



EVALUATION CAPACITY DEVELOPMENT

ECD WORKING PAPER SERIES NO. 1: MAY 1998

Lessons from
National Experience

Keith Mackay



Introduction

The development community in general, and the World Bank in particular, has recognized the importance of governance issues and the need to support capacity building and institutional development. An aspect of capacity building that has received some attention in the past, but arguably less than it should have, is the link between public sector performance and the role of evaluation. National or sectoral evaluation systems can provide a powerful support to high-quality decision-making and performance in the public sector and, via these means, can support higher standards of governance and enhance development effectiveness.

The World Bank has been in providing technical assistance on evaluation systems to a number of countries, including in recent years Indonesia, Colombia, China, Zimbabwe, the Philippines and Argentina — other organizations have also been involved in these and other countries. Donors such as UNDP, EU, USAID, SDC, Danida, AfrDB, IDB, AusAID and AsDB are active or have firm plans to take a more active role in fostering national evaluation systems.

The World Bank is strongly supportive of efforts to develop evaluation capacity in developing country governments, and has discussed with a number of donors opportunities for closer collaboration and partnership in this area. One modality which was identified was the more systematic sharing of understanding and experience in developing national evaluation systems (*evaluation capacity development, or ECD*).

The Bank is therefore launching a new series of short papers on ECD issues, as part of its ECD Program. This paper is the first of the series, and reports key themes, findings and conclusions reached at a seminar which the Bank organized on ECD issues, held in Washington, DC, in April 1998. Several other papers from that seminar, reporting country case studies, are also being published. Future papers will discuss key ECD issues and additional case studies. A collection of selected papers and transcripts from the seminar will be published soon (see endpage for details concerning how to obtain a copy).

This series will help to present lessons learned in different contexts — i.e., what worked, what did not, and why — and thus contribute to a growing 'library' of ECD experience. An active approach will be taken to disseminate the papers in this series to share these lessons with World Bank staff, with other development agencies, and with developing country governments.

Elizabeth McAllister
Director
Operations Evaluation Department
May 1998

Lessons from National Experience

Keith Mackay

The priority for developing evaluation capacity

The potential for evaluation can be understood from 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 argued, evaluation is in many ways central to the effectiveness of development assistance:¹

- the development assistance community has turned to results-based management at the project, country and global levels. For this approach to work requires solid institutional capacity in countries;
- partnerships and coalitions among development agencies to help support country programs and institutions also require a common approach to evaluation and assessment; and
- it is evident that there is a growing need to demonstrate the effectiveness of development interventions to the electorates of industrial democracies.

The development of national evaluation capacities is a means to ensure that evaluation findings are available to assist countries in three key areas. First, evaluation findings can be an important input into government decision-making and prioritization, particularly in the budget process. Second, evaluation assists managers by revealing the performance of ongoing activities at the project, program or sector levels — it is therefore a management tool which leads to learning and improvement in the future (i.e., results-based management). Evaluation can also be used to assess the performance of organizations, institutional reform processes, and so on. 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 also be held accountable for performance. There may well exist trade-offs between these three uses of evaluation findings.

The concept of performance is central. It encompasses the efficiency of a 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.

There are many different types of evaluation tools; and they can be used in a variety of ways. These tools are related, but they can lead to confusion exacerbated by the different terminologies employed by evaluation practitioners. They all address performance measurement: ongoing monitoring and performance information; project and program evaluation — ex ante, ongoing/formative and ex post; performance (or value-for-money) audits; financial auditing, etc. 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. Unless otherwise stated, the term ‘evaluation’ is used in a generic, shorthand sense in this note to encompass these various terms and concepts.

Experts from different backgrounds — such as economists (who have traditionally focused on project evaluation), program evaluators (who typically have a broad social science background) and auditors (who in the past have emphasized financial compliance) — tend to use different concepts but with often similar or even identical nomenclature. It is little wonder that those new to the area of evaluation are often confused by the panoply of concepts and jargon.

Performance measurement is a valuable activity not least because it provides an opportunity and a framework for asking fundamental questions about any activity, such as: ‘what are you trying to achieve?’; ‘what does “success” look like?’; how will you know if or when you’ve achieved it?

“No public sector can afford to overlook the importance of clearly defining its objectives and priorities, assessing performance against well-defined benchmarks, and changing the bureaucratic culture into one that stresses client service and achievement of results. Rather than an imposed requirement of donor agencies, evaluation now becomes a key instrument of good governance and institutional development within our client countries. We all have a responsibility to make sure this function is nurtured and supported, as it has been within our own institutions.”

Mark Baird, ‘The role of evaluation’, in The World Bank, *Public Sector Performance - the Critical Role of Evaluation*, 1998 (forthcoming).

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 the areas of:

- budgetary financial management systems, including financial reporting and auditing;
- intergovernmental fiscal relations, and the extent to which they encompass a focus on performance;
- commercialization and the private sector delivery of public services — i.e., where governments exit from public service delivery. This requires a clear understanding of program objectives, and also ex-ante, ongoing and ex-post assessments of performance, to be successful;
- the setting of explicit customer service standards by service delivery agencies, and monitoring the extent to which these are actually achieved;
- civil service reform, including 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 or not that advice encompasses whatever evaluation findings exist or can be commissioned;
- participation and the ‘voice’ of civil society — this addresses

the issue of 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 also relates 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, in areas such as improved financial management systems, performance reporting, strengthening of watchdog agencies, and greater transparency in policymaking and implementation.

Country experience

There is a growing volume of experience with national evaluation systems in developed and developing countries. The experience of developed countries illustrates the potential links between national evaluation capacity and good governance. It reflects the opportunities and difficulties in achieving cultural change in a government — winning hearts and minds is a slow business. And it illustrates the different dimensions which must be developed to achieve a robust national evaluation system: in the areas of demand, supply, and information infrastructure.

The main *precondition* for the development of a national evaluation system is country demand — an evaluation system cannot be effectively foisted on an unwilling government. There are particular risks if the impetus for an evaluation system is donor-driven; this is not to say that donors cannot take the lead in ‘selling’ the merits of evaluation systems to countries, but that unless and until countries accept the strength of such arguments, or reach their own conclusions that evaluation has much to offer, an evaluation system is unlikely to be sustainable.

Two building blocks for effective demand are the sensitization of key stakeholders to the need for and benefits from evaluation, and the building of awareness of suitable techniques and approaches.

Experience tells us that the main barriers to developing evaluation systems in developing countries have been: lack of genuine demand and ownership in countries; lack of a culture of accountability, often related to issues of ethics or corruption; lack of evaluation, accounting, or auditing skills — there is a need to develop the supply of these skills and systems to match demand as and when it grows; poor quality of financial and other performance informa-

tion, and of accounting/auditing standards and systems; lack of evaluation feedback mechanisms into decision-making processes; and the need for the efforts to develop evaluation systems to have some minimum critical mass to be sustainable. (These issues are discussed in The World Bank, *Evaluation Capacity Development*, Report of the Task Force, Washington DC, 1994.)

What is not yet clear is whether a government or ministry needs to possess some minimum level of overall capability before an evaluation system can realistically be contemplated.

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 do not support it. If the framework and incentives are insufficient then this is a strong argument for ensuring that the development of an evaluation system is part of a broader initiative to develop governance; this approach recognizes the strong synergies between performance measurement/evaluation and performance management.

Success factors

Experience has revealed that there exist a number of success factors for the development of an evaluation system, and that this should be pursued only if a number of them already exist or if there are reasonable prospects for developing or creating them.

One factor is the powerful role which a ‘champion’ ministry or agency can have in supporting, encouraging and pushing the development of an evaluation system — this is related to the challenge of winning hearts and minds within government, and this requires a lead agency with power and influence, such as a finance or planning ministry (as in Australia and Canada, for example), or a national audit office. Countries such as Indonesia have found it valuable to have had the support of powerful central agencies, as well as to create a core group or network of committed supporters in ministries. The existence of an explicit and high-profile evaluation strategy can also be effective in selling the message, as can the support of individual ministers or the Cabinet as a whole — particularly via ministerial or presidential decrees.

Sustained government commitment is also important. An evalua-

tion system cannot be developed overnight; indeed experience indicates that it can take at least a decade at the whole-of-government level to embed such a system in a sustainable manner — it simply takes time to win hearts and minds, to develop necessary skills, and to set up civil service structures and systems to make full use of evaluation findings. The counterpart of the need for sustained government support is the need for sustained support from development assistance agencies — Alain Barberie has emphasized the value of having had the ongoing and active support of the World Bank in Indonesia.

“An active, visible and continuous contribution by World Bank staff to the efforts of the Government of Indonesia brought legitimacy to the evaluation function and reassured decision-makers about the importance and benefits of evaluation.”

Alain Barberie, ‘Indonesia’s national evaluation system’, in The World Bank, *Public Sector Performance - the Critical Role of Evaluation*, 1998 (forthcoming).

A whole-of-government approach (e.g., Indonesia, Chile, Canada, Australia) has advantages in terms of achieving momentum and helping to ensure that laggard ministries endeavor to keep up with leading ministries. A whole-of-government approach might be especially feasible if a major series of reforms is being contemplated, such as major changes to public expenditure management (ie budgetary processes and decision-making) in response to budgetary imperatives.

But a whole-of-government approach might be unrealistic in a number of developing countries — it can be difficult to achieve. In many countries a more achievable and modest approach might be to start with an initial focus on ongoing performance monitoring, possibly in particular sectors or ministries in order to create a demonstration effect, and then seek to spread the approach to other sectors/ministries and to other performance measurement tools (such as evaluation) as opportunities present themselves or can be created.

This type of sequencing implicitly depends on perceptions of satisfactory benefits vis-a-vis the costs of the initial, more modest approach. Chile has taken such an incremental approach, following an initial focus on performance information; this has helped to

generate demand for answers about program outcomes and causality, and these are the types of question which only summative program evaluations can fully answer.

Perhaps the key message here is the need to tailor the approach for developing evaluation capacity to the circumstances and opportunities in each country. It is definitely not the case that ‘one-size-fits-all’.

The final lesson is that incentives are crucial to ensuring both that an evaluation system is developed and that evaluation findings are actually used. In working with national governments it is important for those providing technical assistance to understand the incentive frameworks in the country. This would typically involve conducting an institutional diagnosis as a precursor to provision of advice, and to ensuring that this work includes a close dialogue with the government. Such an approach has been followed for Zimbabwe, for example. This diagnosis should also examine the extent of evaluation capacity in a government — it is often the case that governments over-estimate the extent and quality of their evaluation activity.

As Eduardo Wiesner has emphasized, existing incentives may work against evaluation — therefore 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 they will be used by finance or planning ministries to reduce budget appropriations or for accountability purposes — this underscores the trade-off between using such findings to assist budget allocations or for accountability, versus their use by managers to improve performance.

This leads to the issue of who should be responsible for measuring performance — impartial outsiders, or expert insiders. The former approach stresses objectivity and independence, while the latter stresses expert knowledge and ownership of the evaluation results, and this in turn is likely to encourage learning by managers and their staff.

One answer could be to emphasize *learning* where civil service performance and capabilities are believed to be reasonably good and the institutional environment is judged to be conducive to improvement; emphasize *accountability* especially where perfor-

mance is very poor and/or corruption is a problem.

In the latter case there are opportunities to use performance measurement to expose poor performance of the civil service and thus to increase the pressures for a more responsive public sector. This has been the experience in Bangalore, for example, where an NGO has conducted surveys of citizen views concerning the quality of government services in that city and the extent of corruption in delivering services. There is clearly scope for civil society including NGOs to play a role in performance measurement, and there may be opportunities for bilateral donors to promote and support such activities.

Implications for World Bank and other donors

The development of evaluation capacity is a long-term proposition, not amenable to a simple, two- or three-year project approach. Thus sustained support and commitment, and the active involvement of the country government is necessary. This suggests that development agencies should view the provision of technical assistance in this area as part of a sustained and ongoing partnership and dialogue with the government. The role of donors is important — they can help or hinder the development of evaluation capacity.

They can help by the provision of technical assistance — their advice and sharing of lessons and best practice can be invaluable, as can the provision of funds for training and helping to build evaluation systems. Their participation can also help to build confidence within the government.

They can hinder by excessive, conflicting or multiple donor requirements for evaluation. There is a particular danger that scarce country evaluation capacity will be diverted to satisfy donor requirements for the evaluation of development assistance activities, which might not necessarily align with the greatest areas of benefit from evaluation in the government.

New lending instruments, which provide multi-year funding for technical assistance activities, are likely to be better tailored to meeting country needs for long-term commitment and support in this area.

Collectively, donors have a considerable and growing track record in supporting such activities. The World Bank, for example, has had a separate program to support evaluation capacity development since 1987, and other multilateral donors such as the UNDP and Asian Development Bank, and some bilateral donors, have also been active.

There is considerable scope for closer donor collaboration and partnership in this area. This could take place via: the sharing of existing case studies, which are an invaluable learning tool; identification and analysis of existing good practice or promising practice countries; development of approaches for undertaking country diagnoses and identifying their potential to put an evaluation in place; and sponsorship or other support for regional collaboration including the sharing of experience via conferences, etc.

¹ Robert Picciotto, 'The role of evaluation in development', in The World Bank, *Public Sector Performance - the Critical Role of Evaluation*, 1998 (forthcoming).

Other Forthcoming Papers in this Series

Stephen Brushett, *Evaluation Capacity Development in Zimbabwe*.

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

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

Other Recommended Reading

1. The World Bank, *Colombia: Paving the Way for a Results-Oriented Public Sector*, WB Report no. 15300-CO, January 28, 1997.
2. The World Bank, *Report of the Evaluation Capacity Task Force*, June 1994.
3. Eduardo Wiesner D., *From Macroeconomic Correction to Public Sector Reform: the Critical Role of Evaluation*, WB Discussion Paper no. 214, 1993.
4. Derek Poate, *Measuring & Managing Results: Lessons for Development Cooperation*, SIDA/UNDP, 1997.
5. Robert Picciotto and Eduardo Wiesner D. (eds.), *Evaluation and Development*, Transaction Publishers, New Jersey, 1998.
6. R. Pablo Guerrero O., 'Evaluation capacity development in developing countries: applying the lessons from experience', in R. Boyle and D. Lemaire (eds.), *Building Evaluation Capacity*, Transaction Publishers, New Jersey, 1998.

Operations Evaluations Department Publications

The Operations Evaluations 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 and World Bank discussion papers can be obtained through the World Bank bookstore. All other documents are available from the World Bank Public Information Center (PIC).

Summaries of studies and the full text of the *Précis and Lessons & 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> (click on *publications*).

Internet: <http://www.worldbank.org/>. From the World Bank homepage, select *publications*.

E-mail: pic@worldbank.org

Fax number: (202) 522-1500

Telephone number: (202) 458-5454

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.

The OED Working Paper Series is produced by OEDPK (Partnerships and Knowledge). Copies or inquiries about these or other publications in our series can be obtained by calling 202-473-5365 or email: ecampbell-page@worldbank.org.

If you would like to receive a complimentary copy of the World Bank's forthcoming publication on *Public Sector Performance - the Critical Role of Evaluation*, 1998, please contact Ms Elaine Wylie via email (Bwylie@worldbank.org) or phone 202-473-1718.

OED Study Series

India: The Dairy Revolution

The World Bank's Experience with Post-conflict Reconstruction (1998)

Financial Sector Reform: A Review of World Bank Assistance (1998)

Bangladesh Country Assistance Review (1998)

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)

1995 Evaluation Results (1997)

Zambia Country Assistance Review: Turning an Economy Around (1997)

The Aga Khan Rural 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 Study in 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, 1980-92 (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 of 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)

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)