



**Building Country Capacity for Monitoring
and Evaluation in the Public Sector:
Selected Lessons of International
Experience**

This paper draws on the findings from two in-depth country case studies, and on the growing literature on efforts to help countries strengthen their monitoring and evaluation systems. It identifies a number of lessons about what works, what does not, and why.

ECD Working Paper Series ♦ 13

Salvatore Schiavo-Campo



Copyright 2005
Operations Evaluation Department
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. OED aims to identify and help develop good-practice approaches in countries, and to share the growing body of experience with such work.

The OED 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 Operations Evaluation Department or any other unit of the World Bank, its Executive Directors, or the countries they represent.

TABLE OF CONTENTS

Foreword.....	i
Abbreviations.....	ii
1. INTRODUCTION.....	1
2. THE MEANING OF CAPACITY-BUILDING IN MONITORING AND EVALUATION.....	2
2.1 The Emphasis on Institutions.....	2
2.2 The Concept of Capacity and the Four Pillars of Evaluation Capacity Development (ECD).....	2
3. CRITERIA FOR STRENGTHENING M&E CAPACITY.....	4
3.1 Look at the Legitimacy Superstructure.....	4
3.2 Encourage Affirmative Transparency.....	4
3.3 Assure Appropriate Scope and Timeframe.....	4
3.4 Fit ECD with Local Realities.....	5
3.5 Focus ECD Efforts on ECD.....	5
4. MAJOR ISSUES AND CONSIDERATIONS.....	7
4.1 By Objectives or By Results?.....	7
4.2 In-house or Independent M&E Capacity?.....	7
4.3 Connecting to the Legislature?.....	9
4.4 Performance Indicators: Is the Devil in the Details?.....	10
4.5 Rebalancing Monitoring and Evaluation?.....	11
4.6 Entry Points—Macro, Meso, or Micro?.....	11
4.7 Grand Vision or Basic Plumbing?.....	11
5. A CONCLUDING WORD.....	13
BIBLIOGRAPHY.....	14

FOREWORD

The Operations Evaluation Department (OED) of the World Bank has a long-standing program of support to strengthen monitoring and evaluation (M&E) capacities in developing countries, as an important part of sound governance. As part of this support, OED 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/oed/ecd/>

OED is accumulating considerable experience from these efforts to help countries strengthen their monitoring and evaluation systems—known generically as *evaluation capacity development* (ECD). The insights from this experience have been heightened by a self-evaluation of these ECD efforts, which OED conducted in 2004. This evaluation identified a number of results of OED’s ECD efforts—defined as the outputs and outcomes of this work. A growing number of World Bank country teams and other donors are also giving priority to ECD as an important type of capacity-building work.

This paper was prepared by Salvatore Schiavo-Campo, a retired senior Bank staffer with substantive experience in development issues, particularly governance, public sector management, and monitoring and evaluation. Mr Schiavo-Campo conducted two in-depth reviews of OED’s country-based ECD work, in Egypt and Uganda, and these are included in OED’s ECD self-evaluation report. This paper draws on a number of insights and lessons from these country reviews and from the growing ECD literature. The paper is written from the perspective of international donors providing ECD support to governments; however, it should prove to be of broader interest to others in the development community, and to governments themselves.

OED’s task manager for this paper was Keith Mackay.

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

ABBREVIATIONS

CABRI	Collaborative African Budget Reform Initiative
ECD	Evaluation Capacity Development
ICT	Information & communication technologies
IDEAS	International Development Evaluation Association
IDF	Institutional Development Fund grant
IPDET	International Program for Development Evaluation Training
M&E	Monitoring and evaluation
NGO	Non-governmental organization
OECD	Organisation for Economic Co-operation and Development
OED	Operations Evaluation Department
OMB	Office of Management and Budget
UK	United Kingdom
WBI	World Bank Institute

1. INTRODUCTION

The World Bank's Operations Evaluation Department (OED) recently completed a self-evaluation of its evaluation capacity development (ECD) activities. The objective of these activities is to assist countries to strengthen their own monitoring and evaluation (M&E) capacity. OED's ECD activities are conducted by a dedicated small unit within the Department (OED 2004). The unit has provided "high-intensity"¹ support to two countries (Egypt and Uganda), and various other types of support to an additional 32 countries. The main conclusion of the ECD self-evaluation has been that—despite the small scale of the activities and the obvious difficulty of disentangling OED's contribution from other factors—these activities did contribute to the significant progress made in some countries to develop their own capacity to monitor public action and evaluate its results.

Following completion of the self-evaluation, OED decided to complement it with a small and targeted effort to bring together some of the main lessons learned from ECD activities, and define selected major issues for consideration and as guidance to further work. The present paper is the outcome of this effort. It draws from the OED ECD record, but also from the author's personal experience of almost 40 years in development assistance, including public sector management and M&E activities. A number of caveats and disclaimers are therefore especially pertinent here:

- Although the emphasis of the paper is on the lessons from past experience in helping to build countries' M&E capacity, some of the criteria and issues apply to M&E generally;
- No attempt has been made to address M&E capacity-building issues outside government—among other things, the subject of assessing the performance of public enterprises and quasi-government entities has not been touched;
- The "lessons learned" which are reported here are of necessity selective, and no claim of comprehensiveness can be made, although the cross-cutting themes highlighted in the paper do appear to touch on most of the relevant concerns;
- The main points and conclusions are offered here without extensive argumentation or documentation—in order to keep the paper to a reasonable length, and for the readers' convenience; and, most importantly,
- The treatment of the issues is frankly subjective and, as noted, stems partly from personal experience—hence, several of the lessons presented should be seen more as informed hypotheses than as firm conclusions.

¹ OED has defined its country-based ECD work as high intensity where it involves a broad range of different types of capacity-building support, and where that support is provided over a sustained period. Support activities include, for example, M&E training, in-country presentations on ECD, diagnostic work and preparation of action plans, sponsorship of scholarships for government officials and representatives of civil society, and support for national evaluation associations. See OED 2004.

2. THE MEANING OF CAPACITY-BUILDING IN MONITORING AND EVALUATION

2.1 The Emphasis on Institutions

Notwithstanding the reality that the *new institutional economics* is now at least 30 years old,² the term *institution* is still too often confused with *organization*, with the main consequence, among others, of evading difficult institutional problems by reshuffling organizational structures. But the behavior of human beings is governed by the rules and incentives within which they are expected to behave—the real *institutions*—and not by the happenstance of their location in one or another box of the organization. In turn, because institutions comprise both formal, visible rules and informal rules, usually less visible, the total accumulated stock of institutions in any country is very large. From this stems the phenomenon Douglass North called “path dependence”, i.e., the heavy weight of accumulated rules and regulations, formal and informal, and the resulting constraints on the scope and speed of institutional change.³

The link to M&E arises from the reality that most incentives and penalties in the public sector attach to the formulation and early phase of execution of policies and programs, with much less weight given to effective monitoring and to evaluation of their results—let alone to tying the feedback loop back to the formulation of new programs. ECD efforts must therefore pay special attention to introducing or strengthening actual incentives for M&E, as they are likely to be weak or non-existent in the first place. It also follows from the emphasis on institutions that ECD efforts—like all activities to foster institutional change—need to be sustained over time if lasting M&E capacity is to be built, as explored at greater length later in this paper.

2.2 The Concept of Capacity and the Four Pillars of Evaluation Capacity Development (ECD)

It is frequently overlooked that, conceptually and practically, *capacity is relative*, and mainly in relation to the complexity of the tasks the system is asked to perform. The 1970s notion of “appropriate technology”, largely bypassed by events in most fields, is still alive and well in public administration. Indeed, experience over the past 50 years shows a troublesome supply-driven dynamic at work, whereby external technical assistance sometimes pushes complex new practices onto a reasonably functioning system, and thus creates capacity constraints where none may have existed. In turn, these “capacity limitations” are then used to justify the need for continuing assistance.

The intention here is certainly not to argue against the need for improvements, nor to contest the obvious reality that, in most developing countries, weak administrative capacity sets an important constraint to public sector efficiency and effectiveness. The intention here is only to underline that capacity-building in general and M&E in particular is far more than just training, and requires complementary improvements in four major directions. These four pillars of M&E capacity building are improvements in: (i) institutional capacity, i.e., a move from less efficient to more efficient accountability rules and incentives; (ii) organizational capacity, i.e., the tailoring and adaptation of the organizational architecture of M&E government entities to the new and

² The main original contributions were by Ronald Coase (whose most important articles are found in Coase 1988) and Armen Alchian (see primarily Alchian 1977).

³ See North 1991.

more efficient accountability rules and incentives; (iii) information & communication technology (ICT) capacity, using informatics for better and more timely information on results; and, of course (iv) human capacity, through training in M&E, but targeted on the sort of skills that are suited to the particular institutional and organizational context, and thus will actually be used and reinforced after they are imparted.

3. CRITERIA FOR STRENGTHENING M&E CAPACITY

3.1 Look at the Legitimacy Superstructure

Experience shows that almost any regime can produce short-lived results, but only governments with a sound basis of legitimacy can produce sustainable results. Because in the public sector the sustainability of results can be verified only after many years, sometimes even decades, it is that much more important to pay attention at the start to the quality of the decision-making processes and, more broadly, to the governance context within which M&E capacity is to be built. This criterion is underlined by the propensity and ability of unrepresentative regimes to manipulate facts and “results”—either as whitewash for the benefit the ruling elite (e.g., “Potemkin Villages”) or for external donor agencies. There is never a good substitute for reliable reality checks, but a healthy suspicion is appropriate when the “results” are produced in a context of weak governance and accountability. Obviously, it would be neither practical nor desirable for the ECD activities to themselves incorporate a governance assessment. However, plenty of secondary-source information is available on the quality of governance in most countries, and even a cursory look at that information can strengthen the foundation for ECD activities.

3.2 Encourage Affirmative Transparency

Related to the previous criterion is the fact that the sustainability of results depends largely on the credibility of public decisions, which, in turn, requires that such decisions be *understandable*. Transparency, the handmaiden of accountability, calls for affirmative outreach, not just for passive openness.

The link of this point to our subject is that efforts at developing evaluation capacity must be mindful of the transparency dimension—and mainly, of the strength of the “communications bridges” between core and line ministries; between central and subnational levels of government; and between government and the service users, as well as the groups expected to help implement the decisions. Again, to try and go too far in this direction would compromise the chances of producing concrete M&E results. The criterion here calls only for keeping *affirmative transparency* in mind when evaluating the likelihood of long-term sustainability of public management policies and programs, and fostering such transparency as and when appropriate.

3.3 Assure Appropriate Scope and Timeframe

The sterile old debate between “big bang” and “gradualism” is happily being replaced by the more scientific and commonsensical approach of taking advantage of windows of opportunity when they do exist, while building slowly a broad basis for M&E in the public sector. Indeed, even the term *reform* is inappropriate to ECD efforts; *improvement* is a much clearer and more realistic objective.⁴ The term “reform” should be reserved to changes of broad scope and affecting many interrelated parts of the public sector. Efficiency/effectiveness in the public sector can also be raised over time by meaningful specific improvements.

⁴ This was one of the conclusions of a recent conference of the Collaborative African Budget Reform Initiative (CABRI), hosted by the South African Treasury in Pretoria on Dec 1-3, 2004.

The link of this point to our subject is that the *timing* of M&E capacity-building assistance should be carefully tailored to the scope and timeframe of the systemic institutional improvements under way in the public sector. In developing countries, where processes are generally more fluid, M&E should be brought to bear at an early stage—even if in partial manner and only focusing on some of the key issues. This is the flip side of the long-gestation issue raised at the start of this section. While there is no substitute for allowing sufficient time for the evaluation of long-ripening outcomes, the *habits* of M&E should be built as soon as possible in the reform process—and preferably as an integral part of the reforms themselves (as for example in Uganda, Chile and, to a lesser extent, Ghana and Mozambique). Evaluation of effectiveness, let alone impact, must come long after the completion of the activities themselves; but evaluation of process and outputs can take place at a very early stage and feed more quickly into the ECD effort. Paradoxically, then, early M&E intervention is most applicable to both ends of the public service continuum. It is at the “bottom end”—the interface with the citizens—that the connection between physical outputs and accountability is clearest and most immediate (e.g., trash collection, pest control, water purification.) But it is at the “top end” of policy review and program formulation that process indicators are most relevant—and *performance* can be assessed by judicious assessment of the views of the main participants in the process.

3.4 Fit ECD with Local Realities

It is very difficult to define the appropriate scope and timeframe of ECD assistance efforts unless there is a good understanding of local realities—and especially institutional and administrative capacity. Recalling that capacity is a relative concept, international experience (the failures more than the successes) has shown that *simplicity* is a guiding criterion for efforts to introduce M&E successfully. As well stated in a recent World Bank Institute (WBI) publication: “Instead of dwelling on politically unrealistic ‘best practice’ reforms, the focus shifts to a ‘good fit’ approach using modest, viable initiatives, with observable results”.⁵ The ECD assistance-provider should particularly be on the lookout for evidence of a supply-side dynamic at work, for, as noted earlier, it is mainly by advocating inappropriately complex M&E methods that the advocates assure themselves of a continuing market for their expertise. It is the case that the Bank has not always followed its own prescription in this respect.⁶ This is especially regrettable as, with its lack of vested interest, the World Bank has a unique role to play as debunker of fashionable myths.

One more observation: the capacity to monitor and evaluate government action is too important to be left to government. The assessment of M&E capacity should go beyond government, and should consider also possibilities for using the users of government services themselves to provide feedback and contestability. Appropriate participation by civil society can augment limited governmental capacity for M&E. The role of NGOs is especially relevant here. The Uganda experience, among others, has shown the potential contribution of NGOs to effective M&E as well as their concern with the risk of being co-opted. The issue is delicate, but a balance between cooperation and independence can be struck.

3.5 Focus ECD Efforts on ECD

It is difficult in practice to keep a bright line between helping to develop local capacity for M&E in certain subject matters, and providing technical advice in the subject matters themselves. It is,

⁵ Levy and Kpundeh 2004.

⁶ A frequent observation made, among others, by participants to the CABRI Conference (see footnote 4).

however, necessary to practice self-restraint and resist the temptation to derail the focus from ECD to provision of substantive technical assistance in the subject. It is necessary, first, because intrusions by M&E specialists through the ECD backdoor into advice on matters within the competence of the international donor's operational unit concerned can be resented, with the result that the door may slam shut onto the ECD activities themselves. And it is necessary, second, because M&E specialists do not necessarily possess the requisite technical competence in the subject matter. On the other hand, it is also risky to build a firewall between technical advice in the subject matter and advice on how to build the capacity to monitor and evaluate activities in the same area. The solution to this dilemma rests on two pillars: clear ownership of the ECD agenda by the government, and good cooperation between ECD efforts and the aims of the donor operational unit concerned.

4. MAJOR ISSUES AND CONSIDERATIONS

Although suggestions are advanced in this section on how to approach various issues, the purpose is mainly to identify those that should be addressed frontally and at a very early stage in ECD activities. Errors or oversights can always be corrected at a later time, but it is far more cost-effective to confront the relevant questions **at entry**—i.e., on commencing ECD efforts in a country—primarily in light of local institutional and administrative realities, as noted earlier.

As Box 1 illustrates, many of the issues discussed in this section came to the fore from the evaluation of OED’s ECD efforts in various countries, and particularly in Egypt and Uganda, the two cases of “high-intensity” ECD support by OED.⁷

4.1 By Objectives or By Results?

It is well known that a classic approach to evaluation (to assess the degree of achievement of the objectives stated at inception) and a pragmatic approach (to assess the positive results actually achieved, regardless of whether they match the initial objectives) do not necessarily produce the same verdict. The classic approach has been criticized for lending itself to excessive formalism, and enabling a mutation of simple and useful ideas into monsters of red tape.⁸ But, because M&E capacity improvements are long-gestating, as noted earlier, the pragmatic approach can degrade into an alibi for perennial postponement of reckoning and accountability. On balance, there is probably no alternative to the classic approach of evaluation by objectives, provided that it is complemented by some form of mid-course assessments. Thus, evaluation shades into supervision, strengthening the oft-mentioned need for closer collaboration between the M&E capacity efforts and the donor operational unit concerned.⁹

4.2 In-house or Independent M&E Capacity?

The standard assumption in most ECD programs is that M&E capacity should be created within the government itself. Whether this is correct or not, it is surely fallacious to assume from the start that because evaluation **of** government activities is important, it therefore must be conducted **by** government. In-house evaluation has the obvious advantage of inside expertise, savvy, and intimate operational knowledge of the programs being evaluated.¹⁰ The other side of the coin is a natural tendency to overstate results,¹¹ and, where accountability is ambiguous or loose, even to provide a coat of respectability and whitewash to failed programs.¹² The advantages of external

⁷ The two case studies are in Annexes D and E of OED’s ECD self-evaluation report (OED 2004).

⁸ The frequent abuse of the logical framework concept is a case in point.

⁹ Ideally, some M&E capacity-building should be routinely included in project supervision—whether through direct participation in missions or through other forms of assistance and advice.

¹⁰ This is the approach currently followed, for example, in Australia, where line departments have responsibility for conducting any evaluations of the programs they manage (see Mackay 2004).

¹¹ This occurs even in highly-developed systems with strong accountability mechanisms, as for example in the UK—see Box 2.

¹² The United States has created a framework to endeavor to address this potential problem. Line departments and agencies are required to rate the performance of all their programs, based on their available M&E findings. These self-ratings are reviewed—and often over-ridden—by the Office of Management and Budget (OMB), which manages the budget process in the federal government. OMB’s reviews of the departments’ self-ratings include an assessment of the reliability of the departments’ M&E findings (and this constitutes, de facto, a critique of departments’ M&E methods).

evaluation are, first, its presumptively stronger independence and, second, the greater probability that the evaluators are familiar with similar programs in other sectors or other countries.¹³

Box 1: Diverse Lessons From the ECD Country Assistance Experience*

- A clear and agreed mandate for the external agency providing ECD support (in this case, OED) is important for the effectiveness and sustainability of the ECD work, as in Uganda;
- Sometimes, placing M&E on the government agenda is itself a significant accomplishment, as in Sri Lanka and Malawi;
- Also a significant accomplishment is to help build a common M&E language and conceptual understanding, as in the Egypt program;
- Cross-fertilization of ideas and country comparisons can be helpful to generate demand, as in the effective use of the Chile experience;
- Assessing the cost-effectiveness of ECD work demands proper costing, but non-Bank funding is sometimes not included, while in other cases the ECD effort had positive but unaccounted-for externalities in supporting other development activities;
- Over-focus on “macro” or “big-picture” public sector management issues is inimical to robust M&E “at the coal face”—better links of ECD activities with specific sectoral staff and service providers would be important;
- Similarly, focusing M&E on the provision of services of specific sectors can be a highly promising entry point, which is often neglected;
- Dedicated funding can be an important component of the initial phase of ECD activities, as in the Brazil and Mexico Institutional Development Fund (IDF) grants;
- Conversely, the availability of dedicated funding is insufficient to advance the M&E agenda if it is not targeted clearly on M&E capacity building, as the Egypt IDF;
- Excessive attempts at monitoring, through a large number of indicators, produce little effective monitoring, as in Uganda;
- Inattention to bureaucratic realities produces delays, as in Uganda, or weak ownership, as in Egypt;
- Strict length constraints are important to produce meaningful mission statements, as the mixed experience with the budgeting pilots in Egypt shows; and finally
- It is not advisable to rely overmuch on one-off M&E workshops, or similar events. While these events can be important to put M&E “on the map” at the start, sustained capacity-building efforts are required to improve the performance of the public sector on a lasting basis—as stressed in the concluding section of this paper.

** Drawn from the material in OED’s 2004 ECD self-evaluation report, particularly the Uganda and Egypt case-studies in Annexes D and E prepared by the author (OED 2004).*

¹³ Chile is one of a small number of countries which rely largely on commissioning independent evaluations, although with the process managed by a government ministry.

These advantages are not exclusive, however. The literature on organizational effectiveness and the experience of countries has shown that in-house evaluation organs can be assured of a degree of independence very close to that enjoyed by external entities.¹⁴ Conversely, if external evaluators are afforded the chance to contribute on a regular basis, they will develop the understanding of operations needed for an informed assessment. The disadvantages, too, are not exclusive: in particular, if the governance climate is not conducive to robust and candid evaluations, it is a virtual certainty that even the best external evaluations will be hidden away, or distorted to produce the desired results.

The choice is therefore an entirely pragmatic one. Experience shows that thorough evaluations (especially policy evaluations) require a substantial input by economists, researchers, auditors, etc., and a highly-developed database.¹⁵ It is evident that such skills do not exist in most developing countries' governments, and where they do exist they are best used in designing and running sound programs—not in evaluating them. It is both inevitable and desirable that M&E in developing countries should be conducted largely on the basis of expertise and inputs external to the government and, in many cases, external to the country. At the same time, exclusive reliance on external expertise will result in evaluations that are irrelevant—no matter how technically sound—because of the absence of an organic link to the administrative apparatus. Helping to create M&E capacity in developing country governments must therefore rest on two complementary approaches: (i) fostering evaluation capacity outside the government, through a variety of educational or other activities;¹⁶ and (ii) helping directly to create a small but strong in-house capacity to design, guide, contract, *and monitor* the external evaluators. Such in-house capacity must not be confined to a separate small “evaluation ghetto”, but requires systematic connections to the public finance function and to the line ministries, in whatever manner is effective in the specific country.

4.3 Connecting to the Legislature?

While intra-governmental connections to the evaluation function are essential, they are not sufficient for a lasting positive impact of M&E on the efficiency and effectiveness of public resources. It is another unfortunate lesson of experience that neither the executive branch nor the international donors bring the legislative branch into the process of monitoring and evaluating the results of government activity. In countries of the Anglo-Saxon administrative tradition, parliaments usually have both a public accounts committee and a committee on public enterprises—to monitor the actions of government and those of parastatals, respectively. These committees are traditionally headed by the opposition party. Similarly, the supreme audit institution reports to parliament rather than to the executive. However, even where the role of the legislature is influential, it is largely focused on integrity of input use and corruption issues. The same is true of the supreme audit institution, with its traditional emphasis on financial audit and compliance.

¹⁴ At the Bank, OED itself is a case in point, with sufficient safeguards precluding its ECD activities from compromising in any way the integrity and independence of OED's ex post evaluations.

¹⁵ In the already-cited case of the UK, the number of professionals involved in evaluation trebled.

¹⁶ As, for example, through the World Bank Institute's (WBI) introductory course on M&E, the OED-supported *International Program for Development Evaluation Training* (IPDET) program, or, indirectly, through the activities of the *International Development Evaluation Association* (IDEAS), in addition to possible country-specific activities linked to local higher-education institutions. Each of these examples is described in OED 2004.

This focus on financial integrity is entirely appropriate in most developing countries, where indeed the major challenge remains to assure that the public's money is neither stolen nor misallocated. At the same time, however, the existing opportunities to enlist the legislative branch as a partner in M&E, by and large have not been explored. Especially in African countries, the executive branch has a pervasive view of the legislature as an adversary or, at best, a nuisance to be manipulated and pacified.¹⁷ This attitude may perhaps be understandable in light of the troubled governance history of much of the continent—but is nonetheless worrisome for the future of M&E capacity-building on the continent. Indeed, the primary consumers of good M&E should be the parliaments themselves, as representatives of the population whose resources are mobilized for public purposes. Much is made of the need to strengthen legislative capacity to interact intelligently with the process of formulation of the budget—i.e., the definition of the uses to which public money is to be put. Little is made of the equally weighty need to strengthen their capacity to assess the execution of the budget—i.e., the results of the uses of public money. Yet, without such capacity, the feedback loop into the formulation of the next budget cannot be closed; legislative oversight will remain weak no matter how much “capacity” is increased at the front end of the budget process; and the best ex post evaluations will remain detached from budgetary realities—“butterflies encased in amber”.

That said, it will be essential to be realistic and highly selective in pursuing initiatives to strengthen legislatures' capacity to evaluate the efficiency and effectiveness of public action. In most developing countries the clear priority continues to be the need to assure constructive legislative participation in formulating a sound and realistic budget to begin with. Although it is possible to execute badly a well-formulated budget, it is impossible to execute “well” a badly-formulated budget.¹⁸

4.4 Performance Indicators: Is the Devil in the Details?

While on the subject of bad information, the best organizational arrangements for monitoring and evaluating government programs rest on sand unless strong and realistic results indicators are elaborated to begin with. The literature on performance measurement in the public sector is too vast and extensive to even attempt to summarize here. However, international experience yields four simple but crucial lessons: (i) simplicity and clarity are key to the definition of good performance indicators; (ii) performance indicators must never be chosen only from the top-down, but their definition should include the integral participation of front-line staff and service users themselves; (iii) the performance of the performance indicators themselves must be kept under constant review in real time, and adjustments made as required; and therefore, (iv) beginning with specific public services, expansion to other services and sectors should be gradual and should internalize the experience of the previous stages.

Most importantly, performance measurement is a means, not an end in itself. The right question is not “how to introduce performance measures”, but “how to strengthen performance orientation”. The answer to the latter question may or may not call for explicit use of quantitative performance indicators. The developing world offers many illustrations of the waste of resources or even counterproductive incentives produced by a confusion of means with ends, and by an unthinking

¹⁷ This attitude was most recently in evidence at the Conference of the Collaborative African Budget Reform Initiative (CABRI), hosted by the South African Treasury in Pretoria during December 1-3, 2004.

¹⁸ See on this point, as well as the general role of the legislature in the budget process, Schiavo-Campo and Tommasi 1999.

push for results measurement as an end in itself.¹⁹ It could be perhaps the highest contribution of external assistance to M&E capacity in developing countries to hammer in this essential point, and help the countries develop a robust dialogue on accountability for results, and introduce a few, simple, clear, and monitorable performance measures when, and only when, they are appropriate.

4.5 Rebalancing Monitoring and Evaluation?

The previous point is related to the distortion found in most countries of an excess of monitoring and a dearth of genuine evaluation. There are two major issues here. *First*, a few relevant and clear indicators are more likely to produce robust monitoring than a complex plethora of indicators. Having too many monitoring indicators means having too little monitoring, not to mention wasting time and resources. This problem was found, for example, in Uganda despite the generally favorable ECD experience in that country, but is common throughout the developing world. *Second*, a surfeit of monitoring initiatives also tends to produce neglect of evaluation—and without that crucial feedback loop to correcting mistakes and exploiting good opportunities the whole aim of M&E is compromised. A clear tradeoff appears: unless ECD efforts can give solid advice on simple and targeted monitoring indicators, neither robust monitoring nor meaningful evaluations are likely to emerge.

4.6 Entry Points—Macro, Meso, or Micro?

The preponderance of ECD work, by the Bank and other donor institutions, has been centered on macro-level, whole-of-government issues, e.g., public expenditure management, and usually has been conducted with core ministry interlocutors. However, support for M&E depends partly on visible results on the ground, which call for attention to improving monitoring and evaluation at meso level—i.e., at the level of sector ministries—or even for specific services to the public—the micro level. It is true that, in the absence of an enabling macro environment and support from the top, such meso-level and micro-level efforts are likely to remain scattered, and will add up to little over time. But conversely, there is a risk that in the absence of concrete improvements in service delivery—through stronger M&E at meso- and micro-level—the enabling macro environment will eventually collapse for lack of visible impact and public support. Accordingly, without arguing for a change in ECD focus from macro to meso and micro issues, the time may have arrived for assuring complementary M&E assistance activities “in the trenches”. Two of the several implications of this suggestion are, first, the need for closer collaboration with sector and service specialists and, second, the utility of assisting local government in targeted M&E of the basic services provided at local level.

4.7 Grand Vision or Basic Plumbing?

The preponderance of evidence demonstrates the importance for developing countries to focus on the “basic plumbing” of M&E capacity. It would take us too far along to elaborate on this basic point, but one example should suffice. In evaluating public expenditure, the first order of M&E business should be expenditure tracking—not in the sense of value-for-money, but in the pedestrian but critical sense of following the money step by step, from budgetary authorization through the final intended recipient, in order to identify the specific steps in the process where

¹⁹ For an account of the pitfalls of performance measurement in government, see Schiavo-Campo (1999).

“leaks” occur.²⁰ Such tracking is as important in practice as it is pedestrian in concept. Still, even in the most pragmatic of efforts focusing on basic problems, there must be a vision of the eventual ECD end-point to aim for—a vision of what ECD “success” looks like. In the crucial area of fostering performance in the allocation and utilization of government financial resources, such a vision is provided by the UK system of spending reviews, summarized in Box 2.

Box 2: Policy and Expenditure Evaluation Capacity in the UK*

Policy evaluation in the UK rests on substantial earlier capacity-building, and is now well embedded in most departments and is important in defining new policy, with the Treasury playing a key role in creating demand for evaluation. The link between policy and the budget is assured by subjecting virtually all significant new expenditure programs to evaluation at three levels:

- broad long-term priorities
- spending reviews every two years
- specific programs

The heart of the exercise is in the spending reviews, carried out in four phases:

- preparation (strategic priorities, cross-cutting reviews, etc)
- departmental submission
- scrutiny and allocation of funds
- implementation and delivery

In turn, the departments are asked to address the following basic questions in their submissions:

- what are their priorities, and why are they chosen?
- what outcomes are they trying to achieve, and why?
- what has been proven to work, and why?

Literally hundreds of evaluations are carried out. This has required, among other things, a trebling of both economists and researchers in government, a huge research base, and continuous testing of both process and impact. The benefit-cost ratio of the evaluation process has been found to be positive, but a number of problems have been identified:

- no real piloting
- temptation to overclaim results
- monitoring progress is vital but not always assured
- timeliness is a problem, with politicians anxious for results, while evidence of impact takes time, and results are rarely complete and unambiguous

** Adapted from a presentation by Paul Johnson of the U.K. Treasury, at an evaluation conference in Bogotá, Colombia, September 2004.*

²⁰ An example is the Uganda public expenditure tracking study (PETS). See OED 2005.

5. A CONCLUDING WORD

Even in highly advanced countries, the evaluation system suffers from some of the same M&E pitfalls evident in developing countries—weak piloting, overstated achievements, ambiguous results, timing discontinuities. Moreover, it is worth underlining that the current systems of monitoring, evaluation and program reviews emerged in advanced countries from almost 30 years of experimentation, learning from mistakes, and requiring sustained efforts at strengthening institutional, organizational, and human capacity. As is the case of all other major functions in public sector management, building an effective capacity for monitoring and evaluation is neither easy nor quick. This is the fundamental lesson learned from the ECD experience so far.²¹ The two immediate corollaries of this lesson are the need for steady and sustained support by international donors, and the imperative of intelligent and realistic sequencing of ECD assistance.

²¹ The OED experience provides definite confirmation of the difficulties described some years ago by the OECD (1997).

BIBLIOGRAPHY

Alchian, Armen. 1977. *Economic Forces at Work*. Indianapolis: Liberty Press.

Coase, Ronald. 1988. *The Firm, the Market, and the Law*. Chicago and London: University of Chicago Press.

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

Mackay, Keith. 2004. *Two Generations of Performance Evaluation and Management System in Australia*. ECD Working Paper No. 11. Washington D.C.: Operations Evaluation Department, The World Bank.

North, Douglass C. 1991. 'Institutions'. *Journal of Economic Perspectives* 5 (1): 97-112.

Organisation for Economic Co-operation and Development (OECD). 1997. *In search of Results: Performance Management Practices*. Paris: OECD.

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

_____. 2005. *Influential Evaluations: Detailed Case Studies*. Washington D.C.: The World Bank.

Schiavo-Campo, Salvatore. 1999. 'Performance in the public sector'. *Asian Journal of Political Science* 7 (2).

_____ and Daniel Tommasi. 1999. *Managing Government Expenditure*. Manila, Philippines: Asian Development 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*.

Other Recommended Reading

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, *Monitoring and Evaluation: Some Tools, Methods and Approaches*, Second Edition, 2004.

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 OED's ECD Website:

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