



EVALUATION CAPACITY DEVELOPMENT

ECD WORKING PAPER SERIES No. 6: JANUARY 1999

A Diagnostic Guide and Action Framework

Keith Mackay



Evaluation Capacity Development: A Diagnostic Guide and Action Framework

Keith Mackay

| Contents | Page |
|---------------------------------------|------|
| I. Overview: Role of this Guide | 1 |
| II. Steps for Identifying ECD Options | 6 |
| III. Conclusions | 17 |
| Annex | 19 |
| Endnotes | 25 |

Overview: Role of this Guide

This *Guide* is intended to assist those governments and development agencies which have decided to develop a national or sectoral evaluation system. The *Guide* provides a detailed checklist of issues to be considered in developing a country's evaluation capacity. While the *Guide* is intended to be as comprehensive as possible, it will be necessary to adapt it to the circumstances of the country. The principal focus of the *Guide* is the development of a national evaluation system. In some countries such an approach might be infeasible or too expensive. In such circumstances a sectoral approach—or even one which focuses on major projects—might be judged to be more realistic, and the detailed diagnosis presented here would need to be simplified and tailored accordingly.

The *Guide* is organized into nine separate but related steps to assist in the identification of ECD options and the development of an ECD action plan (the relationship between these steps is mapped later in this *Guide*, in Figure 1). These steps involve:

1. Identification of key ministries and other bodies—these are the organizations which are key stakeholders in the government's approach to performance management. They may be currently involved in performance measurement (broadly defined, and including evaluation). Another component of this step is some understanding of the formal, stated functions of these various organizations, and of their formal interrelationships.
2. Diagnosis of the public sector environment in which individual ministries and public servants work. This diagnosis encompasses issues such as public sector ethics, incentives, and the possible existence of corruption. The public sector environment is important to the extent that it influences the way in which government ministries and departments operate, and also influences their interrelationships and performance. This includes an understanding of the rules systems and incentives which shape the behavior of public servants.
3. Development of an understanding of the factors that influence budget decisionmaking and line management decisions at the level of individual ministries. Identification of the actual—as distinct from the formal, stated—functions, and the extent of autonomy of key central and line ministries. Together, Steps 1 and 3 provide a contrast between the stated and actual processes of decisionmaking within government, both at the whole-of-government (budget) level and at the level of individual key ministries.
4. Determination of the extent of existing demand within government for measuring the performance of government activities. This entails asking whether, to what extent, and how evaluation influences budget decisionmaking and line management within individual ministries.
5. Assessment of the evaluation activities and capabilities of central and line ministries and other organizations (such as universities, research institutes and the private sector). This step relates to the extent of supply of evaluation, including an assessment of processes and systems for

making that information available—that is, the information infrastructure.

6. Consideration of the evaluation activities of multilateral and bilateral development assistance agencies in the country. These agencies can have a significant influence on governments through the loans, grants and advice they provide. This influence may be manifested via the formal evaluation requirements of the development agencies, and via the advisory services which they provide.
7. Identification of major public sector reforms in the government in recent years as well as reforms in prospect, especially those which might affect performance management and measurement. These reforms can provide opportunities for pursuing ECD.
8. Mapping of options for developing evaluation capacity.
9. Preparation of a realistic ECD action plan.

Together, the nine steps help in the identification of opportunities and limitations in pursuing ECD. Although the checklist approach systematizes the analysis of issues, the quality of judgment of those who undertake this diagnosis is crucial.

The diagnosis would require detailed consultations and discussions with government officials and probably Ministers, with development assistance agencies, academics, and with representatives of civil society including NGOs. It would require a review of key documents such as: government budget, sectoral and investment planning documents; country assistance strategies; public expenditure reviews (PERs); technical assistance loans or grants to assist public expenditure management and for other public sector management reforms. In particular, thorough, good-practice PERs would cover much of the ground under Steps 1, 3, 4 and 9, and desirably Steps 2, 5 and 8.¹

This *Guide* draws on several strands of work within the World Bank: public expenditure management, governance and public sector reform, including organizational capacity-building, and institutional assessment. Comments and suggestions on this draft are welcome. OED intends, at some future date, to apply the *Guide* to a number of case studies from developing countries.

What is ECD and why is it important?

The development of national or sectoral evaluation systems—known as *evaluation capacity development (ECD)*—is increasingly recognized as an important aid to sound governance and as a means to help achieve high levels of public sector performance. ECD ensures that evaluation findings are available to assist countries in four key areas.

First, evaluation findings can be an important input to government resource allocation—planning, decisionmaking and prioritization, particularly in the budget process.

Second, evaluation assists government managers by revealing the performance of ongoing activities at the sector, program or project levels—it is therefore a management tool which leads to learning and improvement in the future (results-based management).

Third, evaluation findings are an input to accountability mechanisms—so that managers can be held accountable for the performance of the activities which they manage, and so that governments can be held accountable for performance. The notion of accountability encompasses the recognition that economic governance and a sound public sector are central to national economic competitiveness—markets reward countries able to manage and screen public expenditures, and evaluation offers a tool to help do that. As Robert Picciotto has said:

The quality of public expenditures is as important as fiscal balance. Large scale, low return projects (which tend to flourish where evaluation is weak) tend to aggravate financial crises.²

A fourth use of evaluation findings is in demonstrating the extent to which development activities have been successful. This is proving to be increasingly important for countries in attracting external resources, particularly given the pressures on international development assistance agencies to channel their assistance to countries where past development efforts have been successful. Moreover, the increasing emphasis by development assistance agencies on a whole-of-government approach to

development increases the premium on having country-wide measures of performance available.

The concept of performance is central. It encompasses the efficiency of a sector, project or activity—the ability to undertake an activity at the minimum cost possible. It also includes the effectiveness of an activity—whether the activity is achieving the objectives which were set for it. Performance measurement is an essential activity because it provides an opportunity and a framework for asking fundamental questions such as: ‘What are you trying to achieve?’; ‘What does “success” look like?’; ‘How will you know if or when you’ve achieved it?’

There are many different types of evaluation tool and they can be used in a variety of ways. Although these tools are related, the different terminologies employed by evaluation practitioners can lead to confusion. The tools all address performance measurement: ongoing monitoring and performance indicators; project and program evaluation—ex ante, ongoing/formative and ex post/summative; performance (or value-for-money) audits; financial auditing. Each of these tools provides information on the performance of an activity, and each can be used in a variety of contexts. This broad spectrum of performance measurement activities is also known by other generic labels, such as monitoring & evaluation. The terms ‘performance measurement’ and ‘evaluation’ are used interchangeably in this *Guide* in a generic, shorthand sense to encompass these various terms and concepts.

The support which the development of evaluation capacity can make to broader governance, institutional development and public sector reform is often not fully appreciated. There are links and commonalities in:

- budgetary financial management systems, and financial reporting;
- intergovernmental fiscal relations, and the extent to which they encompass a focus on government performance;
- accountability institutions, and audit organizations;
- commercialization, and the private sector delivery of public services—that is, where governments exit from public service delivery. Success in achieving this requires a clear understanding of program objectives, and ex ante, ongoing and ex post assessments of performance;

- setting explicit customer service standards by service delivery agencies, and monitoring the extent to which these are achieved;
- civil service reform which includes devolution to managers of increased levels of responsibility and autonomy (the ‘new public management’);
- civil service reform which focuses on personnel performance, management and appraisal—recognizing that individual performance is reflected to some extent in project or program performance;
- the quality of civil service policy advice, and whether that advice encompasses whatever evaluation findings exist or can be commissioned;
- participation and the ‘voice’ of civil society—this draws on the views and expectations of ordinary citizens concerning the performance of government activities. It is related to concepts and tools of participatory evaluation and client surveys. It relates also to the role and contribution of non-government organizations (NGOs); and
- anti-corruption efforts—there are many commonalities and links with the development of evaluation systems, by way of improved financial management systems, timely performance reporting, strengthening of watchdog agencies, and greater transparency in policymaking and implementation.

Country experience

A number of developed countries have sound evaluation systems, and their experience illustrates the potential links between national evaluation capacity and good governance. And experience also reflects the opportunities and difficulties in achieving cultural change in a government—winning hearts and minds is a slow business. Experience illustrates the different dimensions which must be developed to achieve a robust evaluation system: these include demand, supply, and information infrastructure. A growing number of developing countries are now giving increased priority to building sound evaluation systems, for example, Chile, Zimbabwe, the Philippines, Colombia, and Indonesia.

The main *precondition* for the development of a national evaluation system is country demand—an evaluation system cannot be effectively imposed on an unwilling government.

There are also particular risks if the impetus for an evaluation system is driven by external development agencies; this is not to say that development agencies cannot take the lead in promoting the merits of evaluation systems to countries, but that until countries accept the strength of such arguments, or themselves conclude that evaluation has much to offer, an evaluation system is unlikely to be sustainable.

Experience tells us that the main barriers to building evaluation systems in developing countries have been: lack of genuine demand and ownership in countries; lack of a modern culture of fact-based accountability, often related to issues of ethics or corruption; lack of evaluation, accounting, or auditing skills; poor quality of financial and other performance information, and of accounting/auditing standards and systems; lack of evaluation feedback mechanisms into decisionmaking processes; and the lack of critical mass needed to develop sustain-

able evaluation systems.³ This leads to the question of whether a government or ministry needs some minimum level of overall capability before an evaluation system can realistically be contemplated, and if so, what is that level?

Another lesson from experience is that the development of an evaluation system should not be viewed as a stand-alone activity—it would be unrealistic to attempt to simply ‘bolt on’ an evaluation system to an existing structure of governance if the institutional framework and incentives cannot support it. If the framework and incentives are currently insufficient, this is a strong argument for ensuring that development of an evaluation system is part of a broader initiative to develop governance; this comprehensive approach recognizes the strong synergies between performance measurement/evaluation and performance management.

II. Steps for Identifying ECD Options

An analytical framework for applying the *Guide* is provided in Figure 1, which sets out the nine steps for identifying ECD options and preparing an action plan. Diagnostic checklists for each step are provided in Annex A.

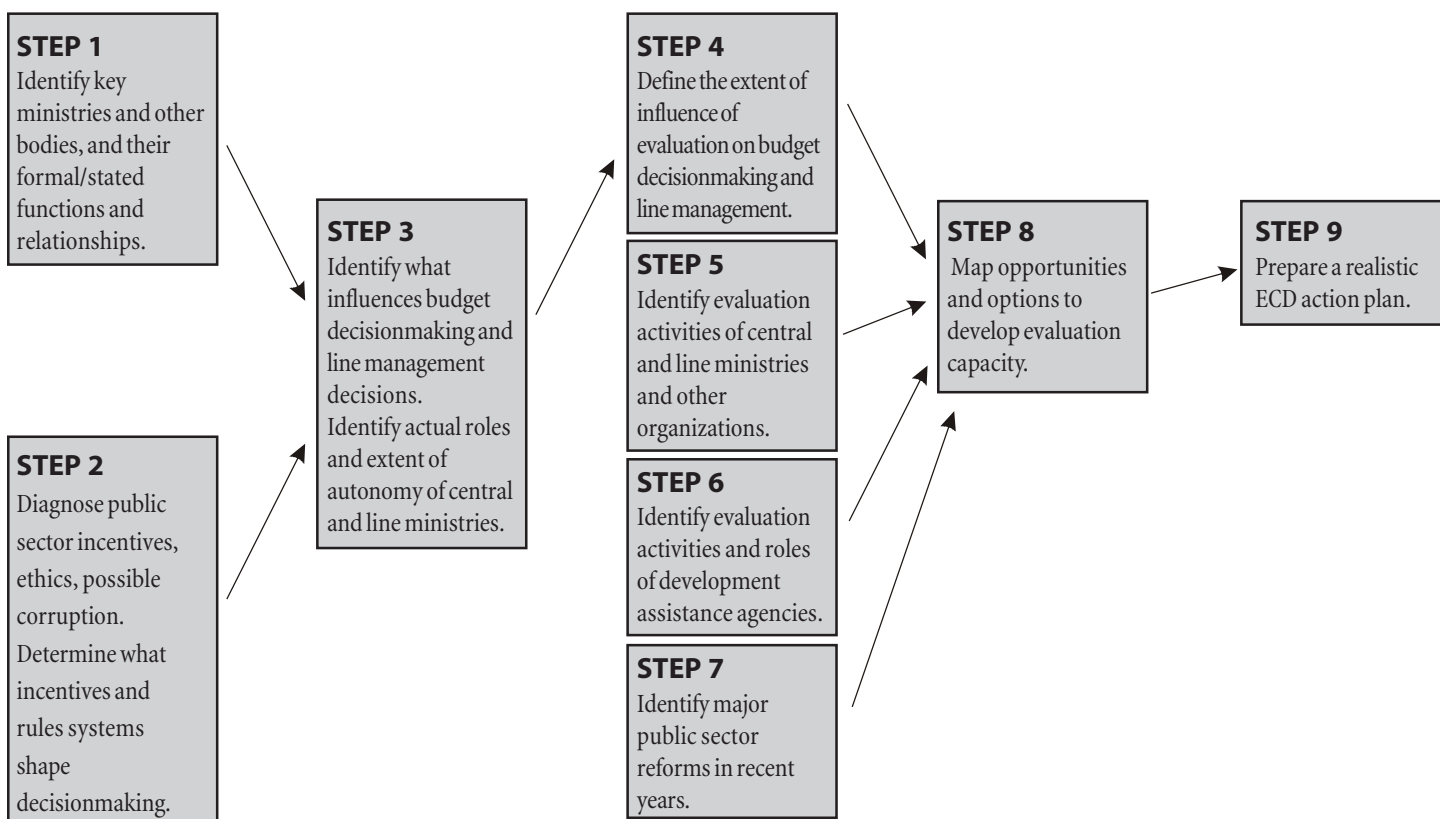
Step 1: Identification of key ministries and other bodies, and their formal, stated functions and relationships

Key organizations are those which are important to performance management within the government—that is, they have a place in the government’s resource allocation decisions, or in the ongoing management of sectors, programs and major projects, or in accountability relationships. They have responsibilities in one or more of the following areas: public expenditure (budget) management;⁴ strategic planning; determination of public investment programs; policy formulation; policy advice; line management; monitoring and evaluation; auditing.

These organizations include central ministries such as finance, planning, Offices of the President/Prime Minister/cabinet, the national audit office, parliament, and large line ministries such as education, health, and agriculture, development assistance agencies, powerful NGOs and the private sector, where this has a formal responsibility for advising government on expenditure levels, priorities or performance (via, for example, deliberation councils).

Central ministries have a powerful influence on the decisions government takes, and on the activities and behavior of line ministries. Large line ministries are particularly important to ECD if they can be persuaded to invest time and effort in

Figure 1: Steps for Identifying ECD Options



developing their evaluation capacity. This will provide a powerful demonstration effect to the rest of the bureaucracy.

It is also helpful to delineate the formal roles of elected officials—especially Ministers—and of bodies such as parliament, in relation to the civil service. It may be appropriate to consider the tasks and functions of other levels of government, especially if the country has a federal system of government.

While consideration of formal responsibilities is relevant to identifying options for ECD, it is only a starting point. The actual responsibilities may well diverge from stated ones—this could arise for a number of reasons, particularly where actual capabilities do not match formal functions and responsibilities (these are considered further under Step 3).

One important and illustrative area where formal and actual roles can diverge markedly is in public expenditure management. In many countries the formal links and coordination between budget planning, resource allocation and implementation either do not exist in reality, or do so only in an attenuated form. Public sector performance can be undermined by weaknesses in the budget system such as:⁵

- poor planning;
- poor expenditure control;
- inadequate funding of operations and maintenance;
- little relationship between budget as formulated and budget as executed;
- unreliability in the flow of budgeted funds to ministries and lower levels of government;
- poor management of external aid;
- inadequate reporting of performance.

Such weaknesses severely undercut planning and management both at the whole-of-government level and at the ministry and agency levels. In such a public sector environment it is difficult to complete budgeted activities and projects in a timely and satisfactory manner. These significant constraints on management's ability to manage, typically generate a working climate in which performance and its measurement are not regarded as important. These considerations lead to Step 2.

Step 2: Diagnosis of the public sector environment, including incentives and ethics

An environment in which it is difficult for managers to perform to high standards and to perform consistently is hostile to evaluation. Managers can do little more than focus on day-to-day management tasks, narrowly defined. Managers have little incentive to see themselves as accountable for performance where they do not have some surety of the resources available to them, or where they do not have substantial control over the outputs of their activities. In this environment, evaluation is understandably seen by managers as irrelevant, as potentially (and probably) unfair to them, and as a threat rather than an aid.

In an ideal world, a civil service culture would embody a set of values strongly supportive of a performance orientation. These values would include a strong desire to serve the public—to help the clients and beneficiaries of government programs—and a focus on performance and improving it. A key set of factors which determine the presence or absence of such a 'performance culture' are the norms, values, expectations, incentives, rules, flexibilities and restraints facing civil servants.

An institutional assessment of the incentives and 'rules of the game' within a civil service would encompass analysis of:

- formal rules of behavior, informal norms and their enforcement;
- the extent of autonomy and authority of managers to control the resources for which they are responsible;⁶
- the sanctions that exist for poor managerial performance;
- the reputation of the civil service for competence, skill, capabilities, impartiality, absence of corruption, extent of merit-based recruitment and promotion, existence of esprit de corps;
- the basis on which management decisions are actually made—a concern to 'do good', monetary rewards, factors that affect personal wealth or that of relations or colleagues, authority, or other intangible rewards such as status or recognition;

-
- civil service salaries in relation to their private sector counterparts (good remuneration helps to attract talent to the civil service and reduces pressures to be corrupt);
 - strength of accountability institutions—such as the national audit office, parliament, the judiciary, audit and procurement rules, financial regulations;
 - extent to which the ‘voice’ of civil society is important—extent to which NGOs, the media and public opinion highlight problems with public sector performance;
 - competitive pressures within and outside the public sector which affect public sector performance, for example, via competitive tendering and contracting, specification of performance contracts for senior management.

Some of these issues would need to be assessed very much on an ‘impressionistic’ basis—relying, for example, on the judgments and experience of country officials, development assistance agency staff and NGOs. Past country analysis in areas of public sector management would be an invaluable source of information and analysis.

This institutional assessment is designed to answer the questions: ‘Does a performance culture exist?’; and ‘What are the roadblocks to achieving it?’

Step 3: How are budget and line management decisions taken, and what are the actual roles of ministries?

The public sector environment is a powerful determinant of the way in which the various arms of the government operate. It influences their interrelationships and performance. It interacts with the formal, stated roles of individual ministries to help determine their *actual* roles and functions. There may well exist a wide gulf between stated and actual roles.

There may be various reasons for such a gulf. The actual capabilities of individual ministries may be inadequate or inappropriate. There may also exist structural flaws in the systems and processes of government, especially in budget processes. These flaws both undercut formal roles and have a strong influence on incentives.

The nature of the budget system is particularly important. A system with fiscal discipline and strong links between strategic priorities and resource allocation provides an enabling environment for public sector performance. Conversely, a system where the sources and uses of funds are balkanized, or where government decisionmaking is ad hoc, is likely to impede rational policymaking, budget planning and implementation—with negative consequences on actual performance.⁷ A sound budget system requires full information on the costs (immediate and in future years) of budget decisions, a clear understanding of the probable implications of alternative policy decisions, and the ability to learn from the experience of policies, programs and projects; having evaluation findings available is thus a key aid to budget decisionmaking.

For evaluation findings to be most effective it is necessary that civil servants draw on them when providing policy advice. Line ministries can use evaluation findings in formulating detailed policy proposals for government’s consideration in the budget process; central ministries (such as finance, and planning) can use such findings in commenting on the merits of policy proposals and in reaching judgments on them.

The budget system has a strong influence on the performance of line ministries. Ideally, the system will provide clarity concerning the government’s policy priorities and the actual allocation of resources available, and will provide some surety over future flows of funding to enable timely completion of investment projects and to assist the management of ongoing activities. It would be incorrect to view the budget system as solely being a top-down process, however—line ministers and ministries have an important task in helping to shape policy and in providing cost and policy information to the finance ministry in coordinating the budget process.

The formal rules and guidelines (which finance and planning ministries require line ministries to follow) help to determine how well the budget system actually operates. These instructions set out the procedures for submitting policy proposals, the types of cost information and analysis which they should contain, their timing, information flows on actual expenditures and commitments, and so on. But for formal rules to be effective, there have to be appropriate incentives to induce line Ministers and ministries to comply—otherwise ‘playing by the rules’ might imply that formal instructions will be ignored or only paid lip service.

For most formal rules to work, then, it is necessary to have in place appropriate incentives and capabilities, together with some central monitoring, support and enforcement. The work of accountability organizations is important here. These include the finance ministry and national audit office.

The capabilities and functions of key line ministries need to be assessed—this is the counterpart of the identification of key ministries and their functions under Step 1. These key ministries have oversight of important sectors such as education, health, agriculture and industry. To what extent do these ministries have control over the resources budgeted to them? What is the nature of their responsibilities and accountabilities? What are the incentives in play? Do capabilities match responsibilities? How sound are their planning, resource allocation, management and monitoring of performance?— these issues are considered in more depth in Steps 4 and 5.

Step 4: Extent of influence of evaluation on budget and line management decisions?

The objective in building evaluation capacity is to help ensure that performance measurement becomes an input into budget and line management decisionmaking, and to satisfy accountability requirements. It is important to understand the extent to which this already occurs, and how it occurs. Step 4 therefore entails assessing the extent of existing demand within government for measuring the performance of its activities and for using this information to the best effect.

The various performance measurement tools which might be used in decisionmaking include: project and program evaluation—ex ante, ongoing/process/formative and ex post/summative; performance audits (also known as value-for-money audits or efficiency audits); ongoing monitoring and performance indicators; and surveys of clients or citizens. Other types of information which aid decisionmaking and can be used to provide measures of performance include: cost/expenditure data; financial/compliance audits; and national statistical data.

Various players are part of the process of budget decision-making, and the uses of evaluation and related information which each make should be separately identified:⁸

- Line ministries and line Ministers, in developing policy options for government's consideration;
- Central ministries, particularly the finance and planning ministries, in framing development plans and providing policy advice to government; and
- Government/cabinet in its actual decisionmaking—that is, in its prioritization and resource allocation.

In principle, information on program/project performance can help to inform most types of budget decisions. In practice, it is often used only in certain contexts, such as decisions on investment projects, or decisions on whether and how much to borrow from development assistance agencies.

Project and program management is typically the responsibility of line ministries and outrider agencies. Evaluation can assist the management of ongoing activities through: identifying emerging operational problems; providing benchmark comparisons with other organizations or over time; and by revealing the cost-effectiveness of different activities. Perhaps high-quality evaluation is being used in particular sectors or major projects, and these could provide examples of good practice which could be used as 'demonstration models' for other parts of the government (see Step 8).

Another use of evaluation is for accountability purposes, in reporting performance to parliament, the finance and planning ministries, and the public, and development assistance agencies. Performance information might appear in budget documents—reporting on outputs and outcomes, together with expenditures and forward estimates, and in public expenditure reviews, country assistance strategies or implementation completion reports.

In making this assessment of demand it would be helpful to include considering whether any central or line ministry could be legitimately classified as a 'champion' of evaluation. When mapping opportunities and options for developing evaluation capacity within the government as a whole (Step 8), such support could be invaluable.

Step 5: What are the evaluation activities and capabilities of ministries and other organizations?

This step relates to the extent of supply of evaluation, including an assessment of processes and systems for making that information available—that is, the information infrastructure. It includes an assessment of the evaluation capabilities and activities of central and line ministries and other organizations such as universities and research institutes.

Capabilities depend on the skills and experience of staff with a variety of performance measurement tools, the existence and capacity of specialist M&E or audit units, and so on.⁹ Such staff may possess formal training in social science research skills, program evaluation, cost-benefit analysis, statistical analysis, audit, or they may be generalists who have acquired relevant skills and experience on-the-job. The evaluation experience and understanding of staff can vary widely. When conducting a country diagnosis, it is necessary to identify the existence of specialist units, to discover the number of staff they contain, and to gain some knowledge of their skill mix and experience.¹⁰ Even rapid review of evaluation reports and information produced by these units can provide some understanding of their capabilities.

Many evaluation units are set up to assist with a major project, especially for donor-funded investments, and the unit may not survive beyond the project's completion. While the staff may then disperse to other civil service activities, their evaluation capacities and training might be drawn on—this could be useful when identifying options for developing evaluation capacity (Step 8).

Specialist agencies such as national audit offices and national statistical agencies may have staff with evaluation skills. Statistical agencies collect socioeconomic studies and may conduct household surveys; these provide useful data for evaluations.

Universities and research institutions may be additional sources of evaluation skills, and may also offer (or have the capacity to offer) training in a range of skills and evaluative techniques. Local consultants are another source of evaluation skills.

One way of making a rapid—but admittedly only partial—assessment of the strength of these various sources of supply is to determine which of the following have prepared evaluation studies in the past (and how many studies) line ministries, central ministries, other government agencies; universities/research institutions, local consultants, expatriate consultants, and development agencies. The quality of these studies also provides necessary insights into existing capabilities—and this helps in identifying opportunities and options to develop evaluation capacity.

Additional sources of performance data are invaluable and often essential inputs to evaluations and to the analysis of policies, sectors, programs and projects. These include financial data on costs (actual expenditures, budget allocations, and forward estimates) and data on outputs—such as number of hospital patients or number of schoolchildren. Some governments have difficulty in collecting even basic cost and output data; others have difficulty in producing data which are reliable, timely, or with the necessary degree of disaggregation. Some assessment of the availability and quality of these basic cost and output data should also be made. If the basic cost and output data are inadequate, then it will be necessary to develop them before attempting to develop evaluation capacity.¹¹

Evaluation is often viewed as a stand-alone, technical activity. But it is important that there be close links between decisionmakers and evaluation units. These links relate to evaluation planning and the determination of evaluation priorities; they also include the means by which evaluation findings are fed back to users, clients and other stakeholders. There is no point in undertaking evaluations and collecting other performance information if the data are not used. And if the outputs from evaluation capacity are not used, the capacity will not be sustainable.

Thus an important part of diagnosing the capabilities and activities of ministries is to assess the ways in which evaluation priorities are set, by whom, and for what purposes. It is illuminating to know the extent to which senior ministry management is involved in determining evaluation priorities, and whether evaluation findings and other performance information are regularly sent to them. Who determines the frequency and

coverage of evaluations? Are any evaluations or policy reviews commissioned by cabinet or by central ministries?

It is important to know what use is made of evaluation findings, and by whom. Who are the clients and stakeholders? Do they include line ministry senior management, policy analysts, line managers and operational staff, and central ministry analysts? Ideally, each group would have at least some understanding of evaluation techniques (including their strengths and weaknesses) and evaluation's potential usefulness.

Step 6: What are the evaluation activities and roles of development assistance agencies?

Development assistance agencies have a long history of evaluating the loans or grants that they make, at both the ex ante and ex post stage. Project implementation is monitored through the collection of performance indicators. And evaluations and reviews are also conducted at the country and sector levels, and for thematic issues such as gender and other human resources issues, public sector management, and so on. These activities also provide support and information to developing country governments concerning the merits or dangers of proposed and actual investments.

Some major development projects involve the creation of evaluation units for the duration of the project. One danger, however, is that scarce country capacity will be diverted to satisfy donor requirements for the evaluation of development assistance activities, which might not necessarily align with the areas of greatest potential benefit from evaluation. Donors might hinder the development of local capacity by excessive, conflicting or multiple donor requirements for evaluation. Even where local capacity is built up in relation to discrete projects, their completion may well lead to the dispersal of those skills.

Donors can also help develop evaluation capacity by means of technical assistance for this purpose or for related activities such as public expenditure management. The World Bank, for example, has provided ECD assistance to a number of countries including the Philippines, Zimbabwe, China, Colombia, Indonesia and Argentina. Other donors have also provided assistance. The types of assistance have included advice, expertise, organiza-

tion of seminars, training, loans and grants, identification of qualified consultants and the preparation of guidance material including case study examples. In the past, much of this assistance has been provided for the limited duration of a loan or grant—however, donors have increasingly come to understand that the development of evaluation capacity and other organizational capacities is a long-term proposition. Sustained support and commitment is necessary, and the World Bank is developing some new lending instruments to provide this.

Step 6 entails mapping those donor activities which may have contributed to the development of in-country capabilities in recent years—within ministries (such as the creation or support of evaluation units), the expansion of local skills through the provision of training and the hiring of local consultants, as well as by giving support for universities and research institutions. Much of this mapping would have taken place under Step 5; but it is worthwhile identifying the special function of donors both as a key to understanding the development of the government's capacity, and to help clarify opportunities and options for further development (Step 8).

Step 7: What major public sector reforms have been implemented in recent years or are likely in future?

This step entails the identification of major public sector reforms in the government in recent years as well as reforms in prospect, especially those which might affect performance management and measurement, and which could provide opportunities for ECD. This would include reforms in: budgetary systems; intergovernmental fiscal relations including decentralization; commercialization and private sector delivery of public services; customer/client service standards; civil service reform;¹² quality of civil service policy advice; participation and the 'voice' of civil society; anti-corruption efforts; and accountability institutions including the professionalization of national audit offices and accounting standards. A list of possible types of public sector reform is shown in Figure 2.

These types of reform can have commonalities and synergies with ECD: there may be economies of scale from undertaking such reforms jointly with ECD. Synergies can arise because efforts to develop evaluation capacity can enhance reforms in

other areas and increase their prospects for success. And this can be a two-way traffic. For example, ECD would typically include a focus on performance reporting, the strengthening of watchdog agencies and the encouragement of greater transparency in policymaking and implementation; these would provide powerful support for anti-corruption initiatives.

A review of key documents and discussion with government officials and development agency officials should confirm the existence of such reforms, or the intention or possibility of pursuing them.

Step 8: What opportunities and options exist to pursue ECD?

Steps 1-7 are designed to facilitate the collection of a solid base of information to help identify ECD options. The steps involve the collection of information on:

- stated and actual roles of key central and line ministries;
- incentives and other influences on decisionmaking, at the budget (whole-of-government) and line management levels;

Figure 2: Related Areas of Public Sector Management Reform

| Public expenditure management | Results-based public management and performance measurement tools | Civil service reform | Anti-corruption | 'Voice' (for example, citizen surveys) | 'Exit' (for example, private sector delivery of public services) | Financial management/ accounting |
|--|---|---|----------------------------|--|--|--|
| macroeconomic/ fiscal balance | specification of objectives | merit-based personnel policies (recruitment, promotion) | causes/ consequences | citizen charters | competitive tendering and contracting | accounting standards |
| national resource allocation | program logic, results chain, logframe | salaries | measurement | client surveys | | financial management information systems and links to budget process |
| sectoral resource allocation | performance budgeting | skills/training | role of watchdog agencies | involvement of NGOs | | cash/funds management |
| policy analysis/ review | performance contracting | ethics/ethos | civil service reform | | | internal control and audit |
| management/ implementation of resources | performance/ efficiency auditing | quality of policy advice | legal/judicial reform | | | external financial audit |
| agency management | benchmarking | | role of civil society | | | efficiency/ performance audit |
| project management | performance information/ monitoring | | national integrity systems | | | procurement |
| budget system architecture ¹³ | evaluation (country/ thematic/ sectoral; programs; projects) | | | | | |
| organizational and institutional diagnosis | performance reporting | | | | | |

- relationships between central and line ministries, and the extent of autonomy of ministries;
- extent to which evaluation influences budget decision-making and line management;
- evaluation activities of central and line ministries, and of other organizations;
- evaluation activities of development assistance agencies; and
- public sector reforms in recent years and in prospect.

This information could be used as the basis for undertaking a SWOT analysis for identifying ECD options in a country—strengths, weaknesses, opportunities and threats.

The approach outlined here is similar to a SWOT analysis, but focuses on: the demand for ECD; the supply of needed skills and capabilities; and issues to consider in mapping out an ECD strategy, including its opportunities and options.

Demand

As noted in Part I, *the main precondition for ECD is country demand*—an evaluation system imposed on an unwilling government is highly unlikely to be sustainable. So it is necessary to understand where the impetus for ECD in the country is coming from, as well as the overall extent of demand. Are there domestic sources of demand, or is the government interested in developing its evaluation capacity solely or largely because of advice or pressure from development agencies such as the World Bank? How committed is the government to ECD? Is there a domestic ‘champion’ ministry—or potential champion—of ECD, and is that champion a key decisionmaker in public sector reform? On which of the four objectives of ECD is the country demand focused (ECD as an aid to resource allocation? To assist line management? For accountability purposes? To demonstrate development effectiveness to donors?) Each of these objectives is likely to have a corresponding—and different—source of demand and each is likely to require different types of evaluation tool to satisfy that demand.

Answers to these questions will be the key in determining which ECD options are feasible. When developing a country’s evaluation capacity, typically demand has to be carefully fostered. Champion ministers, senior officials and ministries usually have to develop some consensus that ECD is a worthwhile objective. The implementation in a government of any reform initiative entails substantial

costs and effort, and it can only be expected to be embraced if the potential benefits are regarded as sufficiently great to warrant this effort. It would be highly unrealistic to expect ministers and officials to view the desirability of ECD as a self-evident truth. Thus it is likely to be necessary, as part of the ECD effort, to develop an explicit strategy for winning hearts and minds, and to accept that this is a slow business requiring sustained effort and commitment.

In assessing demand it is also important to identify possible bureaucratic ‘roadblocks’—especially hostile ministries. Because existing incentives often work against evaluation, it is important to identify who stands to lose from having information available on the performance of government activities. Line ministries may perceive evaluation findings as a threat because of fears that these will be used by finance or planning ministries to reduce budget appropriations, or for intensifying accountability. Central ministries not involved in ECD activities might well regard ECD as a threat to their power in the budget and planning processes. Thus it is helpful to identify the likely winners and losers from ECD.

For ECD to be a realistic proposition, there needs to be a supportive culture or set of values—or at least the possibility of developing such a culture. Institutional change takes time and depends on learning and leadership. If ECD is successful, it will have a powerful influence on the incentive structure within the government—reflecting the objective of creating a performance culture. Seen from this perspective, ECD can be a vehicle for achieving fundamental cultural change.

Supply

The supply of evaluation skills and capabilities is an important consideration in mapping out ECD options, but it is argued here that supply is not as crucial as demand. The supply side encompasses the availability of skills and resources needed to conduct evaluation activities. It includes the skills and resources necessary to provide socioeconomic statistics and to undertake auditing and accounting functions. These provide basic information on which evaluation depends—they are necessary for the estimation of data on the inputs to and outputs from government programs and projects.

Various dimensions of supply include staff levels, financial resources, professional skills, methods and standards, and information flows. Demand helps to ensure the availability of

sufficient numbers of staff and financial resources, and to ensure that information infrastructures can be set up to provide needed information flows. It is critically important to develop the supply of skills to match growing demand.

Simple evaluation tools and techniques such as ongoing monitoring and performance indicators provide a useful starting point in addressing skill levels for ECD. The supply of more sophisticated skills may require a long-term effort involving training organizations—the diagnosis in Steps 5 and 6 will be relevant here. Training in policy-advising skills is one means of providing civil servants with some understanding of evaluation tools and their uses.

The need for ECD to be sustainable means that reliance on domestic skills—or their development—is necessary, rather than reliance on expatriate consultants or experts from development agencies. Training might be provided by universities and research institutions (in-country or regional), the World Bank and other development agencies, and government training bodies such as the civil service commission or the finance ministry.

ECD Options

A menu of a ECD options is presented in Figure 3—this maps options in four corner situations: where overall demand and supply are strong or weak.¹⁴

The menu should not be interpreted as providing a blueprint for action: it provides only a summary of a range of possible options. The feasible options in any country will depend closely on such factors as local circumstances and opportunities, the different actors involved, institutional frameworks, the strength of demand and supply in different ministries and sectors, and evaluation skills.

There is no optimal sequence for developing evaluation capacity. But a first step is to ensure sound country capacity for collecting and analyzing data on costs and simple outputs and service quality indicators. These are also basic elements of a public expenditure management system; however, once that capacity is in place, the next steps have to be chosen opportunistically, according to the realities of demand and supply in-country. ECD is in one sense an incremental process, but it may also be path-dependent. It certainly requires a long-term, multi-year framework or plan, necessitating the identification of clear targets to which the government is committed.

Realism is an important component of an ECD action plan: experience from developed countries, where civil service capabilities are typically strong, shows that ECD takes at least a decade to achieve. Developing countries are likely to require longer; this indicates the importance of sustained commitment and perseverance, on the part both of governments and of development agencies.

A basic issue is how comprehensive ECD needs to be. One viewpoint holds that it has to be comprehensive—a ‘big bang’ approach—to be successful. This would stress the synergies and economies of scale involved in pursuing all three possible objectives of ECD (aid to resource allocation decisionmaking; aid to line management; accountability), particularly where a whole-of-government approach is involved. Such an approach has the potential to generate sufficient momentum to sustain progress in all ministries; even ministries which suffer some external setbacks can be induced to keep up with their peers if there is sufficient government-wide pressure. But such an approach can be unrealistic, particularly in developing countries where basic civil service capabilities are often uneven or weak.

An alternative view is to pursue ECD modestly, with a focus on those ministries, sectors or major projects where there is some demand for evaluation.¹⁵ These could be developed on a trial basis to help demonstrate their feasibility. Opportunities could subsequently be created to spread successful approaches to other sectors. Evaluation systems previously developed for particular sectors or major projects could well provide islands of skills and capabilities to assist other parts of the public sector; these capabilities would most probably have been developed through carrying out project appraisals and developing performance indicators.

A variant of this more modest approach would develop basic evaluation tools—such as performance indicators—as a requirement for all ministries, and to seek to develop the application and use of more sophisticated evaluation tools later.

These incremental approaches have much to commend them; but it would be necessary to guard against the danger that some initiatives might be too small to achieve the critical mass necessary for sustainability. A particular danger is that a trial approach might exist only because of the support of one or two committed senior officials; if these were to move elsewhere,

Figure 3: Menu of Some ECD Options

| | | Demand | |
|---------------|---------------|--|--|
| | | Strong | Weak |
| Supply | Strong | <ul style="list-style-type: none"> • support evaluation of policies, programs and projects • establish links between evaluation, strategic planning, resource allocation and budgets • use expert commissions to evaluate policies • strengthen evaluation in the legislature (parliament) • disseminate evaluation results to the public • organize and systematize the evaluation function | <ul style="list-style-type: none"> • disseminate evaluation methods and practices • support ongoing evaluation of programs and projects • participate in evaluations done by external funding agencies • support professional development in evaluation • support research institutions in carrying out evaluation • link evaluation into ongoing public sector management (governance) reforms • link evaluation into public expenditure management (budget system) reform • build an evaluation network within government and/or regionally • organize ECD seminars for senior officials, to help build consensus • mandate evaluation by decrees, regulations |
| | Weak | <ul style="list-style-type: none"> • support financial and information (monitoring) systems • disseminate lessons of experience and evaluation best practice • set up commissions to evaluate important projects or programs • train and use private sector organizations in evaluation • support university training in evaluation, including via short courses or seminars • provide technical assistance to government agencies • build an evaluation network within government • develop twinning arrangements with governments in the region • promote evaluation by non-government agencies • introduce evaluation standards • promote study tours to other governments • establish twinning arrangements with other governments | <ul style="list-style-type: none"> • create a ministerial decree or cabinet requirement for systematic evaluation • strengthen audit and accounting • carry out joint evaluations with funding agencies • disseminate national and international lessons of experience • support evaluation training in educational institutions • promote cross-country cooperation in evaluation • raise awareness among decisionmakers • encourage NGOs to undertake citizen surveys • encourage the media to report on government performance |

there could be insufficient impetus remaining to continue the trial—thus the issue of sustainability should be addressed explicitly, together with an analysis of risks.

The prospects for sustainability would be greater with other, related public sector reforms underway in performance management, particularly in public expenditure management. And once again, synergies operate in both directions: budget reform on its own is unlikely to be either effective or rapid—it should be part of an integrated strategy for achieving high levels of public sector performance. Thus the World Bank has argued that:

When output measurement is strengthened and credible controls over inputs are instituted, managers can be granted more operational flexibility in exchange for greater accountability for results... Well-functioning policymaking mechanisms make transparent the costs of competing policies and encourage debate and consultation among all stakeholders.¹⁶

While it would be unrealistic to be over-prescriptive about what constitutes desirable approaches to ECD, the need for a clear diagnosis of the problems and for a plan of action is undisputed:

Evidence shows that there are many more examples of failed reform efforts where the approach has not been comprehensive, and the initiative has addressed only parts of the problem. Where comprehensive reform efforts have failed, it has mainly been because of poor analysis of the underlying problem. But comprehensive reform will involve a longer time horizon than will component reforms.

Much of the skepticism about comprehensiveness might lie in a misconception of the term. Comprehensiveness is not about trying to do everything at once. Rather, it is about taking a holistic approach to diagnosing the problems, understanding all the interlinkages and evaluating the institutional impediments to performance, and then finding the most appropriate entry point to launch a phased reform process. Phasing can be fast or slow, depending on country conditions, and could eventually expand to become comprehensive.¹⁷

A key aspect of ECD efforts is the continuing support of development assistance agencies—their funding, advice and general support—reflecting the long-term nature of ECD. Unfortunately, most of the technical assistance provided by development agencies has been in the form of 2-3 year projects. For these reasons, the World Bank is currently examining the feasibility of longer-term loan instruments—such as Adaptable Program Lending Loans (APLs) and Public Expenditure Reform Loans (PERLs)—to support governance reforms. One factor to consider in designing an ECD strategy is the likely duration and the nature of support from development agencies. Can the government develop its evaluation capacity on its own, or with only start-up involvement from development agencies? Chile provides one example of a country whose ECD initiative was essentially home-grown.¹⁸

Step 9: Preparation of a realistic ECD action plan

Preparation of an action plan—however ambitious or modest it might be—provides a useful focus for ECD efforts. An action plan provides a basis for discussion and debate of ECD options by stakeholders, and also provides a yardstick by which a government's ECD efforts can be assessed in the future. There is no single “correct” way to prepare an action plan, but a plan would desirably include explicit focus on:

- demand—how to build demand and create consensus, and where to build demand;
- supply—how to broaden and deepen evaluation skills;
- level of ECD envisaged—at the national or sector levels, or for selected major projects;
- evaluation infrastructure—ways of planning and managing evaluations, and for ensuring that results are provided to users;
- needed support from development assistance agencies;
- timelines, sequencing and speed of implementation;
- sustainability of the new evaluation capacity; and finally,
- list of main ECD initiatives.

A summary list of some issues to consider under each of these headings is contained in Annex A.

III. Conclusions

At the start of this *Guide* it was noted that the diagnostic checklist is intended to assist those governments and development agencies which have decided to develop a national or sectoral evaluation system. A precursor to such a decision is the judgment that ECD is a realistic proposition within that country.

The performance of a public sector is determined by factors including the relationships between: stated roles and functions; budget process—especially planning, budgeting and implementation; incentives and the extent of a performance orientation or culture; the role of accountability organizations; and the actual capabilities of individual ministries and of the civil service as a whole.

ECD is a long and complex business, requiring substantial and sustained country commitment. Some countries might become less enthusiastic about embarking on this reform once they realize the challenges that lie ahead. That would be unfortunate because evaluation—the regular measurement of public sector performance—has the potential to help create a performance culture and is an essential institution of good governance.

The importance of ECD to public sector performance and to performance management suggests that it should be regarded as a central component of the World Bank's advice to developing countries on governance.

Annex A

Checklist for Step 1: Identify Key Government and Other Organizations and Their Stated Functions

- What are the formal responsibilities of Ministers and the parliament in major resource allocation decisions, in the management of sectors, programs or major projects, in accountability issues?
- Which central and line ministries such as finance, planning, offices of the president/prime minister/cabinet, the national audit office, and large line ministries (such as education, health, and agriculture) have a major, formal responsibility in these issues?
- What is the formal role of development assistance agencies in resource allocation decisions?
- What are the formal, stated links between policymaking, planning and budgeting?
- What formal mechanisms exist to achieve necessary trade-offs between aggregate fiscal discipline and strategic resource allocation priorities?
- What formal processes and requirements are there for the monitoring and reporting of budget allocations, actual expenditures, outputs and outcomes?
- Are there any formal requirements and agreements for the undertaking and reporting of performance monitoring and evaluation?¹⁹
- Which ministries have responsibility for monitoring and regulating the institutional arrangements in the civil service?

Checklist for Step 2: Diagnose the Public Sector Environment, Including Institutional Behavior

- How much autonomy and authority do public sector managers have?
- What sanctions exist for poor managerial performance?
- How capable and skilled is the civil service in areas such as project and program management and oversight, policy analysis, and budget management?

- What is the civil service's reputation for impartiality, resistance to corruption, extent of merit-based recruitment and promotion, existence of loyalty to a profession or group?
- How well motivated are civil servants?
- What is believed to provide the motivation for management decisions—Compliance with formal rules? Concern to 'do good'? Monetary rewards? Factors that affect personal wealth or that of relations or colleagues? Authority? Intangible rewards such as status or recognition?
- What is believed to provide the motivation for Ministers and the government?
- How strong/influential are accountability institutions such as the national audit office, parliament, the judiciary, audit and procurement rules, financial regulations? Do any special watchdog organizations exist?
- How important is the 'voice' of civil society—role of NGOs, the media and public opinion in highlighting problems with public sector performance? How open is public debate on issues concerning poor performance of government agencies? Do NGOs conduct customer or citizen surveys, and, if so, are their results publicly available?
- Do public sector agencies face competitive pressures, either inside or outside the public sector—via mechanisms such as competitive tendering and contracting, specification of performance contracts for senior management?

Checklist for Step 3: Clarify How Budget and Line Management Decisions Are Taken, and the Actual Roles of Ministries

- How important is the macroeconomic environment in determining the level of government spending? Is there a high level of fiscal discipline? How are government policy priorities decided, and how are they reflected in budget decisions and resource allocations?²⁰ How are competing demands for scarce budget resources resolved? Are budget decisions made in a considered, systematic manner?
- What kind of issues guide government's (cabinet's) decision-making?—poverty alleviation, lobby groups, factors that affect personal wealth or that of relations or colleagues?

-
- To what extent does information on the actual or expected performance of government activities influence government's decisionmaking?
- Does the budget system require an assessment of the performance of sectors/programs/ projects as an input to resource allocation decisions?
 - Is there a system for costing-out policy proposals and subjecting them to rigorous scrutiny (policy analysis)? Which ministries play the key roles in advising government on policy and expenditure priorities? Does their advice routinely draw on information concerning the performance of government activities? What is the quality of policy advice provided by central and line ministries?
 - What is the role of development assistance agencies, NGOs and private sector organizations (for example, via deliberation councils) in resource allocation decisions?
 - Are government decisions on development assistance loans/grants taken separately from domestic budget resource allocation decisions?
 - Is budget information prepared and published on actual expenditures, outputs and outcomes? Who prepares it, and how accurate is it? (See also Step 5.)
 - What are the roles of key central ministries with regard to line ministries? How actively are central ministries involved in monitoring and supporting line ministries, and in enforcing rules such as expenditure controls? How much autonomy do line ministries actually have? To what extent do line ministries have control over resources budgeted to them?
 - What surety do line ministries have concerning budget allocations and resource flows (how reliable are these—for example, how frequently are within-year variations made to approved spending levels)?
 - Do central and line ministries 'play by the rules'?
 - How capable are key central and line ministries—how well do they perform their functions at the level of autonomy they have been given? Do line managers have a clear understanding of their roles and of the objectives of the activities for which they are responsible? Do managers have the authority/autonomy and control over resources to manage so as to achieve these objectives?
 - What are the particular incentives which apply to key central and line ministries?

- Which ministries monitor and regulate the institutional arrangements in the civil service?
- Do there exist individual ministries which represent islands of 'good practice' within the civil service, in terms of their approach to performance management—for example, ministries which emphasize results-based management or which scrutinize the performance of individual managers via employment or performance contracts? If so, what have been the factors which have influenced their development?
- Do ministries or ministers pay attention to client or citizen surveys conducted by ministries or NGOs—to seek views/opinions on government services—especially line ministries? Do any ministries have client charters, and if so, what is their function?
- What roles do other accountability organizations play? Is internal and external audit taken seriously?

Checklist for Step 4: Extent of Influence of Evaluation on Budget and Line Management Decisions?

- Do line ministries commission and use information on the performance of sectors/programs/ projects in developing policy options?
- Is information on the performance of government activities reported in policy proposals sent by line ministers to government/cabinet? What types of proposal make use of such information for example, on investment projects whether and how much to borrow from development assistance agencies?
- What use do the finance and planning ministries make of evaluation in their policy analysis and policy advising work?
- Does cabinet commission evaluations or other types of performance measurement?
- Do line ministries and outrider agencies commission and make use of performance measurement in their ongoing management of sectors/programs/projects? Are there particular 'good practice' examples?
- Are evaluation findings and other performance information (such as performance audits) reported to parliament and development assistance agencies? What means are used to

report this information (for example, budget documentation; ICRs)? Is this information published?

Checklist for Step 5: Identify In-country Evaluation Capabilities

- Which organizations have supplied M&E studies in the past, and how many studies: line ministries, central ministries, other government agencies; universities/research institutions, local consultants, expatriate consultants, and development agencies? What has been the quality of these studies?
- Which ministries have specialist M&E and audit units? How many staff are employed in each? What are their areas of skill and experience (for example, social science research, program evaluation, cost-benefit analysis, statistical analysis, audit)? How many evaluation reports have they produced? What other types of performance information do they produce?
- What is the role of these units?
- Do specialist M&E units have their own evaluation guidelines, methods, policies and quality assurance/control procedures?
- What is the quality of the evaluation reports and other information which these units produce?
- Have any M&E units been set up for the life of a project only, and if so what are the skills and areas of experience of their staff?
- How are evaluation priorities determined and for what purposes? What involvement does senior ministry management have? What evaluation planning is undertaken? Do planning committees exist in line or central ministries? Do the finance and planning ministries have a role? Who determines the frequency and coverage of evaluations?
- Does cabinet commission any evaluations or policy reviews? If the latter, what role do evaluations play in them?
- How are evaluation findings and other performance information provided to users, clients and other stakeholders? Do the clients include line ministry senior management, policy analysts, line managers and operational staff, and central ministry analysts?
- What is the role of the national audit office and national statistical agency? What types of data or other information do they collect? How many of their staff have skills in performance measurement?

- What types of data exist on the costs and outputs of government sectors/programs/projects? What is the quality of these data (reliability, timeliness, degree of disaggregation)? Who produces these data (line and central ministries)?
- Do the universities and research institutions have staff with evaluation skills? Do they offer or have the capacity to offer training in a range of skills and evaluative techniques?
- How well developed is the market for local consultants? How good are their skills?

Checklist for Step 6: Identify Evaluation Activities/Roles of Development Assistance Agencies

- Which development agencies have provided support for ECD in the country?
- Have evaluation units been set up within the government specifically to service donor evaluation requirements (especially at the project or sector level)? If so, how long have they been in existence?; how long can they be expected to remain in existence? What happens to evaluation unit staff on the completion of donor projects?
- How many government officials work full-time (or equivalent) on servicing donor evaluation requirements? What are their evaluation skills/backgrounds?
- Is there any evidence that donor evaluation requirements either conflict with each other or impose duplicative layers of reporting on the government?
- Have donor evaluation requirements led to sustained increases in local evaluation capacity (by expanding local skills and their utilization)? If so, how? If not, why not?
- What other types of donor support for ECD have there been in the country?
- How long-term and sustained has donor support been?
- Has donor support been effective in developing sustained evaluation capacity?

Checklist for Step 7: Identify Recent and Future Public Sector Reforms

- What reforms have been undertaken in recent years in the areas of public sector management and governance? Have there been any reforms in the areas of budgetary systems

(public expenditure management), intergovernmental fiscal relations, commercialization and private sector delivery of public services, customer/client service standards, civil service reform, quality of civil service policy advice, participation and the ‘voice’ of civil society, and anti-corruption efforts?

- What reforms are planned or likely in coming years?

Checklist for Step 8: Map Options for Developing Evaluation Capacity

Note: For a menu of ECD options see Figure 3.

- Where is the source of demand for ECD? Are there domestic sources of demand and if so how strong is that demand? Is the push for ECD coming largely from development assistance agencies? How committed is the government to ECD? Is there a domestic champion ministry?
- On which of the three objectives of ECD is the country demand focused? (as an aid to resource allocation, to assist line managers, for accountability purposes)
- Are there potential ECD champion ministries, such as finance or planning ministries, or the national audit office? Should the evaluation regime be anchored in these ministries?
- How can influential central and line ministries be co-opted to support ECD, or at least refrain from opposing it? Which ministries would stand to lose from having information available on the performance of government?
- Which ministers and ministries are involved in public sector reform, and could they be induced to advocate/support ECD? What are the incentives for reform for political leaders and senior officials?
- What is the best strategy for winning hearts and minds within the civil service?
- Are there opportunities to put in place a ministerial decree or cabinet requirement for systematic evaluation?
- What synergies exist (or might be created) with other public sector reforms underway or planned?
- What opportunities exist to broaden and deepen the supply of evaluation skills and experience? What might be the role of training providers—universities and other research

institutions (in-country and regional), the World Bank and other development agencies, and government training bodies such as a civil service commission or a finance ministry?

- What have been the sources of supply of evaluations in the past? How capable are these sources?
- What is the quality and availability of socioeconomic statistics and data on the inputs to and outputs from government programs?
- How much would each option from a range of possible/feasible ECD options cost? What funding and other resources are necessary for ECD?
- How sustainable—and how reversible—would these ECD options be? What risks would the ECD strategy face? How dependent would the ECD strategy be on the continuing support of a small number of ministers or senior officials? Would the ECD strategy be likely to survive a change of government?
- What are the realistic timelines involved? It can take a long time to build demand and consensus, to develop skills, and to build civil service infrastructures—for example, in budget systems and in evaluation planning.
- What types of support could development agencies provide, and at what stages in the development of evaluation capacity? How sustained is that support likely to be?

Checklist for Step 9: Preparation of an ECD Action Plan

Demand

- types of data and evaluations currently available
- current sources of demand and their strength
- means of building demand and creating consensus—potential champions
- scope for creating ministerial decrees, regulations, etc relating to evaluation
- types of evaluation tool for which there would be demand (such as basic socioeconomic statistics; accounting data on costs; monitoring and performance indicators; surveys of clients and citizens; program and project evaluations—ex ante, efficiency and ex post; performance/efficiency audits)
- winners and losers from among different ministries

-
- how to enlist the support of powerful central and line ministries
 - how to enlist the support of ministers and ministries involved in other public sector reforms
 - how to ‘win hearts and minds’ in the civil service?
 - extent of possible demand from NGOs, the media?

Supply

- past sources of the different types of evaluation tool (listed above)
- quality and availability of socioeconomic data on government inputs and outputs
- means of broadening and deepening the supply of evaluation skills
- sources of training for staff (universities, research institutions, civil service commission, consulting firms, development assistance agencies, twinning arrangements/ secondments, etc)
- numbers of people who should be made available to work on monitoring and evaluation—and their needed qualifications and skills
- types of evaluation tool in which they would need to be skilled
- amount of funding which would be necessary/available

Levels

- national, sector, or for major project

Evaluation infrastructure

- choice of planning mechanisms to decide which evaluations to conduct, and mechanisms to ensure evaluations are conducted in a timely manner and their results are provided to those who commissioned them

Role of development assistance agencies

- types of support needed (such as advice, expertise, organization of seminars, training, loans and grants, identification of qualified consultants, support for local/regional universities and research institutions, preparation of guidance material including case studies, greater harmonization of donor evaluation requirements, joint government/donor evaluations)
- stages in the ECD process at which support would be needed

Timelines, sequencing, speed of implementation

- indicative targets for, say, 1, 3, 5 and 10 years into the future

Sustainability of new evaluation capacity

- risks and threats to sustainability

Endnotes

- ¹ An internal review of a small sample of World Bank PERs found that most failed to cover Steps 2, 3, 4, 5 or 8.
- ² Robert Picciotto, “Developing Evaluation Capacity: The Road Ahead”, in Proceedings from the Seminar and Workshop on Evaluation Capacity Development in Africa, Abidjan, African Development Bank and World Bank, forthcoming.
- ³ These issues are discussed in The World Bank, Evaluation Capacity Development, Report of the Task Force, Washington DC, 1994, and The World Bank, Public Sector Performance—the Critical Role of Evaluation, Washington DC, 1998.
- ⁴ A thorough analysis and checklist of budget management issues is contained in the The World Bank, Public Expenditure Management Handbook, Washington DC, 1998.
- ⁵ This list is derived from The World Bank, Public Expenditure Management Handbook, page 5.
- ⁶ ‘Incentives for corrupt behavior arise whenever public officials have wide discretion and little accountability’ (The World Bank, World Development Report 1997: The State in a Changing World, New York, NY: Oxford University Press, page 103.)
- ⁷ An ideal budget system includes fiscal discipline and surety in funding allocations, a unified budget—that is, not separated into capital and operations expenditure allocations, nor by domestic/donor funding splits—and a medium-term expenditure framework which makes visible to decisionmakers and managers the forward-year implications and commitments of spending decisions.
- ⁸ The finance ministry or cabinet office would be able to provide an initial assessment of the extent of the use of evaluation in the budget process.
- ⁹ A more in-depth assessment of the capabilities of individual specialist units would examine whether they have their own evaluation guidelines, methods, policies, and quality assurance/control procedures.
- ¹⁰ It would not usually be feasible to collect information on the evaluation skills of other, less easily-identifiable staff within a ministry unless they are members of an evaluation network or professional association.
- ¹¹ More detailed questions relating to the adequacy of financial management and accounting systems are not explored here. They include: Are there whole-of-government (finance ministry) systems?; Individual ministry systems?; Relationships/links between whole-of-government and ministry systems?; Accuracy and timeliness of data?; Adequacy of accounting and auditing standards?; Existence of professional bodies and supervisory organizations?; Adequacy of budgetary classifications, data and procedures?
- ¹² This includes, for example, merit-based hiring, firing and promotion, individual performance appraisal, and market-based salaries.
- ¹³ Budget coverage/structure, policy/planning, preparation, execution.
- ¹⁴ This figure is based on an action framework developed by R. Pablo Guerrero in Comparative Insights from Colombia, China, and Indonesia, The World Bank, Operations Evaluation Department Working Paper No. 5, January 1999.
- ¹⁵ Many projects funded by international development assistance agencies involve the creation of in-country evaluation units. These units typically exist only for the duration of the project. Unfortunately, there appears to be no real evidence that the existence of such units has a lasting effect on the evaluation capacity of the countries involved.
- ¹⁶ World Development Report 1997: The State in a Changing World, Washington D.C. 1998, page 97.
- ¹⁷ Public Expenditure Management Handbook, page 78.
- ¹⁸ Mario Marcel, “Lessons from Chile”, in The World Bank, Public Sector Performance—the Critical Role of Evaluation, Washington DC, 1998.
- ¹⁹ For example, there is provision in Colombia’s constitution, in a ministerial decree in Indonesia, and there was formerly a cabinet requirement in Australia. Most development assistance loans and grants have formal requirements for evaluation.
- ²⁰ More detailed issues here include: how the budget is structured, whether there is a unified budget (or a capital/operations funding split, or a split into domestic and donor funding, or both), whether there is a medium-term expenditure framework, and how external aid is managed.

Other Papers in this Series

Keith Mackay, *Lessons From National Experience*.

Stephen Brushett, *Zimbabwe: Issues and Opportunities*.

Alain Barberie, *Indonesia's National Evaluation System*.

Keith Mackay, *The Development of Australia's Evaluation System*.

R. Pablo Guerrero O., *Comparative Insights from Colombia, China, and Indonesia*

Other Recommended Reading

1. Keith Mackay (ed.), *Public Sector Performance—The Critical Role of Evaluation*.
2. The World Bank, *Building Evaluation Capacity*, Lessons & Practices no. 4 1994.*
3. The World Bank, *Designing Project Monitoring and Evaluation*, Lessons & Practices no. 8 1996.*
4. The World Bank, *Evaluating Development Outcomes: Methods for Judging Outcomes and Impacts*, Lessons & Practices no. 10 1997.*
5. The World Bank, *Assessing Development Effectiveness*.*

* These publications are soon to be available in both French and Spanish.

Operations Evaluations Department Publications

The Operations Evaluation Department (OED), an independent evaluation unit reporting to the World Bank's Executive Directors, rates the development impact and performance of all the Bank's completed lending operations. Results and recommendations are reported to the Executive Directors and fed back into the design and implementation of new policies and projects. In addition to the individual operations and country assistance programs, OED evaluates the Bank's policies and processes.

Operations evaluation studies, World Bank discussion papers, and all other documents are available from the World Bank InfoShop.

Summaries of studies and the full text of the *Précis* and *Lessons and Practices* can be read on the Internet at <http://www.worldbank.org/html/oed/index.htm>

How To Order OED Publications

Documents listed with a stock number and price code may be obtained through the World Bank's mail order service or from its InfoShop in downtown Washington, DC. For information on all other documents contact the World Bank InfoShop.

Ordering World Bank Publications

Customers in the United States and in territories not served by any of the Bank's publication distributors may send publication orders to:

The World Bank
P.O. Box 960
Herndon, VA 20172-0960
Fax: (703) 661-1501
Telephone: (703) 661-1580

The address for the World Bank publication database on the internet is: <http://www.worldbank.org> (*select publications/project info*).

E-mail: pic@worldbank.org
Fax number: (202) 522-1500
Telephone number: (202) 458-5454

The World Bank InfoShop

The World Bank InfoShop serves walk-in customers only. The InfoShop is located at:

701 18th Street, NW
Washington, DC 20433, USA

All other customers must place their orders through their local distributors.

Ordering via E-Mail

If you have an established account with the World Bank, you may transmit your order via electronic mail on the internet to: **books@worldbank.org** Please include your account number, billing and shipping addresses, the title and order number, quantity, and unit price for each item.

OED Study Series

- Evaluation and Development: The Institutional Dimension (1998)
- 1997 Annual Review of Development Effectiveness (1998)
- India: The Dairy Revolution (1998)
- The World Bank's Experience with Post-Conflict Reconstruction (1998)
- Financial Sector Reform: A Review of World Bank Assistance (1998)
- Rebuilding the Mozambique Economy: Assessment of a Development Partnership (1998)
- Mainstreaming Gender in World Bank Lending: An Update (1997)
- Agricultural Extension and Research: Achievements and Problems in National Systems (1997)
- Fiscal Management in Adjustment Lending (1997)
- Reforming Agriculture: The World Bank Goes to Market (1997)
- Paddy Irrigation and Water Management in Southeast Asia (1997)
- Poland Country Assistance Review: Partnership in a Transition Economy (1997)
- 1995 Evaluation Results (1997)
- Zambia Country Assistance Review: Turning an Economy Around (1997)
- The Aga Khan Support Program: A Third Evaluation (1996)
- Lending for Electric Power in Sub-Saharan Africa (1996)
- Industrial Restructuring: World Bank Experience, Future Challenges (1996)
- Social Dimensions of Adjustment: World Bank Experience, 1980-93 (1996)
- 1994 Evaluation Results (1996)
- Ghana Country Assistance Review: A Review of Development Effectiveness (1995)
- Evaluation and Development: Proceedings of the 1994 World Bank Conference (1995)
- Developing Industrial Technology: Lessons for Policy and Practice (1995)
- The World Bank and Irrigation (1995)
- 1993 Evaluation Results (1995)
- Structural and Sectoral Adjustment: World Bank Experience, 1980-82 (1995)
- Gender Issues in World Bank Lending (1995)
- The World Bank's Role in Human Resource Development in Sub-Saharan Africa: Education, Training, and Technical Assistance (1994)
- 1992 Evaluation Results (1994)
- New Lessons from Old Projects: The Workings on Rural Development in Northeast Brazil (1993, contains summaries in French, Portuguese, and Spanish)
- World Bank Approaches to the Environment in Brazil (1993, contains summaries in French, Portuguese, and Spanish)
- Rapid Appraisal Methods (1993)
- Trade Policy Reforms Under Adjustment Programs (1992)
- World Bank Support for Industrialization in Korea, India, and Indonesia (1992)
- Population and the World Bank: Implications from Eight Case Studies (1992)
- The Aga Khan Rural Support Program in Pakistan: Second Interim Evaluation (1990)