would probably be fair to say that most countries with well-performing M&E systems have not developed them in a linear manner — i.e., starting with a clear understanding of what their system would look like once fully mature, and then working progressively to achieve this vision. Instead, when one examines the experience of countries such as Australia (Mackay, 2004), Canada (Lahey, 2005), Chile (Zaltsman, 2006), Ireland (Boyle, 2005) and the United States (Joyce, 2004; and Lahey, 2005), it is evident that their 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 due partly to the different time it takes to build particular M&E functions — a system of performance indicators vis-à-vis the conduct of program reviews or rigorous impact evaluations. It would also appear to be due to a number of mid-course corrections made as the progress, or lack of progress, with particular M&E initiatives becomes evident. And there is also the important influence of external factors, such as a change of government, which can not only alter the direction of an M&E system but can lead to it being substantially run down or even abandoned — such as in Australia after 1997, and the United States after 1980 (GAO, 1987). Thus 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 those governments which have largely abandoned their M&E system, this would appear to reflect an ideological preference for

¹¹ Information on all successful tenders, on the evaluation terms of reference and the evaluation reports themselves, is available from the ministry's website: http://www.dipres.cl/fr_control.html

‘small government’, rather than a considered decision about the cost-effectiveness of their M&E system; the effects on the M&E system thus appear to simply 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. 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 its M&E system is exemplified by the review it commissioned the World Bank to conduct into the two principal evaluation components of the system (World Bank, 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 the congress. There are several diagnostic guides available to support such evaluations of government M&E systems. The first, and most comprehensive, has been published by the World Bank’s Independent Evaluation Group (OED, 1998).

5. KEY TRENDS INFLUENCING COUNTRY REALITIES: A DONOR PERSPECTIVE

OECD research suggests that there are cycles or trends in the types of public sector reform which countries adopt (for example, OECD, 1995, 1997a, b, 1998b, 2004, 2005). Reform priorities which were emphasized by developed countries as the 1990s progressed have included privatization, service standards, results-based management, contracting out, performance pay, decentralization, and performance budgeting, among others. Similar trends influence developing countries, some of which consciously look to adopt world ‘best practice’ approaches — although this can be a dangerous concept for M&E systems, given the need to tailor them to country circumstances and priorities. The influence of OECD trends on developing countries appears to operate with somewhat of a lag; a benefit is that these countries can learn about the successes and failures of implementation elsewhere. Thus in Latin America, for example, it is evident that a growing number of countries — as many as 20 — are currently working to strengthen their government M&E systems (May et al, 2006). Part of this trend appears to be explained by the role of the leading countries which provide a demonstration effect, including Chile, Colombia, Mexico and Brazil; these countries have, in turn, been influenced by OECD trends. But a common set of economic and social pressures are perhaps more important in Latin America. These include: continuing macroeconomic and budgetary constraints; pressures to improve and to extend both government service delivery and income transfers; and growing pressures for government accountability and for ‘social control’.

In Eastern Europe, those countries which have joined the European Union or are candidate countries are required to strengthen their M&E systems, and this is providing further impetus to the trend (Boyle, 2005).

The initiatives of international donors such as the World Bank are also having a strong influence on borrower countries, particularly those which are more dependent on international aid. The Bank’s debt relief initiative for heavily indebted poor countries has required — as a form of donor conditionality — the preparation of *poverty reduction strategy papers* by the countries, including measures of the extent of the country’s success in poverty-reduction efforts (Mackay, 2002). Donor emphasis on achievement of the millennium development goals (MDGs) is necessitating a similar focus. These in turn have required an analysis of the country’s M&E system, particularly the adequacy of available performance indicators. However, most of these poor countries have found it difficult to strengthen their monitoring systems, both in terms of data production and especially in terms of data utilization (World Bank and International Monetary Fund, 2004; Bedi et al., 2006).

There are also strong accountability pressures on international donors themselves to demonstrate results from the billions of dollars in aid spending each year. For the World Bank, these pressures have led to its *results agenda*, which entails among other things the requirement that the Bank’s country assistance strategies be focused firmly on the extent to which results are actually achieved, and the contribution of the Bank to them (World Bank, 2004). This initiative is leading to a considerably greater focus on the availability of M&E information about the performance of Bank projects in countries, as well as on broader issues of country performance vis-à-vis development objectives such as the MDGs. This in turn is necessitating a greater reliance on country M&E systems and the information they produce. And weaknesses in these systems are prompting the Bank to put more effort into provision of support to strengthen them, via Bank loans, grants and technical assistance. At the same time, there is somewhat of a changing emphasis in the loans made by the Bank and other donors, away from narrowly-defined projects

and towards programmatic lending — this entails provision of block funding (in effect, broad budget support). The absence of clearly-defined project activities and outputs from such lending also requires a focus on ‘big-picture’ results or outcomes of development assistance, and this in turn requires a greater reliance on country systems for national statistics, and for monitoring and evaluation of government programs.

These factors have combined to increase the level of donor involvement in helping developing countries build or strengthen their M&E systems. Thus by 2002, the World Bank was working with over 30 countries in this area (OED, 2002); the number has continued to increase since that time. And the Inter-American Development Bank in 2005 initiated a program of support to help countries in the Latin America and Caribbean region build their M&E systems; about 20 countries are expected to receive grant support in the next year or two. Other donors, such as the United Kingdom’s aid agency, the Department for International Development (DFID), are also increasingly active in this area. DFID, for example, has had a particular focus on poverty monitoring systems and on the use of performance information to support the budget process (for example, Booth and Lucas, 2001 a, b; Roberts, 2003).

One final trend which is influencing the focus on M&E is the growth in the number and membership of national, regional and global evaluation associations. In Africa, for example, there are now about 16 national associations, and some of these (such as Niger, Rwanda, Kenya and South Africa) have been particularly active in recent years, although sustaining their level of activity is a continuing challenge, as it depends very much on the presence and energy of local champions. There are also several regional associations, such as the African Evaluation Association (AfrEA) and, in Latin America, Preval and the new regional association, ReLAC. At the global level there is the International Development Evaluation Association (IDEAS) and the International Organisation for Cooperation in Evaluation (IOCE), which comprises the heads of regional and national evaluation associations. Multilateral and bilateral donors, including the World Bank, have provided funding and other support for a number of these evaluation associations.

These associations reflect, in part, the growing interest in M&E, and also the growing number of individuals working in the field. Such communities of practice have the potential to influence the quality of M&E work, and thus to facilitate the efforts of governments to strengthen their M&E systems. Some national associations, such as the one for Niger (RenSE), have involved close collaboration between academics, consultants, government officials and donor officials; and the major conferences of regional and global evaluation associations, such as AfrEA and IDEAS, are also bringing these constituencies together. This has the potential to spread awareness and knowledge of M&E among government officials, and thus to increase demand for it.

6. THE SPECIAL CASE OF AFRICA

The experience of African countries is relevant to poor countries in other regions, especially those preparing poverty reduction strategies. Africa also provides lessons on how to build M&E capacities incrementally, especially when there is the possibility of intensive donor assistance.

It is widely accepted that the extreme poverty situation facing most African countries provides a clear priority for intensive development support. Over 30 African countries have prepared an interim or final *poverty reduction strategy paper* (PRSP), which is required for access to debt relief. As noted earlier, these papers set development targets and are intended to report on results achieved. In practice, this has meant a focus on the extent to which a country has achieved the millennium development goals. Measurement of progress against the MDGs puts a premium on having adequate national statistics, which in turn is leading to donor support for statistical capacity-building, such as assistance for population censuses and household surveys. Countries, particularly their national statistical offices, appear keen to accept this support. PRSPs usually discuss their national monitoring (i.e., statistical) systems as synonymous with ‘M&E’, and the priority for ‘M&E’ has become a mantra which is widely accepted by governments and donors alike. In many cases, however, national monitoring systems are principally designed to meet donor data requirements (OED, 2004b). Moreover, what PRSPs end up focusing on is the amount of budget and other resources spent on national priorities, and national progress against the MDGs. These two issues are certainly important, but what is absent from this focus is what Booth and Lucas (2001, a, b) have termed the ‘missing middle’: the absence of performance information on the intervening steps in the results chain, involving government activities, outputs and services provided, and their outcomes; and the absence of much in-depth evaluative evidence linking government actions to actual results on the ground.

Some African governments, such as Uganda and Tanzania, understand well the importance of having reliable and comprehensive performance information available, and these are used intensively in preparing their national plans and in determining budget priorities (see, for example, Government of Tanzania, 2001; Ssentongo, 2004; Government of Uganda, 2004).

Such countries face at least three key challenges with their performance information, as discussed above. First, they rely heavily on administrative data whose quality is often poor. Second, an excessive volume of underutilized data is collected. One estimate for three sectors in Uganda is that their management information systems included a total of 1,000 performance indicators, requiring about 300,000 data entries by the typical service facility each year (Hauge, 2003). Third, there is often a plethora of uncoordinated sector and sub-sector data systems, using different data definitions, periodicity, etc. These are not trivial barriers to an efficiently-functioning monitoring system, and their importance has been recognized by the Ugandan government, which is setting up an integrated national M&E system to better coordinate these activities. Moreover, the World Bank and a number of other donors have agreed to rely on Uganda’s national system as the primary source of performance indicators for measuring the performance of their budget support and much of their other development assistance to Uganda.

Most African countries are simply too poor to be able to conduct evaluations and reviews, relying instead on donors for such work. A difficulty is the heavy burden placed on countries from meeting the evaluation requirements of donors, in terms of inspection missions, unharmonized donor evaluation criteria and methods, etc (OED, 2003b). However, donor cooperation can be facilitated through means such as sector-wide approaches (SWAPs). In Tanzania, for example, there is a health sector working group comprising government and donors which not only

analyzes sector performance and policies, but has also reviewed sector M&E systems and identified M&E capacity-building priorities. The sector working group also commissions evaluation or research into selected issues (Stout, 2001).

In addition, the move towards greater use of programmatic lending to countries also provides one way to mitigate the harmonization problem, because it reduces the scope for project-specific M&E, and thus the scope for balkanized donor M&E.

Although it would be unrealistic to expect most African countries to build comprehensive M&E systems, there are a number of important elements which are feasible. What follows is a list from which African and other countries preparing a PRSP could draw:

- Financial management information systems — to support better financial tracking
- Public expenditure tracking surveys — these enable ‘leakage’ or the effects of corruption to be traced
- Service delivery surveys
- Rapid appraisals — for example, of ‘problem’ projects or programs
- National and sectoral statistical collections — especially relating to national priorities such as the MDGs
- Sector ministries’ administrative data.

The only caveat with this list is that, in some senses, ‘less is more’. One danger to avoid is the tendency to over-engineer whatever M&E system is being created.

7. CONCLUSIONS, AND CHALLENGES FOR THE FUTURE

This paper argues that demand for M&E is key to the creation of a successful country M&E system. ‘Success’ is viewed here as having four dimensions: first, reliable monitoring information and evaluation findings; second, a high level of utilization of M&E findings; third, sustainability into the future; and fourth, country ownership.

But where demand is weak, the priority is to work to strengthen it by raising awareness of what M&E encompasses, what it can offer, and of other countries’ models. Evidence of highly influential evaluations is particularly persuasive. It makes sense to strengthen the supply side — number of evaluators, skills, and systems — as country demand increases, but there are real dangers from a purely supply-driven approach. Donors have a strong role to play, through direct support to governments and evaluation associations for M&E capacity-building, and also by means of advocacy and persuasion (‘jawboning’), and through the demonstration effect of good-practice M&E approaches. A challenge for donors is to broaden and deepen their support for M&E in countries.

Evaluation associations have a role to play. They can not only help share experience and expertise, but they also can provide a forum for greater dialogue between civil society, academia, governments and donors. But building national evaluation associations itself takes time, commitment, and at least a minimum level of resources; and sustaining them is likewise a challenge. Experience with evaluation associations in many regions is that they have depended on dynamic, committed individuals, and with this type of leadership, associations can be very active and — by any reasonable yardstick — highly successful.

This paper has examined a range of difficulties and challenges in achieving country-driven and country-owned capacities in M&E. But it would not be fair to reach a pessimistic conclusion. There are a growing number of countries with strong M&E systems, with a more committed set of stakeholders including government ministers, senior officials, donors and academia, and with well-functioning evaluation associations. If we compare the situation now with, say, ten years ago then an enormous change is evident. The priority for institutionalizing monitoring and evaluation is more widely recognized. Although there are no ‘quick fixes’ to achieve this objective, there is a rapidly-growing body of experience about how to do this, and what to avoid.

REFERENCES

African Development Bank and World Bank (1998) *Evaluation Capacity Development in Africa*. Washington, D.C.: The World Bank.

Bamberger, Michael, Keith Mackay and Elaine Ooi (2004) *Influential Evaluations: Evaluations that Improved Performance and Impacts of Development Programs*. Washington, D.C.: Operations Evaluation Department, The World Bank.

Beck, Thorsten, Asli Demirguc-Kunt and Ross Levine (2002) *Law and Finance: Why Does Legal Origin Matter?* NBER working paper no. 9379. Cambridge: National Bureau of Economic Research.

Bedi, Tara, Aline Coudouel, Marcus Cox, Markus Goldstein and Nigel Thornton (2006) *Beyond the Numbers: Understanding the Institutions for Monitoring Poverty Reduction Strategies*. Washington, D.C.: The World Bank, forthcoming.

Booth, David and Henry Lucas (2001a) *Desk Study of Good Practice in the Development of PRSP Indicators and Monitoring Systems: Initial Review of PRSP Documentation*. London: Overseas Development Institute.

____ (2001b) *Desk Study of Good Practice in the Development of PRSP Indicators and Monitoring Systems: Final Report*. London: Overseas Development Institute.

Boyle, Richard (2005) *Evaluation Capacity Development in the Republic of Ireland*. Operations Evaluation Department ECD working paper no. 14. Washington, D.C.: Operations Evaluation Department, The World Bank.

Castro, Manuel Fernando (2006) 'Colombia: Country Presentation', *Towards Institutionalizing Monitoring and Evaluation Systems in Latin America and the Caribbean*. Washington, D.C.: The World Bank and the Inter-American Development Bank (forthcoming).

Compton, Donald W., Michael Baizerman and Stacey H. Stockdill, eds (2002) *The Art, Craft, and Science of Evaluation Capacity Building*. American Evaluation Association, New Directions for Evaluation. San Francisco: Jossey-Bass.

Development Bank of Southern Africa, African Development Bank and World Bank (2000) *Monitoring and Evaluation Capacity Development in Africa*. Johannesburg: Development Bank of Southern Africa.

GAO (Government Accounting Office) (1987) *Federal Evaluation: Fewer Units, Reduced Resources, Different Studies From 1980*. (PEMD-87-9). Washington, D.C.: General Accounting Office.

____ (2004) *Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget*. (GAO-04-174). Washington, D.C.: General Accounting Office.

Government of Tanzania (2001) *Poverty Monitoring Master Plan*. Dar es Salaam: Government of Tanzania.

Government of Uganda (2004) *Poverty Eradication Action Plan (2004/5 – 2007/8)*. Kampala: Ministry of Finance, Planning and Economic Development, Government of Uganda.

____ (2005) *NIMES: National Integrated Monitoring and Evaluation Strategy*. Kampala: Office of the Prime Minister, Government of Uganda.

Guzman, Marcela (2003) *Systems of Management Control and Results-Based Budgeting: The Chilean Experience*. Santiago: Ministry of Finance, Government of Chile.

Hauge, Arild (2003) *The Development of Monitoring and Evaluation Capacities to Improve Government Performance in Uganda*. Operations Evaluation Department ECD working paper no. 10. Washington, D.C.: Operations Evaluation Department, The World Bank.

Hernandez, Gonzalo (2006) 'Mexico: Country Presentation', in E. May et al. (eds.) *Towards Institutionalizing Monitoring and Evaluation Systems in Latin America and the Caribbean*. Washington, D.C.: The World Bank and the Inter-American Development Bank (forthcoming).

Joyce, Philip G. (2004) *Linking Performance and Budgeting: Opportunities in the Federal Budget Process*. Washington, D.C.: IBM Center for The Business of Government.

Kiryegyera, Ben, Augustus Nuwagaba and Eric A. Ochen (2005) Results-Based Monitoring and Evaluation Plan for the Poverty Eradication Action Plan (PEAP). Unpublished report prepared for the Office of the Prime Minister, Uganda.

Lahey, Robert (2005) A Comparative Analysis of Monitoring and Evaluation in Four Selected Countries: Canada, United States, Australia and United Kingdom. Unpublished manuscript.

Levy, Brian and Sahr Kpundeh (2004) *Building State Capacity in Africa: New Approaches, Emerging Lessons*. Washington, D.C.: The World Bank.

Mackay, Keith, ed (1998) *Public Sector Performance — the Critical Role of Evaluation*. Washington, D.C.: Operations Evaluation Department, The World Bank.

____ (2002) *Evaluation Capacity Development (ECD) and the Poverty Reduction Strategy Initiative: Emerging Opportunities*. Proceedings of the AfrEA 2002 Conference. Nairobi: African Evaluation Association.

____ (2004) *Two Generations of Performance Evaluation and Management System in Australia*. Operations Evaluation Department ECD working paper no. 11. Washington, D.C.: Operations Evaluation Department, The World Bank.

May, Ernesto, David Shand, Keith Mackay, Fernando Rojas and Jaime Saavedra (eds.) (2006) *Towards Institutionalizing Monitoring and Evaluation Systems in Latin America and the Caribbean*. Washington, D.C.: The World Bank and the Inter-American Development Bank (forthcoming).

Mayne, John and P. Wilkins (2005) "Believe it or Not?": The Emergence of Performance Information Auditing', in Robert Schwartz and John Mayne (eds.) *Quality Matters: Seeking Confidence in Evaluating, Auditing, and Performance Reporting*, pp. 237-260. New Jersey: Transaction Publishers.

OECD (Organisation for Economic Co-operation and Development) (1995) *Governance in Transition: Public Management Reforms in OECD Countries*. Paris: OECD.

____ (1997a) *In Search of Results: Performance Management Practices*. Paris: OECD.

____ (1997b) *Issues and Development in Public Management: Survey 1996-1997*. Paris: OECD.

____ (1997c) *Promoting the Use of Programme Evaluation*. Paris: OECD.

____ (1998a) *Best Practice Guidelines for Evaluation*. PUMA policy brief no. 5. Paris: OECD.

____ (1998b) *Public Management Reform and Economic and Social Development*. Paris: OECD.

____ (2004) *Public Sector Modernisation: Governing for Performance*. OECD policy brief. Paris: OECD.

____ (2005) *Modernising Government: The Way Forward*. Paris: OECD.

OED (Operations Evaluation Department) (1998) *Evaluation Capacity Development: A Diagnostic Guide and Action Framework*. Operations Evaluation Department ECD working paper no. 6. Washington, D.C.: Operations Evaluation Department, The World Bank.

____ (2002) *2002 Annual Report on Evaluation Capacity Development*. Washington, D.C.: Operations Evaluation Department, The World Bank.

____ (2003a) *World Bank Operations Evaluation Department: The First 30 Years*. Washington, D.C.: The World Bank.

____ (2003b) *Toward Country-led Development: A Multi-Partner Evaluation of the Comprehensive Development Framework*. Washington, D.C.: Operations Evaluation Department, The World Bank.

____ (2004a) *Evaluation Capacity Development: OED Self-Evaluation*. Washington, D.C.: Operations Evaluation Department, The World Bank.

____ (2004b) *The Poverty Reduction Strategy Initiative: An Independent Evaluation of the World Bank's Support Through 2003*. Washington, D.C.: Operations Evaluation Department, The World Bank.

Roberts, John (2003) *Managing Public Expenditure for Development Results and Poverty Reduction*. ODI working paper no. 203. London: Overseas Development Institute.

Schiavo-Campo, Salvatore (2005) *Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience*. Operations Evaluation Department ECD working paper no. 13. Washington, D.C.: Operations Evaluation Department, The World Bank.

Ssentongo, Peter (2004) *The National Integrated Monitoring and Evaluation Strategy*. Unpublished paper presented at the 2004 conference of the African Evaluation Association.

Stout, Susan (2001) *Tanzania: Rapid Assessment of Monitoring and Evaluation in the Health Sector*. Draft report.

UNDP (United Nations Development Programme) (2000) *Evaluation Capacity Development in Asia*. New York: UNDP.

World Bank (1997) *World Development Report 1997: The State in a Changing World*. Oxford: Oxford University Press.

_____ (2004) *The World Bank Annual Report 2004*. Washington, D.C.: The World Bank.

_____ (2005) *Chile: Study of Evaluation Program — Impact Evaluation and Evaluations of Government Programs*. Washington, D.C.: The World Bank.

_____ and International Monetary Fund (2004) *Poverty Reduction Strategy Papers — Progress in Implementation*. Washington, D.C.: The World Bank and International Monetary Fund.

Zaltsman, Ariel (2006) *Experience with Institutionalizing Monitoring and Evaluation Systems in Five Latin American Countries: Argentina, Chile, Colombia, Costa Rica and Uruguay*. Independent Evaluation Group ECD working paper no. 16. Washington, D.C.: Independent Evaluation Group, The World Bank.

Other Papers in This Series

- #1: Keith Mackay. 1998. *Lessons from National Experience*.
- #2: Stephen Brushett. 1998. *Zimbabwe: Issues and Opportunities*.
- #3: Alain Barberie. 1998. *Indonesia's National Evaluation System*.
- #4: Keith Mackay. 1998. *The Development of Australia's Evaluation System*.
- #5: R. Pablo Guerrero O. 1999. *Comparative Insights from Colombia, China and Indonesia*.
- #6: Keith Mackay. 1999. *Evaluation Capacity Development: A Diagnostic Guide and Action Framework*.
- #7: Mark Schacter. 2000. *Sub-Saharan Africa: Lessons from Experience in Supporting Sound Governance*.
- #8: Arild Hauge. 2001. *Strengthening Capacity for Monitoring and Evaluation in Uganda: A Results Based Management Perspective*.
- #9: Marie-Hélène Adrien. 2003. *Guide to Conducting Reviews of Organizations Supplying M&E Training*.
- #10: Arild Hauge. 2003. *The Development of Monitoring and Evaluation Capacities to Improve Government Performance in Uganda*.
- #11: Keith Mackay. 2004. *Two Generations of Performance Evaluation and Management System in Australia*.
- #12: Adikeshavalu Ravindra. 2004. *An Assessment of the Impact of Bangalore Citizen Report Cards on the Performance of Public Agencies*.
- #13: Salvatore Schiavo-Campo. 2005. *Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience*.
- #14: Richard Boyle. 2005. *Evaluation Capacity Development in the Republic of Ireland*.

Other Recommended Reading

Operations Evaluation Department (OED). 2004. *Evaluation Capacity Development: OED Self-Evaluation*.

OED. 2002. *Annual Report on Evaluation Capacity Development*.

OED. 2004. *Influential Evaluations: Evaluations that Improved Performance and Impacts of Development Programs*.

OED. 2005. *Influential Evaluations: Detailed Case Studies*.

OED. 2004. *Monitoring and Evaluation: Some Tools, Methods and Approaches*. 2nd Edition.

Independent Evaluation Group (IEG). 2006. *Conducting Quality Impact Evaluations Under Budget, Time and Data Constraints*, forthcoming.

Development Bank of Southern Africa, African Development Bank and The World Bank. 2000. *Developing African Capacity for Monitoring and Evaluation*.

K. Mackay and S. Gariba (eds.) 2000. *The Role of Civil Society in Assessing Public Sector Performance in Ghana*. OED.

Other relevant publications can be downloaded from IEG's ECD Website:

><http://www.worldbank.org/ieg/ecd/><