

The logo features the number '10' in a large, white, sans-serif font. The word 'YEARS' is written in a smaller, white, sans-serif font across the middle of the '0'. The background is a grid of squares with a color gradient from light green on the left to light blue on the right.

Independent Evaluation  
at the Asian Development Bank

# Evaluation for Better Results

Independent  
Evaluation 

The ADB logo consists of the letters 'ADB' in white, bold, sans-serif font, set against a dark blue square background.



**10**  
**YEARS**

Independent Evaluation  
at the Asian Development Bank

---

# **Evaluation for Better Results**

Independent  
Evaluation 

© 2014 Asian Development Bank

All rights reserved. Published 2014.  
Printed in the Philippines.

ISBN 978-92-9254-580-2 (Print), 978-92-9254-581-9 (PDF)  
Publication Stock No. BKK146674-2

Cataloging-In-Publication Data

Asian Development Bank.

Evaluation for better results

Mandaluyong City, Philippines: Asian Development Bank, 2014.

1. Accountability. 2. Learning. I. Asian Development Bank.

The views expressed in this publication are those of the authors and do not necessarily reflect the views and policies of the Asian Development Bank (ADB) or its Board of Governors or the governments they represent.

ADB does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use.

By making any designation of or reference to a particular territory or geographic area, or by using the term “country” in this publication, ADB does not intend to make any judgments as to the legal or other status of any territory or area.

ADB encourages printing or copying information exclusively for personal and noncommercial use with proper acknowledgment of ADB. Users are restricted from reselling, redistributing, or creating derivative works for commercial purposes without the express, written consent of ADB.

Note:

In this publication, “\$” refers to US dollars unless otherwise stated.

6 ADB Avenue, Mandaluyong City  
1550 Metro Manila, Philippines  
Tel +63 2 632 4444  
Fax +63 2 636 2444  
[www.adb.org](http://www.adb.org)

For orders, please contact:  
Public Information Center  
Fax +63 2 636 2584  
[adbpub@adb.org](mailto:adbpub@adb.org)

 Printed on recycled paper

# Contents

Foreword by <i>Vinod Thomas</i>	v
Preface by <i>Takehiko Nakao</i>	ix
Acknowledgments	xiii
Abbreviations	xv

## PART I: DEVELOPMENT EVALUATION CONTEXT

<b>1</b>	<b>Development Evaluation in Asia and the Pacific</b>	1
	<i>Haruhiko Kuroda</i>	
<b>2</b>	<b>What Is Development Effectiveness?</b>	5
	<i>Robert Picciotto</i>	
<b>3</b>	<b>Evidence-Based Policies in Asia's Development</b>	11
	<i>Peter Petri and Vinod Thomas</i>	

## PART II: MEETING EVALUATION CHALLENGES

<b>4</b>	<b>Evaluation and Results in Asia</b>	21
	<i>Zhongjing Wang</i>	
<b>5</b>	<b>The Journey to Independent Evaluation at ADB</b>	27
	<i>Drawing on contributions from directors general over the past 10 years</i>	
<b>6</b>	<b>Accountability and Learning: Two Sides of the Same Coin</b>	47
	<i>Cheryl Gray; Caroline Heider; Rachel Meghir, Horst Wattenbach, and Luigi Cuna; Ashwani Muthoo; Rakesh Nangia and Samer Hachem; Moises Schwartz; and Jan Willem van der Kaaij, Ivory Yong-Prötzel, and Bastiaan de Laat</i>	

**PART III: EVALUATION APPROACHES**

---

- 7 Corporate Evaluations in the World Bank Group: Approaches to Enhance Rigor and Evaluation Impact** 69  
*Anis Dani*
- 8 Influential Evaluations at ADB** 87  
*Hemamala Hettige*
- 9 Developing a Gender-Responsive Monitoring and Evaluation System for International Development Agencies** 101  
*Michael Bamberger*

**PART IV: STRENGTHENING EVALUATION CAPACITY AND COUNTRY EVALUATION SYSTEMS**

---

- 10 India's Independent Evaluation Office: A New Approach** 117  
*Ajay Chhibber*
- 11 Results-Based Evaluation in Malaysia** 127  
*Koshy Thomas*
- 12 Strengthening Evaluation Capacity in Asia and the Pacific** 139  
*Li Kouqing, Peng Runzhong, Zhao Min, and Wu Ningquin*

**REFERENCES**

---

**ABOUT THE CONTRIBUTORS**

---

# Foreword

The 10<sup>th</sup> anniversary of independent evaluation at the Asian Development Bank provides us an opportunity to reflect on the role of evaluation in development, its evolution, and ongoing challenges. This book draws on the experiences of distinguished development practitioners and evaluators who provide unique and personal insights into these changing dimensions of evaluation.

The thread of independence in evaluation runs through this book. Independent evaluation re-enforces evaluation's twin roles as first an instrument of accountability in providing objective evidence-based assessments, and second as an instrument of learning in providing lessons to steer operations toward greater effectiveness. The recent efforts worldwide to establish independent evaluation departments reflect the desire to harness evaluation's full potential. The care in delivering objective assessments and crafting contextually responsive recommendations, alongside success in sharing timely lessons, is a key determinant of effectiveness. The success in communicating evidence-based evaluation results to decision makers and implementers is integral to building consensus for taking action on evaluation's recommendations.

ADB has decisively given the function its independence, while also stressing the need for engagement with operations departments and the development community. Its evaluators have increasingly engaged with operations staff, client countries, civil society organizations, the private sector, and other stakeholders in dialogue, evaluation reviews, and learning events. Experience points to the value in substantive interactions by evaluators with various groups before, during, and after evaluation. Striking this balance between independence and engagement has been at the forefront in using evaluation to catalyze ADB's evolving role as a knowledge institution.

Evaluation cannot be considered in isolation of the development context in which it operates. The Asian experience has shown that the development journey can be a roller coaster where there is no one

direction. As development challenges are becoming more complex, replicating what worked in the past is no guarantee of continued success. Tried and tested approaches offer insights, but even for highly rated projects, they may not be adequate by themselves for dealing with dynamic and unfamiliar situations such as natural disasters or climate change. Evaluation needs to adapt to these emerging trends, connect with these evolving challenges, and inform the directions.

Rapid growth has been the hallmark of some of Asia's development. At the same time, the region has also seen vast poverty, rising inequality, and environmental degradation. Society now demands inclusive and environmentally sustained growth. Governments now accept that social and environmental approaches must be factored in to growth plans. A key ingredient is evaluative evidence that sheds light on how to improve policy and the delivery of services through strengthening linkages among plans, the allocation of resources, the implementation of programs, and the monitoring and evaluation of results. In this context, high-quality evaluation can inform and shape decisions on strategies, programs, and projects of governments, businesses, and development partners.

The need for innovation in evaluation is clearer than ever. Development problems, such as groundwater depletion, quality of education, and income distribution, remain intractable. For these, old solutions no longer suffice. And new problems, such as rising incidence of non-communicable diseases, environmental degradation, and climate change, add to the premium that one would pay for innovation in projects and portfolio.

Using novel approaches in the project cycle may usher in uncertainty, but it also raises recognition and reward for taking risks. Improvements in evaluation techniques are therefore essential. Designing evaluation to give innovators the information they need to discover new pathways, test solutions, and detect the results of their efforts is vital. Adaptive techniques that are sensitive to complexity and amenable to real time feedback and iteration are well suited for assessing innovation. Timely feedback is of great importance, particularly when a quick response is required and when assumptions change over time.

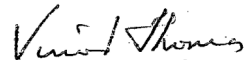
The various chapters in this book examine the themes of development's changing context and how evaluation needs to factor in emerging challenges, most notably the need for inclusive and



environmentally sustainable economic growth. Both are shaping the discourse on a post-2015 Millennium Development Goal agenda. This has highlighted even more the multidimensional nature of roles and challenges for evaluation. Indeed, the value and impact of independent evaluation hinges on how well these roles balance, reconcile, and complement each other.

Whether adequate actions are taken on evaluation's recommendations remains a point of concern. Is there enough follow-up on evaluation findings, lessons, and recommendations? How can the consideration of evaluation findings be mainstreamed into operational work in a similar way to planning, budgeting, and implementation and monitoring?

Evaluation offers considerable promise in informing policies and actions. But especially as evaluation becomes more influential, the field faces tough challenges regarding methods, approaches, and application. In seizing the many opportunities to help shape policy and practice, evaluation must meet these challenges on a continuing basis.



Vinod Thomas  
Director General  
Independent Evaluation  
Asian Development Bank



# Preface

This year, we mark the journey that began in January 2004 when the Asian Development Bank (ADB) initiated bold steps to provide organizational and financial autonomy for independent evaluation at the institution. ADB sought to protect the integrity of the evaluation process and to uphold impartial, credible, and rigorous evaluation findings. For the evaluators, it meant rising to the challenge of presenting evidence-based findings and helping shape decision making to improve development effectiveness.

Coinciding with the early years of independent evaluation was a pronounced concern about managing for development results—a prerequisite to improving the quality of ADB support. Multilateral and bilateral agencies at that time were pressed to demonstrate development effectiveness and to show concrete results. Accordingly, in addition to the evaluation of individual projects, evaluation of country, sector, and themes as well as of ADB's corporate initiatives gained prominence.

Independent evaluation at ADB has supported accountability and learning in many ways. Over time, it has provided insights into the elements of successful and unsuccessful development programs, highlighted the vital role of development effectiveness, and informed policy dialogue. Moreover, it has offered lessons and recommendations, which subsequently have fed into the preparation of country partnership strategies and operational plans, and the design of programs and projects. The launching in 2008 of an innovative computer-based management action record system (MARS) has allowed ADB to track recommendations from evaluations. The system has also provided an avenue for new partnerships with other central evaluation units in multilateral development banks.

For ADB, evaluation findings and recommendations have led to positive results at various levels. At the thematic level, the findings of the work over the past 2 years on natural disasters, social protection, inclusive growth, and Millennium Development Goals (MDGs) have contributed meaningfully to the debate on ADB's operational

directions. The ongoing evaluative work on governance and safeguards is highly relevant to the future direction of the organization.

Asia and the Pacific faces many challenges. Inequalities in income and access to economic and social opportunities are widening. Climate change puts sustainability of growth in jeopardy. A large infrastructure deficit remains. Moreover, many countries that have reached middle-income status face rapid urbanization, aging populations, and the task of sustaining inclusive growth.

A comprehensive midterm review of ADB's long-term strategy—Strategy 2020—called for ADB to sharpen its operational focus to better help its clients overcome challenges. This review has benefited from the findings of independent evaluation, and its follow up will be informed by further assessments. As a result, we now have 10 strategic priorities. Of these, seven seek to sharpen and rebalance ADB operations and strengthen their responsiveness to the changing business environment. These are poverty reduction and inclusive economic growth, environment and climate change, regional cooperation and integration, infrastructure development, middle-income countries, private sector development and operations, and knowledge solutions. The remaining three priorities aim at increasing ADB's capacity and effectiveness: financial resources and partnerships, delivering value for money, and organizing to meet new challenges.

In line with these strategic priorities, ADB must become more innovative, more inclusive, and more integrated. We need to innovate the mobilization of finance from public and private partners to meet the region's financing needs. We need to innovate our processes and products by reforming our business processes to increase our institutional efficiency and effectiveness. Moreover, we must promote innovative thinking and skills in our staff.

Our region can contribute to global development commensurate with its growing economic strength. Its weight in the global economy brings with it a new responsibility—that of contributing to the achievement of the post-2015 MDG agenda. We must be prepared to deal with cross-cutting development issues, support a development agenda that carefully combines the social, environmental, and economic dimensions of development in a single framework, and harness linkages to maximize results. The post-2015 MDG agenda must meet the basic needs of the poor and target critical social and environmental objectives to sustain development.


How does independent evaluation fit in with this process? For an organization like ours to remain relevant and to do the right things, we must continually assess our development effectiveness and uphold our accountability to stakeholders. In this light, we need independent evaluation to alert us to the alignment of our policies and operations with our strategic development agenda and to assess the coherence of our development programs. We need independent evaluation to provide evidence of success and failure based on the achievement of expected outcomes and impacts, and to offer pragmatic lessons and solutions to unfolding development problems.

As the development challenges of the future become increasingly connected, adopting appropriate evaluation methods remains important. Examining cross-boundary and cross-sector links will be essential in development evaluation, including synergies between public and private sector operations. As conditions become more dynamic and as uncertainties grow, evaluation for results will become even more crucial for informing policies and investment decisions.

On this 10th anniversary of independent evaluation at ADB, we recognize the commitment, dedication, and hard work of the evaluators. Evaluation has catalyzed improvements and shaped efforts to improve operational performance.

Given the tough development challenges in Asia and the Pacific, independent evaluation can play a clear and important role. It can help operations to keep emerging development imperatives in mind. It can shed light on the ingredients of development effectiveness even as they vary from country to country. And it can help assess and measure how progress is actually being made and indicate what more must be done.

The agenda and role of evaluation at ADB continue to evolve. We have learned lessons from the experience of countries and other international organizations while ADB's experience should provide useful pointers or lessons for others as well.



Takehiko Nakao  
President

Asian Development Bank

### **Editorial Board**

**Hemamala Hettige**, Senior Advisor, Independent Evaluation  
**Valerie Reppelin-Hill**, Advisor, Independent Evaluation  
**James Nugent**, Director General, Southeast Asia Department  
**Xianbin Yao**, Director General, Pacific Department

# Acknowledgments

The book has been prepared by a team led by Vinod Thomas and comprising Hemamala Hettige, Valerie Reppelin-Hill, Brenda Katon, and Michael Diza. The team is grateful to all contributors for deepening the discourse on evaluation challenges, approaches, and capacity development and for sharing their experiences in evaluation. Our sincere appreciation goes to James Nugent and Xianbin Yao for providing directions and suggestions to improve the coverage, content, and flow of the book. We thank Njoman George Bestari, Jocelyn Tubadeza, and Renato Lumain for their inputs during the early months of manuscript conceptualization and preparation.

We acknowledge valuable information on influential evaluations from the Independent Evaluation Department staff of ADB. We appreciate Tomoyuki Saisu for providing comments on ADB contributions. Likewise, we thank Shireen Lateef for reviewing the chapter on a gender-responsive monitoring and evaluation system.

We acknowledge the editorial support from Alastair McIndoe and Eric van Zant at various stages to make the manuscript reader-friendly. We thank our manuscript readers—Farzana Ahmed, Alastair Dingwall, Bob Finlayson, and Marie Therese Ng—for their valuable comments. We also appreciate Kae Sugawara for proofreading the manuscript. Vital publishing and printing support came from the Publishing and Dissemination Unit of the Department of External Relations and from the Printing Services Unit of the Office of Administrative Services.

This book would not have been possible without collaboration on various fronts. Our deep gratitude goes to all who participated in this undertaking.





# Abbreviations

ADB	- Asian Development Bank
ADF	- Asian Development Fund
AFDC	- Asia-Pacific Finance and Development Center
CLEAR	- Center for Learning on Evaluation and Results
DAC	- Development Assistance Committee
DEC	- Development Effectiveness Committee
EIB	- European Investment Bank
FCAS	- fragile and conflict-affected situation
GHG	- greenhouse gas
IED	- Independent Evaluation Department
IEG	- Independent Evaluation Group
IFAD	- International Fund for Agricultural Development
IMF	- International Monetary Fund
IOE	- Independent Office of Evaluation
MARS	- management action record system
MDB	- multilateral development bank
MDG	- Millennium Development Goals
M&E	- monitoring and evaluation
MFF	- multitranches financing facility
NGO	- nongovernmental organization
OECD	- Organisation for Economic Co-operation and Development
OED	- Operations Evaluation Department
PRC	- People's Republic of China
SHIPDET	- Shanghai International Program for Development Evaluation Training





**Part 1**

# **Development Evaluation Context**

Asia faces a diverse, sometimes unique set of social and economic challenges in the coming decades. Rising economic inequality and inertia undermine efforts to propel economies to higher income levels in many countries; in others, poverty remains deep and widespread; and in most environmental problems are already endemic. Uniting all of these issues is the threat they pose to the sustainability of growth and development, and the role that better evaluation can play in finding appropriate solutions.

This calls for an emphasis on development effectiveness grounded in the evaluation criteria. At the same time, projects and programs going forward must pursue evidence-based development that works toward a triple bottom line—growth, social inclusion, and environmental sustainability.

The first chapter, from former ADB President Haruhiko Kuroda, looks closely at the looming issues and details the broad consensus for areas of action. It argues that in overcoming the litany of risks—from income disparities, through climate change, infrastructure gaps, and rising urbanization, to the middle-income trap—effective use of resources is vital. And, in this, evaluation is a valuable tool for looking objectively at what has worked and what has not, and on the basis of those lessons, to improve development approaches.

In the next chapter, Robert Picciotto, professor at King's College London, takes up the discussion within the emerging global context. “A broader conception of development is taking hold,” he notes, increasing the urgency of cooperation in the development community and for better evaluation.

Development organizations operating in a globalized world, he argues, must focus on development effectiveness that emphasizes outcomes and impacts, rather than inputs and outputs. In this context, the Development Assistance Committee's five development effectiveness criteria (relevance, efficacy, efficiency, sustainability, and impact) best define that goal.

Rounding out the section, Peter Petri and Vinod Thomas, in the third chapter, zero in on the need for evidence-based strategies in the next stage of Asian development. Rigorous evidence and analysis must support proposed solutions to persuade all concerned of the need for action. And action, in turn, must pursue multiple goals simultaneously. Sustaining the next decades of development requires that growth is rapid, inclusive, and environmentally sustainable.

## Chapter 1

# Development Evaluation in Asia and the Pacific

*Haruhiko Kuroda*

The Asia and Pacific region has enjoyed high growth and rapid socioeconomic progress over several consecutive decades. Yet, in the coming decades the region will likely face increasing risks associated with disparities within and across nations, competition for finite natural resources, environmental degradation, climate change, energy access, food security, infrastructure gaps, rising urbanization, water resource constraints, and the middle-income trap. Moreover, developing Asia needs to tackle the challenges of ensuring good governance and building strong institutions that can provide transparency, accountability, and rule of law.

Both in good times and bad, effective use of resources is a vital goal. And evaluation is a valuable tool for looking objectively at what has worked and what has not, and on the basis of those lessons, for helping to improve approaches and their implementation.

The independence of the evaluation function in any organization can add to credibility and the impact of the lessons brought to the table. The Asian Development Bank (ADB) has valued the role of independent evaluation in informing the development effectiveness of its operations, and it is proud of the constructive collaboration and mutual trust between the Independent Evaluation Department and the management and operations departments.

The rapid economic growth of recent decades, studies show, has lifted millions of people out of poverty. But Asia remains home to two-thirds of the world's poor, and its remarkable economic growth has been accompanied by rising inequality and serious environmental consequences. Inequality has widened in a large number of Asian countries, including the three most populous, fast-growing nations—the People's Republic of China, India, and Indonesia. And across

developing Asia, the Gini coefficient has increased, indicating greater inequality and ringing alarm bells.

To sustain economic growth, its benefits need to be broadly shared. Yet, inequality can undermine sustainable growth: it weakens the sense of shared objectives that are necessary for transformational change and reduces social cohesion. An effective way to ensure the region's growth is increasingly inclusive is to create opportunities for remunerative, productive, and fulfilling jobs. By spending more on infrastructure, education, health, and social protection, governments can also promote inclusive growth while fostering economic expansion. We need to support these goals, including through sound evaluation and development that continues to tackle poverty and that is socially inclusive and environmentally sustainable.

Resource-intensive economic growth is revealing the problems of resource constraints and rising disaster effects. This growth pattern may be unsustainable. To move the development agenda forward, the region must become disaster resilient and promote a low-carbon development path. The region typically experiences about one-third of the world's disasters and is home to a disproportionate number of disaster-affected people.

One of the most important factors increasingly affecting natural disasters is climate change and volatility, the evidence for which the world is now seeing. Rising sea levels, higher temperatures, desertification, droughts, and increased floods can all have dramatic effects on people. The annual economic costs of disasters run into billions of dollars, posing potentially major setbacks to development.

Evaluations tell us that the region needs prominent efforts to reduce disaster risk and build resilience to natural hazards. Attaining a resilient posture requires multidisciplinary collaboration across sectors, themes, and boundaries. This requires political commitment, human resource utilization, and knowledge, and poses significant challenges. Asia cannot afford to grow now and clean up later. With commitment, innovation, and appropriate investment, the region can lead the world toward a more sustainable and resilient future.

Regional cooperation can help countries across Asia and the Pacific tackle common challenges and seize opportunities. These include dealing with climate change and environmental degradation,

liberalizing trade within the region, strengthening connectivity through investment in infrastructure, promoting partnerships and collaboration in many priority areas, and building sophisticated financial systems to better channel savings into productive investments at home. Support for regional platforms for creating and exchanging knowledge resources is a key part of ADB's commitment to strengthen regional cooperation.

The region also needs to recognize the importance of knowledge-led growth. To be competitive, economies in Asia and the Pacific must promote entrepreneurship and innovation, use resources more efficiently, and harness new technologies and creative ideas. Moreover, the region must have institutional, governance, and regulatory frameworks that offer incentives for innovation, promote competition, foster research and development, and protect intellectual property rights. Investments in tertiary and vocational education, and research and development are important to help businesses succeed in this highly competitive global environment.

Better knowledge products and services, and improved access to services within the region can be instrumental enabling factors in advancing sustainable development. There has been an increase in regional cooperation for sustainable management of natural resources that are an important part of the global commons. The potential for progress and the benefits through regional cooperation are visible across the region. Asia and the Pacific must continue to deepen cooperation and integration to improve economic resilience and respond effectively to global challenges, such as climate change, food security, water resource constraints, and energy shortages.

ADB embraces a results-focused management system, with a results-focused culture embedded in our operations and at all levels of our work. Hence, it is essential that we have independent evaluation of our work to provide an objective performance assessment that will also allow us to learn lessons from evaluations.

The evaluation feedback loop is integral to poverty fighting efforts in Asia and the Pacific. It contributes to the development effectiveness of ADB operations by providing evaluation feedback on performance, and generates evaluation lessons that contribute to our continuous learning and relentless search for better ways to improve performance and the development effectiveness of our contributions.





## Chapter 2

# What Is Development Effectiveness?

*Robert Picciotto*

*“The challenge of development, in the broadest sense, is to improve the quality of life”...World Development Report, 1991*

**A**t a time of extraordinary turmoil in the global economy, development effectiveness has become a major focus of debate. Both the ends and the means of development policy are being reconsidered. A broader conception of development is taking hold. The more comprehensive development agenda as well as a growing public appetite for results has raised the bar for the development cooperation enterprise—and for evaluation.

The idea of development is still remarkably influential. The basic development effectiveness concepts multilateral development banks now use are sound: they are equally useful at project, country, regional, and global levels. But they only contribute to sound decision making when used within a balanced portfolio of high-quality evaluations that make full use of the evaluation tool kit. Considering the increased volatility and risks of the current operating environment, development evaluation must give adequate emphasis to sustainability and impact criteria.

Given the interconnectedness of global economies, development evaluations must adopt coherence as a development effectiveness criterion. Considering the increasingly fragmented aid architecture, they must embrace the Millennium Development Goals (MDGs) and the Paris Declaration on Aid Effectiveness. And to promote accountability, they must assess the distinctive development contributions of individual partners in achieving development outcomes.

To fully contribute to development effectiveness, development evaluations must be formative as well as summative. They must focus on the actions needed to overcome the policy and institutional obstacles that obstruct sustainable poverty reduction.

## Assessing development effectiveness

Development effectiveness definitions are legion. Some focus on achieving predetermined ends. For the United Nations, development effectiveness is about “bringing about targeted changes in people’s lives” so that measuring development effectiveness can be equated to “an exercise in tracking progress toward development goals” (United Nations Development Programme 2001). Others highlight the means deployed to achieve results. For example, the Paris Declaration focuses on key characteristics of aid delivery (OECD, n.d.), while the Istanbul Principles endorsed by civil society organizations embrace a wide range of policies associated with democratic power sharing, environmental sustainability, human rights, gender equality, and others.

All of these definitions reflect admirable aspirations, but they are not fit for evaluative purposes because the value of achieving predetermined goals depends on their pertinence and on a comprehensive listing of all facets of effective aid. The good society makes the assessment process cumbersome and impractical. The Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development has endorsed a more useful alternative geared to practical development realities. It states that a development intervention can be labeled effective only if it achieves its relevant objectives efficiently.

Virtually all development assistance agencies have endorsed this straightforward performance test (the *outcome* rating). It is grounded in the trilogy of relevance, efficacy,<sup>1</sup> and efficiency that informs international financial institutions’ ratings of development policies and operations. It defines the three dominant development effectiveness terms as follows:

- *Relevance*: the extent to which the objectives of the development intervention are consistent with the beneficiaries’ requirements, country needs, global priorities, and partners’ and donors’ policies.
- *Efficacy*: the extent to which the development intervention’s objectives were achieved or expected to be achieved taking account of their relative importance.

- *Efficiency*: a measure of how economically resources and inputs (funds, expertise, time, and others) are converted to results.

### **Complexity and risk**

Outcome ratings grounded in objective assessments of relevance, efficacy, and efficiency are sometimes criticized on the grounds that they are awarded well before the effects of a development intervention can be fully ascertained. Evaluations normally take place 1 or 2 years after completion of disbursements, or 7–8 years after their approval. But postponing evaluations beyond such a period would not contribute to managerial accountability or organizational learning since evaluation delayed is evaluation denied.

To help compensate for the limitations associated with outcome ratings awarded soon after implementation, two important development effectiveness criteria are included in the development effectiveness tool kit. The DAC has also endorsed these:

- *Sustainability*: the continuation of benefits from a development intervention after major development assistance has been completed.
- *Impact*: the positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.

Multilateral development banks use these criteria in parallel and sometimes in conjunction with the three criteria that make up the development outcome trilogy. But they are not always given the analytical care they deserve. Nor are they routinely assessed using the full panoply of state-of-the-art evaluation methods that seek to focus on risk and complexity. Looking ahead, these criteria must be given their full due and their assessment must take advantage of recent methodological advances in evaluation (Patton 2011).

### **The logic of development effectiveness**

The above criteria have stood the test of time. They are well adapted to a pragmatic and results-oriented stance regarding the development process. They were forged through hard-won lessons of evaluation

practice. They have supplanted prior approaches that mistakenly focused on inputs and outputs rather than on outcomes and impacts. By focusing on the full range of development considerations, they have provided vital complements to the auditing approaches that emphasize compliance with established norms.

All five DAC-endorsed development effectiveness criteria (relevance, efficacy, efficiency, sustainability, and impact) matter to the success of development operations. Relevance is about *doing the right things* while effectiveness and efficiency have to do with *doing things right*. Assessing relevance is critical since achieving the wrong goals efficiently is counterproductive. Efficacy is also vital: excellence of development goals matters little if the vision they embody is not realized. Efficiency matters, too, since reaching even highly relevant operational goals cannot qualify as development success when excessive costs are incurred or scarce resources misallocated. The bottom line is results so that assessing impact and sustainability is a critical dimension of development effectiveness.

The relevance criterion is what most distinguishes evaluation from auditing. To be sure, achieving full consistency with beneficiaries' felt needs, country policies, donor requirements, and global priorities in one fell swoop is rarely feasible. Trade-offs must be struck and selectivity practiced lest goals become unrealistically ambitious (so that failure to achieve them may not be significant) or project designs become so complex that operational efficacy and/or efficiency are threatened. Equally, efficacy judgments must be used judiciously. In particular, they must not be wielded to penalize operations that do not meet overambitious goals if they deal with pertinent goals and deliver value for money; that is, make better use of scarce resources compared to the alternatives.

The multi-criteria development effectiveness concept sketched above helps track progress toward the objectives of development interventions and encourages evaluators to ascertain that the lives of people in developing countries have been tangibly improved. In a world of constrained aid budgets, it induces systematic examination of whether the interventions being evaluated have made economical use of resources—that the same results could not have been achieved more cheaply.

## Development outcomes versus agency performance

Development effectiveness criteria shape organizational incentives. As a *summative* endeavor, evaluation charts development. As a *formative* process, it helps shape it. Whereas goal-free evaluations facilitate a focus on development results as perceived by beneficiaries, goal-oriented evaluations associated with the DAC criteria enhance organizational learning as well as managerial accountability for a simple reason: they relate results to the promises made when loans, credits, and grants are approved.

This said, most development cooperation activities rely on *partnerships* to achieve development outcomes. This is the right thing to do, but it also blurs agency accountability. For example, responsibility for failure may be shirked altogether if it is summarily attributed to poor partner performance. Conversely, responsibility for success may be unfairly captured by the development agency—whether its contribution to the shared objectives justifies it or not.

When program or project failure is ascribed entirely to the development agency, it induces risk aversion: it may even encourage suspension of programs that fail to meet ambitious goals, forsaking the opportunity to adapt them so that they can succeed. It follows that evaluations must take explicit account of partners' distinct accountabilities and reciprocal obligations. Unless agency performance is assessed separately, moral hazard is bound to prevail.

Hence, good evaluations go beyond answering the question of whether a particular program works or not. They enrich development outcome and impact assessments by generating agency performance ratings along with outcome, sustainability, and impact ratings, which deliberately focus on the extent to which an agency has lived up to its distinctive obligations within the broad-based partnership arrangements. Rather than limiting its focus to the attribution question, good evaluations deliberately tackle the contribution question—how well did the individual development partners perform toward the achievement of program or project objectives, and what might be done to improve their performance?

## Impact evaluations complement outcome ratings

At their best, development operations are policy experiments that contribute to learning as well as improved governance and enhanced

accountability. In this context, freestanding and rigorous impact assessments retain their utility since the results directly attributable to development interventions are rarely captured with rigor in standard implementation completion reports.

By contrast, rigorous impact assessments seek to account for exogenous influences to recognize confounding factors and to consider explicit counterfactual scenarios. Carried out on mature and representative operations, they are an integral part of the evaluator's tool kit. Experimental methods may be useful in this context, but they have limitations. They are costly, poorly adapted to interventions, constrained by ethical considerations, and highly demanding in terms of specialized skills. Hence, they must be carried out selectively, and impact evaluations must rely on mixed methods adapted to the circumstances of individual cases (Picciotto 2012).

## Conclusions

The development effectiveness concepts the DAC has endorsed are valid and useful evaluation tools at project, country, regional, and global levels. But their potential cannot be realized outside of a well-conceived program of development evaluations that displays independence and excellence.<sup>2</sup>

Achieving and measuring development effectiveness implies a broad but explicit conception of development with evaluable development interventions. Among the various aspects for emphasis, sustainability and impact criteria would merit a special focus. Complex and nuanced as it may need to be, increased attention to agency performance ratings is essential to getting better results. Selective resort to rigorous impact assessments using mixed methods can enhance the value of the evaluative work.

## Notes

<sup>1</sup> The dominant use of the term effectiveness to connote efficacy helps explain why the term development effectiveness is often conceived far too narrowly by emphasizing achievement of goals without adequate consideration of their merit and worth.

<sup>2</sup> Compliance with the good practices promulgated by the Evaluation Cooperation Group must be encouraged. [https://wpqr1.adb.org/LotusQuickr/ecg/Main.nsf/h\\_Toc/797c4001cbfc7a564825773100281fa6/?OpenDocument](https://wpqr1.adb.org/LotusQuickr/ecg/Main.nsf/h_Toc/797c4001cbfc7a564825773100281fa6/?OpenDocument)

## Chapter 3

# Evidence-Based Policies in Asia's Development

*Peter Petri and Vinod Thomas*

Since 1990, developing Asia has more than tripled its living standards. The incidence of extreme poverty—the percentage of people with expenditures below \$1.25 per day—has fallen from 54% to 21%. This happened, over the last decade, in a region that has accounted for about half of world growth.<sup>1</sup> The recipe for this remarkable record, although richly varied, can be traced to the rise of evidence-based economic decisions in much of Asia—in other words, the pragmatic pursuit of policies that work. This is akin to learning from formative evaluation.

Should Asia simply continue on its established growth path? Evidence strongly argues for widening the region's priorities due to two compelling factors. First, over the last decade, income inequality has worsened substantially in countries accounting for 80% of the region's population. Second, key environmental indicators are deteriorating in much of the region, with Asia now the world's leading emitter of greenhouse gases.

Decades of experience have helped Asian decision makers understand the requirements of growth, but the task of tackling the wider challenges that they now face has just begun. Evidence now argues for controlling the externalities of growth, but not for abandoning growth as a priority. Win-win options exist because of past failures of markets, prices, and policies. But such options may not be enough and may need to be complemented by policies that deliver overall gains at some initial cost.<sup>2</sup>

Despite large potential gains, pursuing a triple bottom line—growth, social inclusion, and environmental sustainability—will be difficult because of resistance from vested interests. Even policies that yield gains on economic and social or environmental objectives, such as eliminating fuel or fertilizer subsidies, involve winners and losers.

In general, social and environmental policies meet more resistance than policies that promote growth—the latter produce profits for market agents, while those that target social and environmental goals often do not.

### **Past policies for growth**

Asia's diversity defies the easy generalizations that are sometimes used in proposing an "Asian development model." Still, the adoption of evidence-based policy making represents an underexplored commonality behind the region's remarkable economic successes. This concept is relatively new in the development literature, but pragmatism, flexibility, and gradualism—fundamental elements of evidence-based policies—have been long recognized as characteristic of Asia's successful economies.<sup>3</sup> In these countries, policies were based on results and were adjusted to changing circumstances.

The trial-and-error approach spread across many Asian economies, eventually leading to a cluster of growth-promoting measures that featured outward orientation, stable macroeconomics, and high savings and investments, including in human capital. In the 1960s, Japan modernized its war-torn economy, developed export markets for increasingly sophisticated products, and sharply raised income levels. Later, the Republic of Korea; Taipei, China; Singapore; and Hong Kong, China adopted similar strategies, at times through the purposeful imitation of Japanese policy (Petri 1988). With variations, similar strategies also emerged in Indonesia, Malaysia, and Thailand, and still later in the People's Republic of China (PRC), the Philippines, and Viet Nam. The approach is now being emulated in Central and West Asia and in South Asia, particularly in countries with low income levels and limited natural resource endowments.

Governments played a central role in most Asian success stories. They did not necessarily do so with direct expenditures or state-owned sectors, but those examples also exist. More generally, they focused on economic goals and mobilized massive resources, with savings rates at times reaching half of output. Some governments made large investments themselves, and others made investment resources available for market agents. Governments assembled capable bureaucracies, monitored performance, and reacted quickly to poor results and new challenges.



Colorful examples illustrate how the approach worked. In the Republic of Korea, former President Park Chung Hee launched his economic growth campaign by arresting 24 business leaders to underline their accountability for results and to government (Jones and Sakong 1980). The government later monitored the largest exporting companies daily and provided access to credit and other forms of government support based on export results. Some measures were provided selectively, while others, such as generous trade credit, were directly tied to performance.<sup>4</sup>

In the PRC, reforms began as experiments in smaller subregions and were later scaled up as part of a “scientific development” strategy. Agricultural reforms—the household responsibility system—were based on models practiced by some collective units. Industrial reforms were also rolled out through a sequence of experiments. Currently, financial experiments are under way in major centers such as Shanghai and Guangdong. Lin and Wang (2012) conclude that, responding to uncertainties inherent in the development process, policy followed “a process of learning, selective adaptation, and innovation.”

An especially intense wave of learning followed the Asian financial crisis of 1997–1998, focusing on financial regulation, market oversight, and foreign exchange policies (ADB 2000a). Countries scrutinized their own policy results, as well as those of others in similar circumstances. Mechanisms were developed to manage bankruptcies and bank failures, to monitor financial markets, and to foster macroeconomic resilience. Exchange rates were allowed to settle at relatively favorable levels and substantial foreign exchange reserves were built up to defend against currency runs. Some countries adopted contingency plans to manage capital flows.

Economic results were not always positive. There were many prominent failures, ranging from the Great Leap Forward in the PRC and the dominance of the License Raj in India, to major setbacks in the Democratic People's Republic of Korea, missed opportunities in the Philippines, and missteps in virtually all countries. Clearly, the common denominator of success was not luck or culture (as some would put it, “what's in the water”), since successes typically emerged from adversity, failure, and hard experience.

## Lessons for the future

Evidence-based strategies will be even more important in the next stage of Asian development. Evidence will be required to establish priorities, to gain public acceptance, and to choose the best policies. Refining the feedback cycle between policies and results will be critical to success, and this will require innovations in monitoring and evaluation. The roles of national and international agencies in the policy process will also need to be reconsidered.

Evidence-based strategies require policy decisions to be based on real results. Accurate monitoring and data are key inputs in this feedback process. Public engagement in monitoring the effects of policies could substantially improve the chances for their success. This points to the need to develop monitoring and evaluation capacity in national agencies.

These insights have spawned developments in a professional field—monitoring and evaluation—and systematic efforts to extract principles and lessons from experience. The Development Assistance Committee of the Organisation for Economic Co-operation and Development has codified criteria for evaluating policy.

The science of evaluation is in its early stages. Some government interventions may be rated highly on narrow, project-oriented criteria, but may still fail to have significant impact. At the other extreme, if assessments are focused on broad effects rather than project-specific results, their conclusions may be inaccurate and dominated by uncontrolled factors. New assessment tools, such as randomized experiments, are designed to avoid these errors (Duflo 2011).

Unfortunately, as monitoring and evaluation becomes more influential in policy making, incentives grow to make its results fit the interests of agencies or individuals. Paradoxically, this calls for insulating data collection and analysis from government influence, despite greater government demands for evidence. Solutions may involve conducting monitoring and evaluation through independent institutes or respected private sector agencies. Media scrutiny and public dissemination offer additional tools—albeit not always reliable ones—to prevent manipulation.

Good evidence is not enough; decision makers must also have the skills and analytical support to interpret and analyze data. This is especially important if policy decisions are separated from data collection and analysis in the interest of objectivity. Small analytical units within decision-making offices can help to bridge these gaps. The Republic of Korea and some other countries have already established such dual systems to collect and analyze data in research institutes, but have also created capabilities within government to sift through external evidence. The PRC is also in the process of separating research and analysis from the formulation of policy.

International institutions and the donor community will be especially important in an era of triple-bottom-line policies. Relative to the traditional priorities of economic growth, the new priorities require broader policy perspectives, more data and information, and greater unconventional financing to implement. In addition to its traditional financing role, the external financing community can help to provide these assets, including financial support for new policies, research, and information, and to improve decision making itself.

Since environmental issues often involve global externalities, solutions will often require financial flows from advanced to developing economies, beyond the traditional project and program financing. These flows may be interpreted as compensation for the free ride that developed economies enjoyed before global constraints tightened. International agencies will be the conduits of such transfers, under formal climate change mechanisms such as the Clean Development Mechanism, as well as development programs. In providing financial support, they will also play a role in brokering, implementing, and monitoring solutions.

International agencies can help to shape and disseminate knowledge. The United Nations Development Programme (through its *Human Development Report*), the United Nations (through the MDG initiative), and the multilateral development banks have already developed analytical frameworks and data for setting objectives beyond growth. Similarly, the Intergovernmental Panel on Climate Change, the International Energy Agency, and the multilateral development banks have contributed research on environmental issues. These agencies are themselves subject to the risk of being captured by special interests, but, ideally, competition among them will provide incentives for objective analysis.

International agencies can contribute to the complex national debates that are emerging in both advanced and emerging economies. Triple-bottom-line priorities, although urgent from a global and societal viewpoint, are, as noted, often opposed by powerful local vested interests. Objective, external sources of information, analysis, and support for policy can make a difference in tilting policy decisions toward the best solutions.

## **Conclusions**

Evidence now argues for policies to control the externalities of growth. This does not mean abandoning growth as an objective. The threat of an economic slowdown in middle-income Asia is real, and accelerating growth in the poorest countries is urgent. Rather, multiple goals have to be dealt with simultaneously. Inclusive, sustainable growth—long an aspiration—has to become a firm priority. In fact, early actions that prioritize inclusion and sustainability are likely to involve win-win policies, that is, initiatives that generate positive distributional and environmental results as well as increased economic efficiency. These options appear to be significant because of a wide range of market failures and distortions that often exist in developing economies.

The challenge of shifting strategies is nevertheless substantial, and will require innovation in many areas of policy. Attractive policy options exist, and many more will emerge, including win-win policies. Yet many policies will face stiff resistance from vested interests. It is important to understand the constraints that stand in the way of better outcomes and the solutions that might relax them.

Smart, effective governance will be critical to making the new priorities work. Three approaches to policy making could make a difference. First, the tool kit of governance has to be updated to the requirements of changing priorities. This will involve not just improved central governance, but also fresh experiments with accountability. Governments can monitor economic results centrally, but they need public feedback to implement social and environmental goals. Fortunately, new technologies are available to facilitate these interactions.

Second, the search for solutions must be supported by rigorous evidence and analysis to persuade Asian societies and the international community that action is needed, and that the potential benefits are

large. Good information will also be essential for making the right choices among varied and technically complicated alternatives. But all this is not enough unless it is also presented effectively to the public.

Third, international financial institutions and the external financing community will have to help support national policies toward the triple bottom line. They can do so by funding cooperation, disseminating knowledge, helping to counter vested interests, and championing international perspectives on transborder issues. Simple as these goals might seem, they will raise questions, from time to time, about the relationships of international agencies and their national partners.

This is an important turning point in Asia's development strategy. Asian priorities are changing, and when they have done so in the past, they led to extensive innovation and exceptional results. Asia has economic momentum, and its people confidently expect advances in their quality of life. The value that evidence-based approaches can contribute in this context cannot be overestimated.

## Notes

<sup>1</sup> These values are calculated from the databases of the World Bank, <http://data.worldbank.org/indicator> (accessed on 10 April 2013) and the International Monetary Fund, <http://www.imf.org/external/pubs/ft/weo/2012/02/weodata/index.aspx> (accessed 18 December 2012).

<sup>2</sup> This note later refers to these as net-win policies.

<sup>3</sup> Evidence-based policy making originated in the United Kingdom as part of an effort to improve the effectiveness of national policy decisions (Pawson 2006) In turn, some of its key concepts were borrowed from evidence-based medicine (Sackett et al. 1996).

<sup>4</sup> The government mobilized savings and used them to fund investments by large-scale enterprises. In the terminology of modern economics, it filled in for missing financial markets, a common market failure in emerging economies. It avoided the negative consequences of such interventions by monitoring the productivity of investments. It may have even amplified the consequences of market signals, such as export success, by using them to drive lending policy.



SAMBUTAN  
ULANGTAHUN  
KE-8  
BANDARAYA  
MELAKA  
BERSEJARAH  
2011



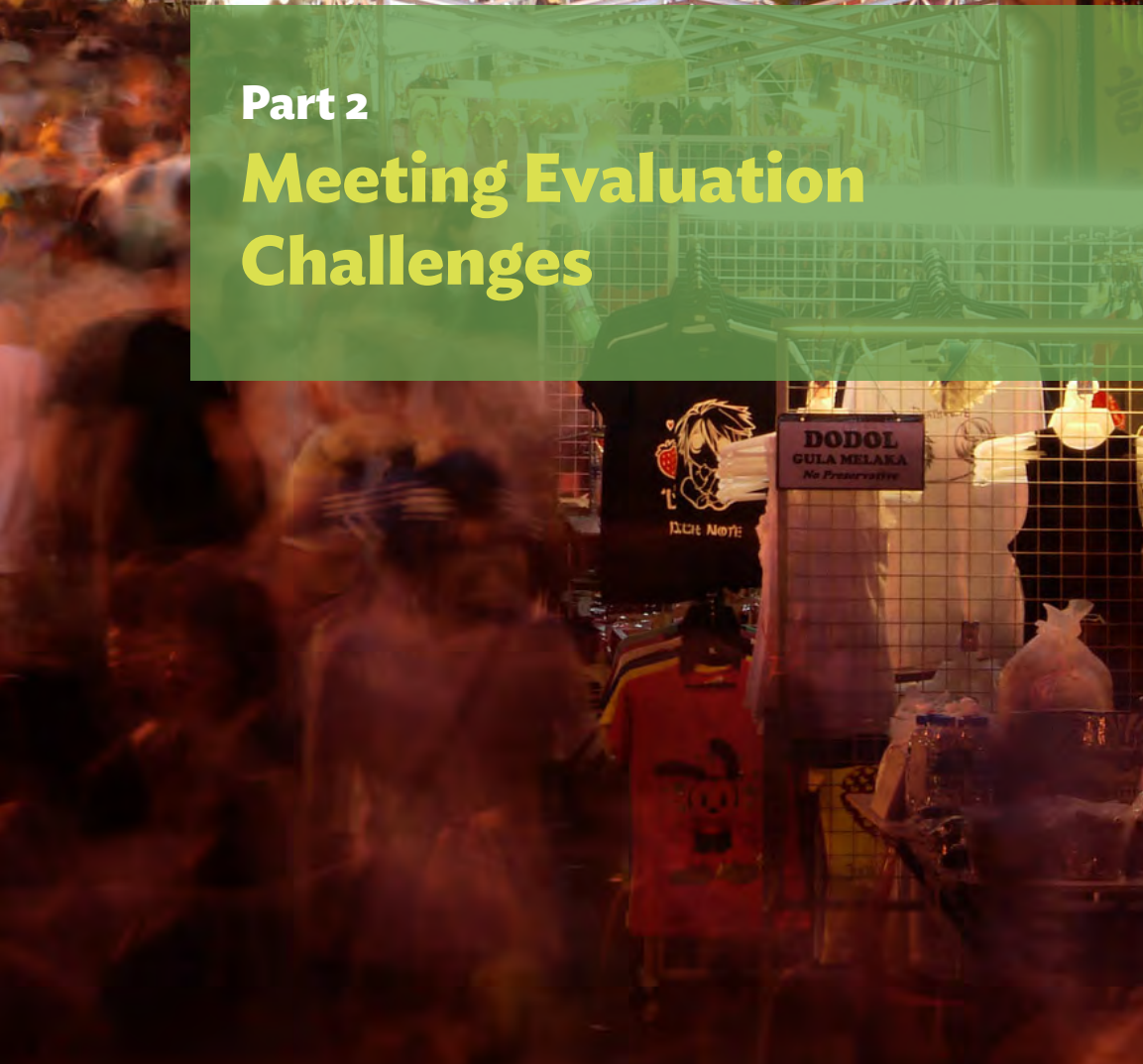
SAMBUTAN  
ULANGTAHUN  
KE-8  
BANDARAYA  
MELAKA  
BERSEJARAH  
2011

100



Part 2

# Meeting Evaluation Challenges



Independent evaluation offers a rich diversity of experience, and stakeholders in development, particularly those outside the evaluation field, can benefit from insight into its inner workings. Evaluation informs policies, strategies, and assistance programs; sheds light on lessons; and challenges conventional wisdom. And by seizing opportunities, it can make a difference in shaping decisions and improving development results. This view, offered in Chapter 4 by Zhongjing Wang, Development Effectiveness Committee (DEC) chair at the Asian Development Bank (ADB), encapsulates the goal of evaluation—to improve the results of development efforts—and opens this section offering candid opinions about the evaluation function from senior officials at ADB and the other major evaluators in global development.

Chapter 4 also explores the function of the DEC as guardian of independent evaluation. Its six members from the ADB Board of Directors began work in 2001. DEC has a general mandate to assist the Board in ensuring that ADB's programs and activities achieve development effectiveness.

The essence of independent evaluations, according to the Organisation for Economic Co-operation and Development (OECD 2002) is “carried out by entities and persons free of control of those responsible for the design and implementation of the development intervention.”

To this understanding, offered in Chapter 5, Vinod Thomas, current director general of independent evaluation at ADB adds: “However, independence should not translate into isolation from operations.” Failure to engage with operations staff and other stakeholders risks missing valuable opportunities. The chapter presents a history of evaluation at ADB, beginning in 1972 under the Economics Office, assuming its broad outline with establishment of independent evaluation in 2004, and subsequent strengthening in 2009. Through views offered from the directors general, past and present, it illuminates many challenges to independent evaluation.

Chapter 6, finally, looks at the twin goals of accountability and learning in improving development effectiveness. Both present opportunities for evaluators, but also tensions. Twelve senior evaluators from seven international financial institutions largely acknowledge the strains between these two functions. But in shining a light on the issues, they make clear that the functions are two sides of one coin, and can be balanced. The chapter offers views from officials at the African Development Bank, Council of Europe Development Bank, European Investment Bank, Inter-American Development Bank, International Fund for Agricultural Development, International Monetary Fund, and World Bank.



## Chapter 4

# Evaluation and Results in Asia

*Zhongjing Wang*

The Asian Development Bank's (ADB) efforts to shift the focus from inputs and outputs to outcomes and impacts is a welcome change for results orientation. For the move to improve operational performance, however, evaluators must help operations focus on and measure results, and provide lessons to improve them (Thomas and Luo 2012).

Independent evaluation at ADB helps various groups—the Board of Directors, Management, and external stakeholders—understand whether the organization has spent resources well and achieved planned outcomes. Independent evaluation covers all aspects of operations and emphasizes feedback on performance and use of lessons to improve ADB's development effectiveness.

Development Effectiveness Committee (DEC) meetings provide a vital channel for dealing with results-related issues. In particular, the DEC helps ensure that (i) the Board is aware of high-priority evaluation insights that have a significant bearing on the relevance, efficiency, and effectiveness of operations; (ii) ADB has an effective and efficient evaluation function; and (iii) lessons learned from evaluations are adequately reflected in operations (see Box 4.1 for a closer look at the DEC).

### **Identifying the results to focus on**

That the final results of a project or program are several steps removed from actual projects or policy interventions is always a challenge. It can make such results difficult to use as the final objective because the links in the chain leading up to them are hard to delineate or pinpoint. That said, for evaluation to be meaningful, its focus needs to be on the results of the actions that can be identified and make a difference to final outcomes.

#### Box 4.1: The Development Effectiveness Committee

The Asian Development Bank's (ADB) Board of Directors established the Development Effectiveness Committee (DEC) in December 2000, and the DEC began activities on behalf of the Board in 2001. It consists of not more than six members of the Board. The ADB President, in consultation with the Board, appoints DEC members for a term of 2 years and designates one as chair.

The DEC meets regularly throughout the year and has a general mandate to assist the Board in ensuring that ADB's program and activities achieve development effectiveness. For the DEC's work, development effectiveness is the measure of (i) whether ADB's program and activities have resulted in the desired outcomes, and (ii) whether these programs and activities have made efficient use of ADB's available resources.

DEC responsibilities include (i) overseeing the recruitment process for the director general of the Independent Evaluation Department (IED); (ii) advising IED in the preparation of its work program and budget and endorsing it to the Board for approval; (iii) reviewing all IED reports and discussing selected major reports, as well as Management responses to any report; (iv) monitoring and evaluating ADB actions taken on IED's recommendations; (v) reporting to the Board on selected development effectiveness issues that have a significant bearing on the achievement of ADB's overarching goal of poverty reduction, and making recommendations on such issues to the Board; and (vi) monitoring and reporting to the Board on the implementation of its decisions.

Through its annual reports, the DEC highlights key findings and conclusions on ADB policies, strategies, and completed operations, based on its review of various studies and discussions with Management and with evaluation staff. It also assesses evaluation activities based on completed deliverables and recommends future evaluation directions.

Overall, the DEC acts as the guardian of IED's independence. In concert with the Board as a whole, it seeks to ensure that (i) IED has an adequate budget to carry out its work; (ii) the quality of IED reports is adequate; and (iii) IED's recommendations are tracked and acted upon. However, the DEC does not interfere with the content and conduct of evaluation reports.

Sources: Development Effectiveness Committee. Various years. *Annual Report of the Development Effectiveness Committee*. Manila: ADB; ADB. 2011. *Terms of Reference of the Development Effectiveness Committee of the Board of Directors*. Manila; ADB. 2008. *Review of the Independence and Effectiveness of the Operations Evaluation Department*. Manila.

Several considerations are needed to ensure that this can take place. For example, immediate concerns must not crowd out what is important

for sustaining results. The response to natural disasters provides a case in point. After calamity strikes, speedy reconstruction is necessary, but rebuilding hastily without complying with building standards can lead to more losses when disasters strike again. The challenge is to rebuild structures such that they can better withstand natural hazards. This understanding is crucial given the increasing frequency and intensity of natural disasters around the region (IED 2013b).

Evaluators also need to recognize the factors that contribute to outcomes and not just to the outputs. Educational interventions make this clear. In these, building classrooms and improving access for children to go to school are important, but equally important are the learning outcomes that draw on the quality of the spending on teacher training, the relevance of curricula, and learning methods. Simply raising enrollment rates without consideration for these issues can mean that learning outcomes and the employability of graduates do not improve. Clearly, the ultimate goal is not just school attendance, but better learning and higher earnings from employment.

Some issues go beyond sectors and require broad-based government and societal action. Reducing carbon emissions from transport projects calls for sound urban planning and management that integrates transport, urban design, land use, and environmental protection (IED 2010a). Doing projects right is an important part of achieving outcomes, but factors beyond projects, such as policies, legislation, integrated planning, and law enforcement, can affect results. How different interventions come together is crucial for evaluating development results.

### **Measuring results**

Measuring results is often a big challenge. The process of developing good databases is a natural ally of evaluation. And evaluative questions, in turn, provide a rich basis for the development of data, as do the evaluation studies.

For the presentation of results, data very often need to be summarized into indicators. At the same time, care is needed in developing and interpreting summary measures. Because they are simple, composite measures are appealing. But these indicators can misdirect if their underlying premises are not transparent.

Arbitrary clustering and weights and converting ratings to rankings in surveys can lead to unclear findings. For example, a limitation of the World Bank's Doing Business indicator lies in its partial focus on the cost to business of having regulations (Manuel 2013). It suggests that businesses are more likely to flourish if regulations are fewer and simpler. However, it neither examines the pertinent aspects of regulations nor looks at their benefits, such as safety, environmental protection, and transparency. Thus, the information user must interpret composite indicators cautiously and collect supplemental reports to better assess the situation.

Likewise, looking only at averages risks missing key constituents if the effectiveness of an intervention differs across the subgroups of a population. For example, the Millennium Development Goals measure progress toward eliminating poverty, but they use average indicators, and higher average income can coexist with lower income for the poorest population segments (Thomas and Luo 2012). Thus, the very poor may have remained so or even become poorer even as a better average for the entire population has been achieved. Focusing only on the average will miss this fact.

Another issue in measurement is the need to align intermediate and final goals to examine development results. An exclusive focus on intermediate objectives, such as building road connectivity, can come at the expense of achieving the final goal of, say, inclusive growth. While intermediate outcomes are often easier to monitor, they may not lead to the desired final results when crucial links are not carefully investigated. Actually achieving inclusive growth demands action on several fronts. Given widening income inequality in many Asian countries, it is essential to balance infrastructure development with complementary investments and partnerships in education, health and social protection, environment, and agriculture (IED 2012a). Efforts to strengthen these investments through reforms in policies, administration, and governance to improve development results—and to track the results—are equally important (IED 2011).

### **Evaluation lessons to improve results**

Lessons are typically drawn from past experiences that add up to tell a consistent story, and independent evaluation at ADB puts together such lessons from operational experiences. Some lessons drawn from other countries are often valid and have value for other countries,

and the challenge here is to make knowledge exchanges effective and timely. This means that the application of evaluation lessons must be contextual both for the country and the time.

That said, development uncertainties and complexities also suggest that in changing environments, applying past findings to future efforts may not be enough. When contexts change, approaches to development issues may also need to change. Growing development complexities call for flexible evaluation strategies that focus on underlying conditions and adapt to dynamic situations, and use innovative and creative approaches to meet future needs.

Evaluation needs to identify opportunities in operational work—a form of counterfactual analysis. It looks not only at what happened, but also at what might have happened, and what could have made a difference if opportunities had not been missed. Going beyond conventional wisdom and looking at situations with a fresh eye could make a difference. In the dynamic Asia and Pacific region, ADB's challenge is to anticipate these opportunities, a process that evaluation needs to feed into.

ADB adopted Strategy 2020 in 2008 to establish its long-term strategic directions (ADB 2008b). Changes in the development landscape prompted a review of Strategy 2020 in how it has been implemented and what might be done to strengthen it. This review provided an opportunity for evaluation to bring to the table suggestions on what might be done differently through mid-course corrections (IED 2014).

## **Conclusion**

The demand for better evaluation results is high, particularly when resources are scarce and when pragmatic solutions to development issues are crucial. Evaluation must aim to get results, measure them, and contribute to learning to improve them. Evaluation for results increases the importance of looking beyond the stated objectives to find synergies and cross-cutting solutions to complex development issues.

Evaluation plays a valuable role in many ways. It informs policies, strategies, and assistance programs; sheds light on lessons; and often challenges conventional wisdom. By seizing opportunities and

synergies, evaluation can make a difference in shaping decisions and improving development results.

Over the medium term, independent evaluation at ADB will devote more effort to higher-level evaluations of value to the Board, Management, and external stakeholders. It will be forward-looking and responsive to current issues and emerging needs in the region. In light of the changing development situation in Asia, evaluation for results will be valuable in steering ADB operations toward greater development effectiveness.

## Chapter 5

# The Journey to Independent Evaluation at ADB

*Drawing on contributions from directors general over the past 10 years*

This chapter traces the main events leading to the independence of the Asian Development Bank's (ADB) evaluation function in 2004 and its strengthening in 2009. Along the way, the evaluation focus shifted significantly from the project level to sectors, policies, strategies, and themes.

The directors general of independent evaluation at ADB over the past 10 years were Bruce Murray (2004–2007), Halady Satish Rao (2008–2011), and current director general Vinod Thomas (appointed 2011). In this short period the evaluation department transformed its approaches to make evaluations more credible and support effectively learning and accountability. It implemented changes to increase transparency and it made knowledge management an increasing part of the evaluation lexicon, with a unit set up to make evaluation findings accessible to a broad audience in the development community and beyond. This process continues as the Independent Evaluation Department (IED) at ADB aims to increase its relevance and effectiveness in a rapidly changing development environment.

The journey in this chapter's title is not finished, however. In examining the dimensions of independent evaluation, contributors draw attention to the areas they see needing further change. ADB's experience with independent evaluation offers insights into this complex and sometimes difficult process to other organizations seeking to strengthen their evaluation functions.

An equally important aim of the chapter is to examine how evaluation has evolved at ADB. Within little more than a decade, the evaluation process underwent a sea change, advancing it from largely ex post assessments of programs and projects to a varied set of prospective and real time tools and innovations in evaluation methods. Evaluation work is now paying greater attention to the role

links play in strengthening evaluation results, involving projects and programs in different areas and the work of different players. Again, this demonstrates a markedly broader development perspective than in the early days of independent evaluation. Indeed, a thread that runs through many evaluative findings in today's IED is the need to connect these dots during project design and implementation for better development results. The need for these has never been stronger at ADB as it confronts newer development challenges, such as inclusive growth and environmental sustainability, to make further headway in the unfinished business of eradicating extreme poverty in the region.

### **Beginnings**

Evaluation at ADB began in 1972 when it reviewed project history and performance. Both the art of evaluation and its practice at ADB and other multilateral development institutions have come a long way since then. Over the years, evaluation at ADB has stepped up to the challenge of providing forward-looking and evidence-based recommendations and insights to inform decision making and shape directions. To see how far ADB has come in this area, consider these evaluation processes. Early evaluation work concentrated on input-output relationships, particularly in investment projects, to assess whether the actual socioeconomic benefits of completed projects were in line with those expected at appraisal (ADB 2008a). This led initially to postevaluation reports and later to project performance audit reports. New evaluation products emerged in the 1980s, with the first impact evaluation study done in 1984 and the first evaluation of a program loan in the following year (Operations Evaluation Department [OED] 2007).

A 1997 survey of ADB's operations staff to assess the relevance and extent of use of evaluation findings led to the then Evaluation Office conducting more thematic studies of operational relevance and expanded country and sector syntheses. A particularly important step was the office starting to review multiple operations rather than, as before, projects and programs individually. The first country assistance program evaluation, on the People's Republic of China, was conducted in 1998. Another milestone of that period was a special evaluation study on ADB's midterm review process for project lending to draw lessons and recommend steps toward improving the effectiveness of the review process.



After independence of the evaluation function in 2004, there was a marked shift to higher order evaluations and a reduced emphasis on projects and programs. From then on, evaluation has focused on a mix of lending and nonlending operations, with increasing attention to relevance, efficiency, effectiveness, sustainability, and impact, which the evaluation community has established as the five main evaluation criteria. Early interactions with stakeholders, use of multiple quantitative and qualitative tools, better dissemination of findings, and the follow-up on evaluation recommendations have led to a more persuasive evaluation function. But the most remarkable aspect of the journey has been strengthened independence.

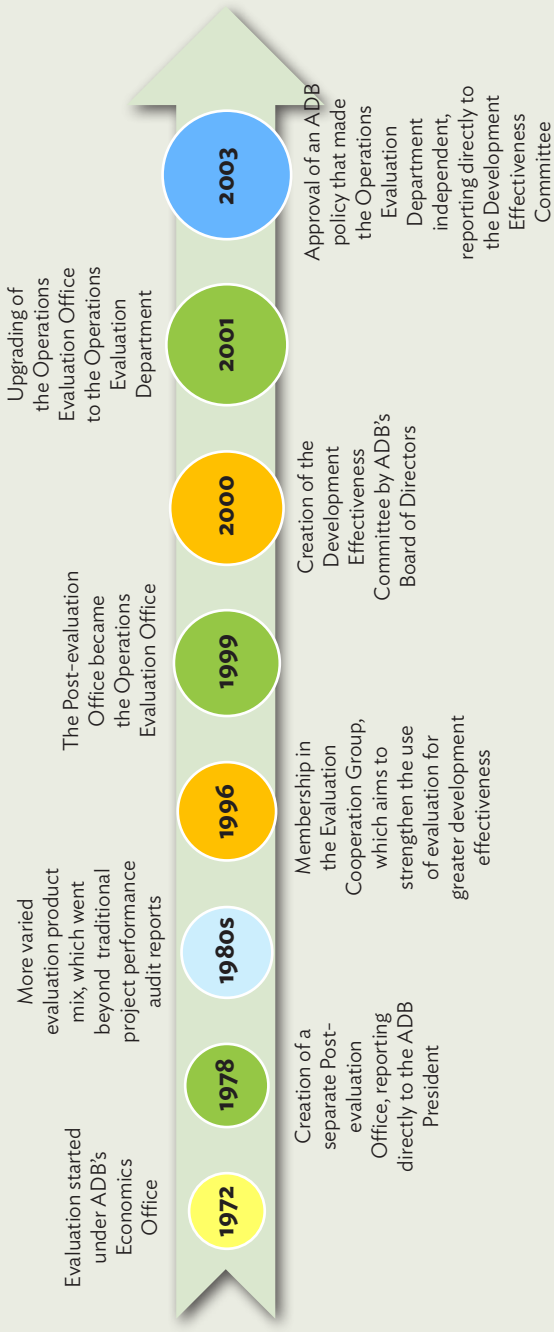
### **Before evaluation independence in 2004**

When ADB was established in 1966, evaluation was neither part of its structure nor its initial activities. Figure 5.1 outlines the beginnings of evaluation in 1972 under the Economics Office, which provided the internal evaluation capacity with the assistance of consultants, and the subsequent operation of a Postevaluation Office in 1978. Selective postevaluation reports were discussed at the Board's Audit Committee.

The progress of evaluation at ADB is closely linked to the increasing importance of operational evaluation within the institution and growing concern in the international development community about the need to improve development effectiveness. Both factors led to the upgrading of the Evaluation Office to the OED in 2001. OED was tasked to help ADB management and decision makers in developing member countries to assess whether resources were well spent and planned outcomes achieved (ADB 2003b).

Donors raised the issue of independent evaluation in 2000 during the seventh replenishment of the Asian Development Fund (ADF), ADB's concessionary lending facility to its poorest member countries. During the replenishment negotiations, they stressed that they expected ADB to be a model for institutional governance, reflecting international best practices in transparency, accountability, participation, and predictability. They encouraged ADB to implement measures to improve its internal governance, including examining the possible strengthening of its evaluation office. Donors emphasized that timely access to evaluation results on ADF-financed operations is necessary to enable them, Management, and the Board to assess

**Figure 5.1: Before evaluation independence in 2004**



ADB = Asian Development Bank  
 Sources: ADB. 2003. *Enhancing the Independence and Effectiveness of the Operations Evaluation Department*. Manila; Operations Evaluation Department. 2007. *Independent Evaluation at the Asian Development Bank*. Manila: ADB.

the stewardship of ADF resources. They felt that existing evaluation systems limited ADB's ability to fully assess the results and impacts from ADF-financed operations and asked ADB to continue its efforts to strengthen such systems to allow for better measurement of impacts (ADB 2000b).

In December 2000, ADB's Board established a new standing committee, the Development Effectiveness Committee (DEC), to assist the Board in ensuring that ADB's programs and activities achieve their desired development objectives and make efficient use of ADB resources. The DEC was charged with reviewing the annual work program of OED, the annual report on evaluation activities, and other selected evaluation reports.

The need for more autonomy in evaluation was reiterated in the ADF VIII midterm review in 2003 (ADB 2003a) and the first meeting of the ADF IX replenishment in October 2003 (ADB 2003c). In its organization structure, OED was subject to the ADB President's oversight. However, its de facto dual accountability, both to the President and the DEC, provided it with a balance between behavioral autonomy and important leverage on the operations departments. Nevertheless, to strengthen the organizational independence of OED, donors recommended that OED should report directly to the Board through the DEC.

In November 2003, the Board approved major changes in OED operations to provide greater credibility in assessing ADB operations and ensure evaluation independence (ADB 2003b). These changes were meant to bring OED to a comparable position with the evaluation units in other development banks. The most significant changes related to OED reporting directly to the Board through the DEC, instead of to the ADB President, and the appointment of OED's director general by the Board rather than the President. The appointment was a joint recommendation by the DEC and the President, with an initial term of 3 years, which could be renewed for a maximum of 2 years.

Changes were also made to staff hiring and work programs, with the director general, in consultation with the Budget, Personnel, and Management Systems Department, responsible for the selection of new OED personnel. A separate resource envelope for OED as part of ADB's overall administrative budget was created, and the preparation

of the department's work program was done in consultation with the DEC and the President. For transparency, changes were made to procedures for the disclosure of OED reports.

### **The early years of independent evaluation (2004–2008)**

Independent evaluation at ADB began on 1 January 2004. With independence, the reporting relationships and roles among the Board, Management, and OED changed. In particular, Management's role shifted from approving evaluation reports to responding to the reports' conclusions. Management responses gradually evolved from a "no comment" nature at the outset to carefully considered commitments to specific improvements and actions, in consultation with relevant departments. Box 5.1 highlights these early years.

Recognizing the need for greater interaction between evaluators and operational staff, OED implemented the following:

- seeking comments on draft evaluation approach papers for the first time,
- soliciting greater consultation on the work program,
- encouraging evaluation officers to interact with operational staff during the evaluation process and to share preliminary findings and recommendations,
- conducting meetings at the director-general level for higher-level evaluations to discuss the executive summary and recommendations and similar meetings at the director level for project performance audit reports,
- improving consultations with regional departments and the Strategy and Policy Department to get the timing right so that evaluations were available in a timely manner to help inform new strategies and policies, and
- undertaking missions to countries to discuss government comments on draft country assistance program evaluations and the feasibility of draft recommendations.

### Box 5.1: The early years of independent evaluation (2004–2008)

The Development Effectiveness Committee (DEC) boosted independent evaluation through several steps in the early years:

*Preparing annual reports to the Board on the activities of the DEC.* The DEC annual reports were prepared with input from the Operations Evaluation Department (OED). These early reports showed a noticeably weak link in the evaluation chain—follow-up action on agreed evaluation recommendations. Although manually tracking follow-up action dates back to 1982, the DEC’s concern on this resulted in the launch in 2008 of a computerized tracking system, MARS (short for management action record system).

*Introducing Management’s response to evaluation reports.* The DEC required that Management prepare written responses on all independent evaluation reports. In 2004, it expressed concern about mainstreaming the Management response. The new dynamics of interaction among OED, Management, and the DEC were formalized. Management responses became important documents that the DEC considered when it reviewed evaluation reports.

*Ensuring that evaluation findings were considered in the formulation of country strategies.* To strengthen feedback from evaluation, the DEC agreed with Management that new country strategies would only be finalized after the previous strategy had been evaluated and considered by the DEC. By 2006, the influence of country assistance program evaluations on the formulation of new country partnership strategies had been mainstreamed, although it was not possible to do country evaluations for all countries. In lieu of these, a report validating the self-evaluation of country strategies started in 2007. Validation is done if the evaluation work program does not include a country assistance program evaluation.

The DEC also required OED to do a prior review for new and revised policies. For example, it recommended that an independent evaluation of the environment, resettlement, and indigenous people’s policies precede the formulation of the new safeguard policy statement. In response, OED advanced the evaluation of safeguard policies from 2008’s work plan to 2006. The DEC also requested that the preparation of the new policy statement be delayed for 9 months to enable evaluation findings to feed into Management’s policy review.

Sources: ADB. 2013. *Operations Manual on Independent Evaluation*. Manila; Independent Evaluation Department (IED). 2005. *Annual Evaluation Review*. Manila: ADB; IED. 2006. DEC Chair’s Summary of the Committee Discussion on 27 September 2006. In the *Special Evaluation Study on Environmental Safeguards*. Manila, ADB; IED. 2011. *Annual Evaluation Review*. Manila: ADB.

Efforts were also made to increase interaction between evaluators and operations staff to confirm the reliability of data and to exchange views on evaluations. However, the final decisions on the conclusions and recommendations in evaluation reports remained firmly with OED, a key tenet of independent evaluation.

OED also made significant changes to its evaluation products and services. The following paragraphs discuss some of the main changes.

**Shift in the product mix.** The introduction of a much wider consultation process was valuable in identifying topics for evaluation. Communiqués were sent to operations departments requesting their inputs to the evaluation work program. Evaluation staff synthesized the feedback and held individual discussions with vice-presidents, members of the DEC, ADB's dean of the Board, and other interested Board members.<sup>1</sup> Based on these, a 3-year rolling work program was developed and finalized after discussions with the DEC. As a result, the composition of the evaluation work program shifted to strategy, policy, thematic, and sector evaluations. In 2007, OED introduced the project completion validation report, with the validation process providing a more independent view of project performance documented in the project completion reports.

**Strengthened evaluation recommendations.** Given the importance that the DEC placed on tracking action taken on evaluation recommendations, it was clear that steps had to be taken to (i) ensure that recommendations followed from the evidence presented; (ii) increase the dialogue between evaluation officers and operations staff to make recommendations more feasible; (iii) identify which department was to take action on recommendations and by when; and (iv) for complex evaluations, hold discussions on evaluation findings at the director-general level after comments had been considered in the report.

Traditionally, evaluations at ADB were conducted to assess accountability for development results. Central to this concern was whether ADB was doing the right thing, that resources were properly allocated and used, and that intended outcomes were achieved. One measure of accountability is the success rate. Country assistance program evaluations and validation of the self-evaluation of completed country partnership strategies, programs, and projects have been instrumental in assessing accountability. The annual

evaluation review presented analytical aspects, and a view of ADB's operational performance. Another measure of accountability was to see how ADB followed up on OED recommendations. The automated MARS launched in 2008 monitored this, initially, with yearly stand-alone reports on Management's actions on recommendations. In 2011, the report on actions taken was consolidated into one report together with the annual evaluation review. Box 5.2 provides details on MARS.

#### Box 5.2: Management action record system

- Under the management action record system, evaluators enter each recommendation agreed to by Management and, in a separate section, those not agreed to by Management.
- All evaluation recommendations agreed to by Management are tracked together with the associated action plans; those not agreed to are not tracked but kept on file. Management, in collaboration with concerned departments, is responsible for monitoring actions taken in response to evaluation recommendations and for recording implementation progress in the system.
- Management acceptance of evaluation recommendations improved from 92% during 2008–2011 to 96% during 2010–2013. For 2013 evaluations, Management accepted 90% of recommendations.
- The validation of completed actions shows that fully or largely adopted actions averaged 72% of completed actions during 2010–2013. Partial compliance with and nonadoption of recommendations are due to limitations in resources; changes in business processes, policies, or priorities; and open-ended targets suggested in the recommendations.

Source: Independent Evaluation Department. 2014. *Annual Evaluation Review*. Manila.

**Knowledge management.** A knowledge management unit was established in 2007 and new knowledge products rolled out (for example, *Learning Curves*, a 2-page reference on the findings and recommendations of evaluation reports). A searchable platform was developed making it easier to access evaluation lessons. The previous postevaluation system did not store evaluation lessons and recommendations in an easily accessible way. But with continuous upgrading of the evaluation information system since 2007, users can search for their preferred combination of parameters in the evaluation information system database. The knowledge management unit ushered in better synthesis of findings, outreach, and improvements in the dissemination of evaluation products through various web links.

**Increased transparency.** During the first year of independence, a policy of making evaluation reports public only after discussions with the DEC undermined credibility with some nongovernment organizations (NGO). This was thrown into sharp focus in 2004 by OED's evaluation of the Nam Leuk Hydropower Project on the Mekong River, which was of considerable interest to NGOs at that time. Many international NGOs requested copies of the project performance audit report. In line with ADB policy, this was denied because the DEC had not yet discussed the report. This caused a stir among concerned NGOs, which suspected that something was going on behind the scenes to modify the report. This suspicion undermined the perception of *independent* evaluation at ADB in the minds of some NGOs. To prevent future misunderstandings, evaluators worked with the Office of External Relations to insert a clause in ADB's public disclosure policy that evaluations would be made public when they were circulated to Management and the Board.<sup>2</sup>

Evaluations were deemed final when the director general of independent evaluation approved them, and were not changed based on Management's response and the discussion at the DEC. The chair's summary of the DEC meeting came out after the evaluation report had been disclosed to the public. If the committee disagreed with findings, it could say so in the chair's summary, which was also made public. One of the DEC's key responsibilities is to ensure that lessons from evaluation are reflected in ADB's policies, procedures, and operations, as well as to monitor the implementation of evaluation recommendations. In this context, the DEC can provide guidance to ADB Management on what actions needed to be taken, based on evaluation findings. That OED had the freedom to disclose evaluation findings was another significant milestone toward independent evaluation at ADB.

Following approval of the 2003 evaluation policy, OED's standing increased, as did the DEC's reliance on OED evaluation studies, which increased in number. However, certain areas of responsibility (including the manner of appointment and term of the director general, OED personnel decisions, its budget preparation, and the performance evaluation of the director general) were not sufficiently stated or delineated in the policy, which led to some difficulties during implementation.



Management's decision not to support the DEC's recommendation to extend the term of the incumbent director general in October 2007 led to tensions between the Board and Management regarding the role of the ADB President in appointing the head of evaluation and raised concerns both internally and externally about OED's independence, just as the ADF X replenishment negotiations were about to start. To tackle these concerns and ensure that the evaluation process within the institution remained impartial and credible, ADB committed in February 2008 to undertake a comprehensive review of its evaluation function. A working group of three Board directors and one Management representative was constituted to review the 2003 evaluation policy. They, in turn, appointed two independent consultants to assist in the review.<sup>3</sup> In the meantime, the second Board-appointed director general took office in April 2008.

### **Strengthening independent evaluation (2009–2014)**

As a result of the review, OED was renamed the Independent Evaluation Department in 2009 to reflect its enhanced independent status. The review culminated in the adoption of measures to strengthen the independence of the evaluation function, among them:

- the appointment of IED's director general by the Board upon the recommendation of the DEC in consultation with the ADB President,
- expanding the nonrenewable term of the director general from 3 to 5 years,
- the selection process for director general to be led by the DEC and strengthened through the use of an executive search firm and advertising the post in the media,
- the director general to be exempted from the formal performance review process and his or her salary to be determined by a formula, and
- IED's budget to be approved by the board separately from ADB's overall administrative budget (ADB 2008a).

Probably the most important of these changes was the appointment process for the head of IED: Management and the DEC would no longer jointly recommend the appointment to the Board, with the sole authority for this being vested with the DEC, albeit in consultation with Management. Box 5.3 shares some reflections from former director general Bruce Murray.

**Box 5.3: Bruce Murray, director general, 2004–2007**

Murray initially viewed independence as largely a state of mind. As long as the director general and Operations Evaluation Department (OED) staff were sound professionals with independent, enquiring minds, their evaluations would be independent and objective. However, he now realizes that this was a misconception. With independence, the reporting relationships and roles of the Board, Management, and OED changed. As a result, evaluation findings and recommendations were taken more seriously.

Looking back, Murray identified six main achievements during the early years of independence: (i) changing product mix with more emphasis on strategy, policy, thematic, and corporate evaluations, which were of greater strategic interest to the Development Effectiveness Committee (DEC) and the Board; (ii) strengthened evaluation recommendations—OED was open to suggested changes to make recommendations more feasible and took steps to improve the clarity and quality of recommendations and to reduce their numbers; (iii) better knowledge management and outreach—OED staff realized that producing more reports was not the way to make OED more effective and influential; (iv) strengthened relations with the nongovernment organization community; (v) increased transparency with the disclosure of OED reports upon circulation to Management and the Board; and (vi) greater attention to corruption issues in evaluation reports.

Murray also recalls various challenges. Independent evaluation departments in multilateral development banks are sometimes at odds with management. And when that happens, he says, boards “must defend the independence of the evaluation function.” For example, disagreements between boards and presidents over the appointment and renewal of heads of evaluation happen. “It is probably part of the normal growing pains as people get accustomed to their new roles and come to grips with what independence really means,” says Murray.

He notes with approval that the Board took firm action to preserve and strengthen the independence of ADB’s evaluation function by revising and strengthening ADB’s evaluation policy. That said, Murray sees a remaining question over the role of the ADB President in appointing the members of the Development Effectiveness Committee and its chair. “In other multilateral development banks, the boards and not the presidents decide the composition and chairs of board committees,” he says.

The essence of independent evaluation as defined by the Organisation for Economic Co-operation and Development is that it is “carried out by entities and persons free of control of those responsible for the design and implementation of the development intervention”<sup>4</sup> to enable impartial evidence-based evaluations that are more credible. In ADB’s case, “free of control” means independence from management and is possibly the most important factor that distinguishes independent evaluation from management’s self-evaluation. Box 5.4 presents the dimensions of independent evaluation at ADB.

#### Box 5.4: Dimensions of independent evaluation at ADB

**Behavioral independence** refers to the extent to which the evaluation unit is able and willing to produce uncompromising reports and to disclose its findings to the Board without Management-imposed restrictions. The Independent Evaluation Department (IED) has such independence. The Board, Management, and staff may comment, but IED ultimately decides which comments to incorporate. IED maintains a high degree of transparency and openness in reporting and disclosing its evaluation findings. Its reports are publicly disclosed on the Asian Development Bank’s (ADB) website as soon as IED’s director general approves them, and before they are discussed by the Development Effectiveness Committee.

**Organizational independence** requires that the evaluation unit and its staff are not under the control or influence of decision makers who have responsibility for the activities being evaluated, and that the scope of evaluation covers all relevant aspects of ADB. IED has such independence. By reporting directly to the Board through the Development Effectiveness Committee, IED is located outside the line functions that it is tasked to evaluate.

**Protection from outside influence** refers to IED’s ability to decide on the design, conduct, and content of evaluations without interference; its access to adequate resources to carry out its mandated responsibilities; and its control over human resources decisions. ADB has set up several safeguards in this area. The Board approves its budget separately from the ADB-wide administrative budget. It also appoints IED’s director general, who can only be removed by the Board on the grounds of inefficiency or misconduct. The selection of IED directors involves Management approval (through ADB’s vice-presidents). For international staff, recruitment is subject to the approval of IED’s director general, the director general of the Budget, Personnel and Management Systems Department, and an independent panelist.

**Avoiding conflict of interest** means that current, immediate future, or prior professional or personal relationships and considerations are not allowed to influence the evaluators’ judgments or create the appearance of a lack of objectivity. IED has such independence. Since 2012, it has implemented guidelines to avoid conflict of interest that are well thought out, sufficiently

stringent, and appropriate. These guidelines require evaluators to inform IED management of any potential conflict of interest or potential perception of conflict of interest, before their assignments are finalized. Evaluators must recuse themselves from any project, program, or activity that they worked on or in which they had line responsibility. Similarly, evaluators must recuse themselves if they had a personal influence or stake in a project, either in a previous capacity at ADB or prior to joining ADB. They also may not evaluate an entity in which they had a significant decision making role or had a financial stake in prior to joining ADB, or in which their future employment is a significant possibility.

Sources: N. Fostvedt. 2012. *Consultant Report for the IED Self-Evaluation*. Manila:ADB; Independent Evaluation Department. 2012. *Self-Evaluation Brief*. Manila: ADB.

Box 5.5 presents reflections from former director general Halady Satish Rao on this period.

#### Box 5.5: Halady Satish Rao, director general, 2008–2011

Rao remembers the Operations Evaluation Department becoming the Independent Evaluation Department (IED) in 2009 “as a fitting proclamation and assertion” of the department’s strengthened independence. Yet, for IED staff seeking career advancement in the Asian Development Bank (ADB), impartiality of evaluations could be susceptible to Management influence, real or perceived. One approach to deal with this issue could be ending rotation of IED staff to ADB and recruiting staff mainly from outside of ADB. Rao, however, feels that this may lead to a “disconnect with ADB and isolation of IED staff,” and hence prefers a healthy interchange of staff with stronger safeguards to ensure evaluations without fear or favor.<sup>a</sup>

Rao believes an even broader debate could at some point emerge at ADB on whether to go beyond independence from Management and reduce or even do away with Board oversight of the evaluation unit, as the International Monetary Fund has done. Rao thinks this is unlikely to happen at ADB, but as evaluation moves toward higher-level topics—for example, an evaluation of the Board itself—the issue of IED’s independence could conceivably arise. “If this were to ever happen,” says Rao, “the question would be whether IED can operate in isolation with no one to champion its cause or take care of its interests.” Indeed, he adds, an even more important question is who IED would be accountable to in that event.

“Independence has been rightly strengthened, but correspondingly, responsibility and the accountability of evaluators need more attention,” says Rao.

-----  
<sup>a</sup> The IED guidelines to avoid conflict of interest subsequently dealt with this issue. See IED. 2012. *Guidelines to Avoid Conflict of Interest in Independent Evaluations*. Manila:ADB.

Since the strengthening of evaluation independence in 2009, there have been significant changes in evaluation activities.

**Strategic and relevant work program.** To sharpen the evaluation focus on ADB's priorities, IED aligned its 3-year work programs with Strategy 2020's development agendas—inclusive growth, environmentally sustainable growth, and regional integration. IED is increasingly looking at the consistency of ADB operations with Strategy 2020, and pursuing more rigorous evaluations and real time feedback applicable to the design and implementation of operations. The goal of improving development effectiveness and future impacts receives greater attention.

**Balancing independence and engagement.** To strengthen its operational relevance and learning function, IED intensified efforts to find the right balance between evaluation independence and engagement with operations staff and other stakeholders. It has pursued building rapport and engaging with stakeholders through discussion and learning events to heighten the sensitivity of evaluators to operating contexts and to help validate the feasibility of proposed recommendations. Indeed, some of IED's most influential evaluations (see Chapter 8 of this volume on influential evaluations) were guided by evaluator engagement with stakeholders at various stages of the evaluation process while maintaining analytical rigor and independence. Box 5.6 shares reflections from Vinod Thomas.

**Good evaluation practices.** Participation in international evaluation networks such as the Evaluation Cooperation Group and the network of the Organisation for Economic Co-operation and Development's Development Assistance Committee has helped IED to adopt good evaluation practices. Being in step with these practices is essential for strengthening the use of evaluation for learning and accountability. IED led the preparation of the Evaluation Cooperation Group's *Big Book of Good Practice Standards*<sup>5</sup>, a living compilation of good practice standards in evaluation begun in 2012; a synthesis report on microfinance, *Making Microfinance Work: Evidence from Evaluations*, published in 2010; and the *Good Practice Standards on Country Strategy and Program Evaluation*, published in 2008. It also contributed to the drawing up of good practice standards for the evaluation of sovereign and private sector operations in the same year, which have all been folded into the *Good Practice Standards Big Book*.

### Box 5.6: Vinod Thomas, director general, 2011 to date

For Thomas, balancing independence and engagement is a continuing challenge to evaluation. “Evaluation must be free from external pressure if it is to provide evidence to support learning and accountable decision making. However, independence should not translate into isolation from operations. Without engagement with operations staff and other stakeholders at various stages of the evaluation process, we will miss invaluable opportunities to interact on complex issues, understand unfolding situations, and make implementable recommendations,” he says.

“The process followed in the mid-term review of Strategy 2020—the Asian Development Bank’s (ADB) long-term strategic framework—provides a recent example of balancing independence and engagement,” says Thomas. Evaluation staff prepared the report in parallel with Management and engaged with ADB’s communities of practice, governments, development partners, and external peer reviewers. Meaningful interactions helped deepen insights into how ADB could sharpen its operational focus and better service the needs of its client countries in Asia and the Pacific.

“With respect to evaluation staff, I support the principle of rotation to and from other parts of ADB due to the importance of maintaining the right balance between insiders and outsiders,” he says. This flexibility in staff movement between the Independent Evaluation Department (IED) and parts of ADB, as is the practice in the World Bank and other multilateral development banks, should help guard against the potential isolation of IED staff and enrich both operations and evaluation activities through cross-fertilization of knowledge and experience. IED has adopted strict guidelines to avoid potential conflicts of interest.

Thomas thinks that striking a balance between independence and engagement is not typically smooth. Yet, it brings in benefits and boosts the prospects for evaluation to make a difference. “Continuing to put a premium on balancing independence and engagement will be essential. This way, we have a chance to optimize opportunities for rigorous assessments and foster processes for innovation and learning to improve development effectiveness,” he says.

**Evaluation guidelines.** Because the reliability of evaluation results is crucial to shaping policies, strategies, and programs, IED continues to refine its evaluation guidelines and prepare new ones. In 2010, for example, it released revised country assistance program evaluation guidelines. In line with international evaluation practice, it issued new guidelines on avoiding conflict of interest in 2012 and, in 2014, worked on guidelines for preparing project performance evaluation reports on nonsovereign operations. IED is active in the Evaluation

Cooperation Group's discussions on updating its guidelines. In the evaluation of Strategy 2020 at midterm, IED affirmed its commitment to encouraging innovation and new solutions through its guidelines.

**Evaluation capacity development.** IED has continued to support the Shanghai International Development Program for Development Evaluation Training, launched in 2007 to help ADB's client governments develop results-based monitoring and evaluation systems. It did this in partnership with the World Bank's Independent Evaluation Group and the Ministry of Finance of the People's Republic of China and its Asia-Pacific Finance and Development Center.<sup>6</sup> IED also helped fund the Regional Centers for Learning on Evaluation and Results Program, a 3-year program (2010–2012) supported by multiple donors.<sup>7</sup> IED's contribution went to the regional center for South Asia.

In 2011, IED introduced a 5-month program to provide on-the-job training in evaluation methods and approaches to government evaluators to institutionalize results-based monitoring and evaluation. As of 2013, officials from Bangladesh, Bhutan, the People's Republic of China, Indonesia, the Maldives, Nepal, the Philippines, and Sri Lanka have completed the program.<sup>8</sup> IED started a 2-year training program in 2014 for executing and implementing agencies of ADB projects, which involves the preparation of project completion reports and technical assistance completion reports to better equip these agencies for conducting evidence-based self-evaluation. For ADB staff, IED has conducted training sessions in preparing completion reports for projects and technical assistance since 2009. It also introduced an evaluation module for program design and management, launched in 2011.<sup>9</sup>

**Improved dissemination of evaluation results.** Since the strengthening of independent evaluation in 2009, evaluation for learning has gained considerable mileage, an encouraging development because this is integral to the search for strategic and innovative ways to improve operations. With the growing importance of building knowledge economies, IED is making its work known to a far broader audience in the development community, in client countries, and national and international media—and doing so with knowledge products that come in easily digestible formats.

As well as its evaluation assessments on ADB's strategies and operations, IED produces forward-looking products. These are *Evaluation Knowledge Briefs*, which, among other things, have looked at the greenhouse gas implications of ADB's energy sector operations and ADB's response to the global financial crisis; and Topical Papers, a new product that deals with prominent development issues. *Learning Lessons* is a 4-page publication synthesizing lessons from evaluation and completion reports and from literature reviews. Recent *Learning Lessons*, and the shorter *Learning Curves*, have covered ADB's response to natural disasters and the vital role links play in the patterns of development. *Evaluation Knowledge Studies* were started in 2012 in response to the specific interests of ADB's regional departments and the Regional and Sustainable Development Department; the series was launched by a study on agricultural value chain development.

## Conclusion

A decade has passed since evaluation at ADB started its journey to independence. Many challenges were encountered and several notable changes have taken place as independent evaluation sought to perform its strategic role in helping achieve development effectiveness. As evaluation moves forward, it must rise to the challenge of staying relevant and effective in an age of rapid change, complexity, and uncertainty. In a dynamic environment where the way forward is often not certain, tailoring evaluation methods to the demands of innovation will be increasingly crucial in finding solutions to emergent development challenges.

## Notes

<sup>1</sup> One meeting took place with the President on the evaluation work plan, but he felt that further meetings with the director general of independent evaluation might be perceived as compromising evaluation independence.

<sup>2</sup> Exceptions to the disclosure policy are the annual evaluation reviews, which are disclosed after the DEC meeting, and private sector evaluations, which are disclosed after sensitive business information is redacted.

<sup>3</sup> See para. 5 of the 2008 IED policy. Asian Development Bank. Review of the Independence and Effectiveness of the Operations Evaluation Department. November 2008. (R297-08).



<sup>4</sup> Independent evaluation as defined in OECD (2002).

<sup>5</sup> [www.ecgnet.org/gpsbigbook](http://www.ecgnet.org/gpsbigbook)

<sup>6</sup> The training program comprises two parts: (i) a core course in development evaluation for middle-level government officials involved in monitoring and evaluation and (ii) workshops on impact evaluation and performance-based budgeting and expenditure for senior officials with technical skills. Since 2009, 139 government officials from Central Asia Regional Economic Cooperation countries (Kazakhstan, Kyrgyz Republic, Tajikistan, and Uzbekistan) have participated in the program, and 82 officials from the Greater Mekong Subregion (Cambodia, Lao People's Democratic Republic, and Viet Nam).

<sup>7</sup> Other donors included the African Development Bank, Department for International Development of the United Kingdom, Inter-American Development Bank, International Fund for Agricultural Development, Swedish International Development Cooperation Agency, and the World Bank Group. Four regional centers in Africa, East and Southeast Asia, Latin America, and South Asia participated in the training program.

<sup>8</sup> From 2011 to 2013, 10 government officials participated in the on-the-job training (4–5 months).

<sup>9</sup> IED has trained over 300 ADB staff in preparing good-quality project completion reports since 2009. Over 200 staff have participated in project design and management program evaluation since 2011.



## Chapter 6

# Accountability and Learning: Two Sides of the Same Coin

**H**ow to achieve the twin goals of accountability and learning in improving development effectiveness is one of the most discussed issues in the evaluation community. Pursuing either one presents opportunities that evaluators can seize, but they also create tensions. On this issue, senior evaluators affiliated with the Evaluation Cooperation Group share their views based on three guiding questions: (i) how do they see the opportunity or tension between accountability and learning, (ii) what is their experience in running in tandem these dual functions, and (iii) what is the best way forward?

### *Cheryl Gray*

When I first went to work in the independent evaluation office of a multilateral development bank (in 2007, at the World Bank's Independent Evaluation Group [IEG]), a prominent operational vice president told me: "The problem with you folks is that you focus too much on accountability and not enough on learning." What he meant was that we devoted too many of our resources to showing what the World Bank was doing wrong—often leading to sensitive or defensive reactions from management—rather than helping management learn how it could do better.

I did not agree with that view but thought it was intriguing nonetheless—and have thought about it a lot during my 6 years since then working in evaluation (about half of those at the World Bank and half at the Inter-American Development Bank). How should independent evaluation groups balance the twin roles of ensuring accountability and promoting learning? Is there tension between them and, if so, how can it be managed constructively? This issue gained in prominence as I became more familiar with evaluation units in other international organizations, whether the United Nations (I have been a member of the external oversight advisory body of a UN agency for 4 years), bilateral aid agencies, or other financial institutions.

In my experience, there is indeed some tension between these roles, and the most appropriate balance between them depends very much on context. Promoting accountability requires “speaking truth to power,” even if it is not what the power structure wants to hear. But promoting learning may call for a more nuanced tone to ensure that management and staff remain open to hearing the messages. It is laudable if different evaluations can promote different goals. In practice, however, one of these roles is likely to predominate. So evaluation groups need to carefully consider what is feasible and how they will be able to add the most value in their particular environments. Three questions are worth asking in this regard.

First, what is the nature of governance in the agency? Does it have strong and effective internal governance mechanisms that will be able to use independent evaluations to promote accountability from management? In addition (or as an alternative), are there external watchdogs that keep an eye on the agency and will it be able to use evaluation findings to apply external pressure for performance? Large international institutions vary widely in their governance arrangements. Some may have more external scrutiny and more internal pressures for accountability, possibly because of the composition and activism of their standing boards. Others have less external scrutiny and may have only part-time boards, boards facing confused incentives (for example, confusion between their oversight and advocacy roles), or no boards at all. Organizations without strong internal or external governance may have less use for evaluations that are geared primarily toward accountability. In such environments, it may make sense to tilt evaluation resources toward learning in the hope that internal stakeholders are motivated toward performance.

Second, do other parts of the organization serve accountability or learning functions? In an ideal situation, the location of an independent evaluation unit in the organization chart is close enough to management to know what is going on but far enough away to allow independent judgment and action. In practice, the location of evaluation groups in a bureaucracy may be influenced by the presence of other internal or external groups whose roles overlap with theirs. In the United Kingdom, for example, the Department for International Development’s internal evaluations unit used to be more independent and saw itself as having a significant accountability role. But when the government set up a wholly external watchdog over aid, the

department's evaluation unit grew closer to management and focused more on its learning role. In other cases, agencies have departments such as research and monitoring and evaluation that focus heavily on the learning agenda, making it less critical for independent evaluation groups to concentrate on that role.

Third, is the institutional culture generally tolerant of constructive criticism? In reality, few organizations—of any type—have managements that appreciate and applaud critical findings or tough recommendations, even if they make sense from a technical or practical perspective. But here again there are differences among organizations, and those that are more resistant to criticism will have a harder time learning from honest and independent evaluation.

In sum, there is inherent tension between the accountability and learning roles of independent evaluation groups. And although it is admirable to try to serve both roles, in practice an evaluation group is likely to lean in one direction or the other. Understanding the environment—the governance situation, the roles of other actors, and the culture of the institution and its tolerance for open debate—can help evaluators determine which way to lean to increase the chances for effectiveness and impact.

### ***Caroline Heider***

“People, you got to fight for independence—it won’t just come for free.” This was not the call of a freedom fighter, but the Asian Development Bank’s (ADB) former director general Eisuke Suzuki, who led the effort to take the institution’s Operations Evaluation Department from reporting directly to the president to reporting instead to the Development Effectiveness Committee (DEC). It was a transition that marked the start of independent evaluation at ADB. I had joined the evaluation office in 1995 and was back there after a short stint in the Strategy and Policy Department, well in time for independence.

So how did it compare, the before and after independence? A question one of the DEC chairs raised when I presented the country assistance program evaluation for Papua New Guinea was: were you more critical when you knew the evaluation office was becoming

independent and tougher on accountability than you would have been? Luckily, the DEC had discussed my earlier evaluation on Mongolia and praised it for being frank and candid. And, indeed, I have been fortunate to work with people and in contexts that valued intellectual independence, and know how important this is as the bedrock for independence.

Having worked in many organizations, I can explain why independence is so important for accountability and learning. Whenever an institution has a corporate culture that rejects the idea that something might go wrong, it reacts adversely to the messengers of bad news. But, more often than not, the negative message is associated with accountability. Such an attitude never fails to surprise me given that development work is risky and does not simply happen (if it did, development institutions like ours would have long gone out of business).

At the World Bank, IEG was the forerunner of independence among multilateral development banks—we at ADB used it as the model when designing ours in 2004—and it instituted measures that are still regarded as gold standards in the evaluation community. These include direct reporting to the executive board, which also appoints the group's director general and approves the work program and budget, and guidelines to prevent conflicts of interest and repercussions on the careers of staff issuing candid evaluations. All of these measures safeguard frank and independent evaluations that hold the World Bank accountable. In short, structural and intellectual independence are necessary conditions for accountability.

As an evaluator for over 20 years, I have never struggled to reconcile accountability and learning. For me, these are two sides of the same coin. Accounting for what happened during implementation when working toward achieving the plans set out in a project design—an agreement and commitment between at least two parties—is the source for learning about the replication of success and avoidance of unnecessary mistakes on the journey to realize expected outcomes. In evaluation, we do not aim to learn in the abstract or test theoretical concepts, but learn about what happened during the process of implementing a policy, strategy, program, or project.

It was when I joined the World Food Programme that accountability and learning appeared to be irreconcilable. In essence the dilemma was: “if you want us to be accountable, we cannot learn” and “if you

want us to learn, don't hold us accountable." Never having seen this argument put so starkly, I sought explanations in the humanitarian sector. Here, both the mission and dedication to save lives seemed to generate a mix that made people sensitive to news about failure, which they took either personally or as an extreme risk to raising resources.

Nonetheless, during my tenure as director of evaluation, we experienced a fundamental shift in attitude when colleagues realized accountability was not about blame. It was about openly and constructively discussing the bottlenecks we faced in doing an incredibly hard and risky job. If we evaluators had been less independent, we could not have raised accountability issues around systems' failures that were well known but not openly discussed or resolved. And we achieved this by setting up a system to increase objectivity and reduce room for opinion-based expert assessments, made possible by the great degree of intellectual independence of my team.

Arriving in IEG, I thought independence, accountability, and learning would not be a hot topic, simply because it had long been independent. But here, too, it was a very lively issue, particularly as we strove to rebalance our focus on the accountability and learning objectives of evaluation. So, can evaluators overcome the hurdle that independence seems to create and reach out for greater learning? My many years of practice tell me "yes." And IEG's recent strides in this area—by expanding learning products, holding joint events with World Bank Group operational teams to share evaluation findings, and piloting a range of media for sharing lessons, to name just a few—are proof that we can without compromising independence or our responsibility for accountability.

### ***Rachel Meghir, Horst Wattenbach, and Luigi Cuna***

Accountability and learning are often presented as the twin goals of independent evaluation, between which a certain balance needs to be struck. Yet there are grounds for considering whether a more fitting conceptual representation would be two goalposts at either end of a spectrum of possible functions of the evaluation unit of a multilateral development bank. Such a unit could potentially

fulfill learning or accountability functions, or a combination of the two. The “mix” would largely reflect the particular characteristics of the organization, its governance structure, operating model, institutional culture, and the degree of engagement of the board, as this largely determines the interest in or demand for accountability and learning.<sup>1</sup>

Accountability requires setting explicit corporate objectives. It also requires an institutional culture that promotes ownership and responsibility to achieve these objectives by advocating transparency within the organization and toward the public in the reporting of results. Accountability can therefore be upward-oriented (to a governing body) and downward/outward-oriented (to external stakeholders and the public).

For multilateral development banks, learning is generally a horizontal concept; indeed, it is hard to imagine a top-down approach in the sense of a headmaster teaching students what to do and how. The learning culture of multilateral development banks requires a combination of formal and informal spaces of interaction, prompt responses to learning needs, and a strong collaborative approach, both inside and outside the institution. It is also important to clarify the type of learning; that is, whether it is for practical improvements to the quality of operations and the processes through which they are designed and implemented (down-to-earth, collaborative learning), or more high-level strategic adjustments (institutional learning).

Based on the above definitions, the twin objectives of learning and accountability are admittedly difficult to fully achieve in parallel. To use the goalpost analogy again, the evaluation function’s position within the institutional space it is granted can evolve as the institution and its evaluation function mature, either to serve the demand for accountability or strengthen learning. In any event, this process must be placed in a dynamic perspective.

The tension or opportunity between accountability and learning can thus be depicted by two scenarios. The first is strong demand for accountability, but with the aim of generating some learning. The second is strong demand for learning, but with the aspiration to serve accountability.



When accountability prevails, the possibility of learning objectives must not be excluded, although the learning spillover in this case is more likely to be of the institutional type, in which accountability is used as a means of reflecting on an organization's strategies and performance. Admittedly, this type of learning may not be of immediate use to those involved in day-to-day development operations, as it has a different audience and knowledge content.

Still, instances of collaborative, down-to-earth learning can also be generated at the microlevel; for example, during the evaluation of specific operations. This can apply when operations being evaluated are of strategic interest to the organization or there is strong interest from evaluatees in the findings, and these are generated in a timely manner.

At the other end of the spectrum, when an evaluation function is mostly geared to learning, the possibility of some form of accountability must likewise not be excluded, even though it is a more challenging case to make. Here, it is the definition of accountability which needs to be nuanced: not the corporate-level accountability noted earlier, but more a microlevel accountability concern within operations that have been selected for in-depth evaluation for learning purposes. In this context, accountability as a spillover effect of learning begs the question: accountable to whom? The operating model of the organization needs to be characterized by a culture of institutional ownership and responsibility for the results achieved—or not achieved—by the foregoing operations, underpinned by adequate instruments of communication and information sharing.

Because of its history, financial structure, and governance model, the accountability framework of the Council of Europe Development Bank—both upward to its governing bodies and downward-outward to external stakeholders and the public—differs from those of larger multilateral development banks. Accountability issues are an essential concern of the Council of Europe Development Bank's governing bodies, as evidenced by the Administrative Council's support for the governor's commitment to strengthen the social orientation and impact of financed operations. In this context, the Evaluation Department noted the positive reception by the Administrative Council of higher-level deliverables by the department (such as evaluation cycle synthesis reports).

However, accountability in the sense of reporting on the overall performance of the institution has not been the focus of the Evaluation Department, since the unit is too small and the necessary complementary systems that would generate credible and reliable information for such reporting are not yet in place. Although the Council of Europe Development Bank finances a wide range of activities, performance information at completion, at this stage, is not sufficiently accessible to enable reporting on higher-level outcomes. Accountability concerns identified within individual operations selected for in-depth evaluation for learning purposes, of course, are addressed.

In practice, the focus at the Council of Europe Development Bank has been on learning through the evaluation of individual projects and programs to improve the quality of operations and the processes through which they are designed and implemented. Awareness is growing of the contribution that evaluation of individual operations can make in this context, and, within the constraints of available resources, the Evaluation Department is increasingly asked to contribute findings upstream to program design platforms and facilitate the use of lessons learned.

Ideally, an institution best pursues the two objectives of learning and accountability in tandem, since it needs both to adapt to a changing world and credibly demonstrate to stakeholders and the public that it is successfully pursuing and fulfilling its mission and mandate. In reality, though, it may not pursue them with the same emphasis because of resource constraints.

Accountability in overall organizational performance is not easily established on a small scale and requires a critical mass of information systems and resources to be credible. In an organizational setting in which adequate instruments for learning are available (such as a research and study department and a credible system of self-evaluation), the natural way forward for an evaluation unit would be to progressively strengthen its accountability orientation and arrange its work program and activities accordingly, thereby gradually shifting the evaluation focus to higher-level corporate and strategic learning.

At the same time, the increasing interest of the public in knowing how well publicly owned institutions perform their missions can create momentum for organizations that have a different level of mix, with greater emphasis on the learning side, to move in the same direction. In this regard, preparing for that moment would certainly be desirable by progressively establishing the necessary preconditions.

### **Ashwani Muthoo**

The traditional view of the accountability and learning functions or goals of evaluation is that there is a trade-off between them: that the stronger the accountability function, the weaker the learning role, and vice versa. However, the experience of the Independent Office of Evaluation (IOE) at the International Fund for Agricultural Development (IFAD) shows that these goals can be complementary. Thus, accountability is a key component of the incentive framework for transforming knowledge generated by independent evaluation into learning.

Put schematically,

**Independent evaluation ⇔ knowledge ⇔ accountability ⇔ learning**

Without accountability, knowledge may not be transformed into learning, with the result that programs and projects and institutions may remain unchanged.

At IFAD, access to knowledge is facilitated through a set of knowledge products, whereas learning is promoted by processes involving operations and key stakeholders. Some of these processes are tools for holding management accountable for applying the knowledge generated by evaluations, and so enhancing development effectiveness.

For each evaluation, the IOE forms a Core Learning Partnership. This comprises representatives of the main users of the evaluation results drawn from government, beneficiary groups, project staff, nongovernment organizations, IFAD management, and the IOE. The Core Learning Partnership provides a platform for stakeholders for

dialogue and reflection during the evaluation process and an ongoing opportunity to deepen understanding of the evaluation findings and subsequent conclusions. The Core Learning Partnership agrees on the recommendations, division of labor, and responsibilities for their implementation. These are recorded in the agreement at completion point, one of the key learning processes, and the end point of a process that aims to determine how well evaluation users understand the recommendations proposed in the independent evaluation, and how they will attempt to make them operational. Interaction among the stakeholders working through the Core Learning Partnership helps deepen understanding of evaluation findings and recommendations, and elicits ownership for implementing the recommendations. The IOE facilitates the agreement at completion point to ensure a full understanding of the evaluation findings and recommendations.

The agreement makes explicit reference to the partners with whom it was concluded. These include all major users of evaluation results, such as IFAD's operational units, project and borrower country authorities, and other stakeholders. All agreements are posted on IFAD's website.

In addition to evaluation reports, IFAD's knowledge products include two types of briefing notes on evaluations, called *Evaluation Profiles* and *Evaluation Insights*. The first are 2-page summaries of the main conclusions and recommendations of the evaluations. These are mainly intended for IFAD's management, government ministers, and decision makers that do not have the time to read every full report. *Evaluation Profiles* give a sampling of evaluation results which, hopefully, will be an incentive for readers to go to the full report. *Evaluation Profiles* also provide early warning of issues identified in an evaluation that require immediate attention.

*Evaluation Insights* focus on learning lessons from corporate, thematic, or country program evaluations. Presenting a hypothesis, they form the basis for debate and discussion among development professionals and policy makers within IFAD and outside the institution. The IOE writes them in collaboration with other members of the Core Learning Partnership. Target *Evaluation Insights* are aimed at a wide audience, including IFAD staff, executive board members, project staff, government officials, cooperating institutions, nongovernment

organizations, and other partners. In addition, the IOE produces *Evaluation Synthesis Reports* that distill lessons learned and good practices.

During executive board sessions when corporate policy or strategy papers are presented, the IOE submits to board representatives “notes with comments on policy or strategy papers” called *IOE Notes*. These are based on previous corporate level evaluations for a particular policy or strategy. *IOE Notes* indicate whether a proposed policy or strategy provides an adequate response to a corresponding evaluation on the same policy area. They present overall alignment of the findings and recommendations with the respective corporate-level evaluation.

In some cases, *IOE Notes* indicate whether a new strategy adequately took account of the recommendations of corporate level evaluations or whether these did not receive sufficient follow-up. Similar types of *IOE Notes* are produced when a country opportunities strategic program is submitted to an executive board session and the IOE has prepared a country program evaluation for the same country. These notes are a means to hold the Project Management Department (Operations) accountable for learning.

Thus, IFAD’s independent evaluation experience shows that the learning and accountability functions of evaluation can be complementary rather than a trade-off. This is important because management expects evaluation offices to contribute to learning processes, and there is a risk of isolation if insufficient attention is paid to this aspect. As such, evaluation offices have the responsibility to take part in internal learning platforms, ensuring this does not compromise the independence of its function.

Independent evaluation will not fulfill its ultimate goals if a balance between accountability and learning is not achieved. This is important to promote institutional transformation, contribute to better results, and generate knowledge the wider development community can use to enhance development effectiveness.

**Rakesh Nangia and Samer Hachem**

The African Development Bank has embarked on its new strategy—*At the Center of Africa's Transformation for the Next 10 Years*. A word the strategy document mentions repeatedly deserves attention: “transformation.” As the African continent moves faster along its development path, the African Development Bank intends to “transform to support the transformation” of Africa and has made this a key tenet of the strategy. While the institution has made the objective clear and explicit, it acknowledges that organizations transform over time, either organically or due to external constraints or pressures. Many organizational theories say much about such transformations, and given the multiple facets and visions of organizations, it can be said that each facet describes a little bit of the reality of any institution. But whatever theory is applied, the ability to transform ultimately hinges on questioning practices for learning and evolving.

It is in the context of transformation that I would like to look at the role of independent evaluation. The independent nature of evaluation functions has traditionally influenced a focus on quality and accountability. Looking at an operation or program using a set of standards enables evaluators to determine what works well and what doesn't, and to propose improvements. That said, we all know there are plenty of excellent reports around that could improve the way an organization works, but that have not been acted upon.

Independent evaluation functions that are operating with standards and are fully immersed in the context of their organizations are what it takes to move from potential to concrete impacts. Every organization has its own norms, standards, and practices. The way structures are defined, processes controlled, and incentives distributed takes root in these underlying norms and standards. Indeed, transformation is not only about changing people, structures, or processes, but also about reviewing underlying norms. By understanding this, while applying rigorous standards, evaluation has the potential to reveal the areas where real issues lie and where tangible progress can be achieved. It is the combination of a rigorous examination of facts with an open-minded elaboration of conclusions that takes into account context that allows usability (or learning, to put it another way).

Accountability and learning, although sometimes seen to be in tension, are fully complementary. It is by continuously learning to do better that accountability takes its full meaning for the simple reason that no human, and therefore no organization structured around human relationships, is foolproof.

What does this mean in real terms for evaluators? In strategic terms, the design of the work program of independent evaluation must be guided by both the strategic focus of the organization and the needs of its governing bodies in terms of management information. In operational terms, evaluation teams must be better equipped to understand the underlying norms and standards in which they are operating. This could be difficult to ask from consultants, who may well be correctly performing a role as “guardians of facts.” From this perspective, internal evaluators play a key role, both through their own understanding, but also by establishing a network of positive relationships within their organizations. The ability of evaluators to fully collaborate with operational staff and managers in designing, understanding, and learning from evaluation is critical to make sure objective facts can then be transformed into practical impacts.

### ***Moises Schwartz***

Independent evaluation in international organizations has attempted to balance the apparently competing objectives of accountability and learning. It is now well recognized that independent evaluation can enhance organizational accountability in public entities where assessing performance is complex, while also serving as a learning tool. Indeed, although there may be tension between these two goals, there are also complementarities—accountability provides incentives for learning and the lessons from evaluation offer a framework for accountability.

I would like to share my views on how the Independent Evaluation Office of the International Monetary Fund (IMF) aims to strike the right balance between these objectives. My experience at the IMF suggests that independent evaluation can and must serve both accountability and learning functions—and the Independent Evaluation Office’s terms of reference clearly call for it to do so.

The office's mandate is to enhance the learning culture within the IMF, strengthen the IMF's external credibility, promote greater understanding of its work throughout the membership, and support the executive board's institutional governance and oversight responsibilities.

The 20 evaluation reports produced since the Independent Evaluation Office's inception in 2001 contain both learning and accountability components. Overall, I am inclined to suggest that these reports have favored learning over accountability. Furthermore, the accountability component in evaluation reports has always taken the form of "institutional accountability" as opposed to "individual accountability." In other words, the evaluative emphasis in these reports is on IMF performance rather than the individuals responsible for carrying out the work.

It seems to me that this has blurred the distinction between learning and accountability in the Independent Evaluation Office's outputs, and the two components have therefore become more difficult to disentangle. As a result, the seeming tension between the two objectives has tended to dissipate. To be precise, when evaluation reports have pointed to instances in which the IMF has fallen short in its performance (the accountability element), the exercise turns into a quest to identify the reason for such behavior, and the findings and conclusions then contribute toward an enhanced organization (the learning element). Focusing on institutional accountability makes the distinction between accountability and learning hardly noticeable and their supposed conflict nonexistent.

Alternatively, if independent evaluation were to stress individual accountability as opposed to institutional accountability, it would then be serving more the purpose of supporting management's ability to administer the institution than a method of institutional learning. By underscoring institutional accountability, evaluation reports have served the double objective of learning from the past, while making the IMF, as a whole, accountable for its performance. I believe that this approach is useful and conducive to the attainment of the two goals of independent evaluation. I see the Independent Evaluation Office performing both roles effectively in a complementary manner and consider that evaluation products must maintain these two objectives.



The impact of independent evaluation on institutional learning depends on factors that likely lie outside the command of independent evaluation. That is, institutional and organizational features beyond the control of the independent evaluation unit have a clear-cut effect on the receptiveness of institutions to the learning that the findings of independent evaluation afford. At the IMF, evaluation reports present a set of findings and recommendations, but the Independent Evaluation Office is not responsible for bringing about required changes. The office plants the seeds for advancing institutional learning, yet the actual responsibility for the IMF to learn from evaluation findings rests on other participants, such as management, staff, and the executive board. In essence, learning by the IMF through evaluation reports takes place through staff and management's assimilation of the results, as well as from the Independent Evaluation Office's interactions with the executive board and other stakeholders.

Evaluation findings also provide a framework for the membership and the executive board to hold management and staff accountable. As with learning, actual utilization of evaluation findings does not depend only on the Independent Evaluation Office's efforts, but on the partaking of the different stakeholders. Making evaluations public help civil society organizations in member countries understand, hold accountable, and reform the IMF.

Notwithstanding the particular institutional layout that may or may not aid the assimilation of lessons and accountability within the IMF, independent evaluation plays a crucial role in fostering the conditions for greater learning and accountability in the institution. The process by which evaluations contribute to the IMF's learning and accountability is thus multifaceted, intricate, and sometimes protracted. This is because evaluation findings need to be absorbed by different participants and then processed through interactions among them. As a result, while some evaluation findings might have an immediate impact on the institution, others may take time to be analyzed and incorporated into the functioning of the IMF.

The policies and processes to utilize independent evaluation at the IMF are evolving. A decade after the Independent Evaluation Office's creation, there is still a need to strengthen this framework to ensure that it can contribute to the organization's learning and accountability

to its fullest potential. In any event, there is ample evidence that independent evaluation has contributed to learning at the IMF, and that the institution is more accountable as a result of the Independent Evaluation Office's presence.

**Jan Willem van der Kaaij, Ivory Yong-Prötzel, and  
Bastiaan de Laat**

Theoreticians have often described the accountability and learning objectives of evaluation as two different, sometimes even irreconcilable, paths that evaluation can pursue. However, practitioners have long faced the need to combine both. Most evaluation functions these days have a double mission of providing information on how successfully policies, programs, or projects have been implemented, and on what works and why and what can be improved in the future. The Roman god Janus has even been used as a symbol for evaluation as he is usually represented with two profiles, simultaneously looking to the future and to the past. Fortunately for practitioners, different scholars increasingly argue that these two aims go hand in hand (Guijt 2010). In other words, there cannot be accountability without learning, and learning needs accountability.

At the European Investment Bank (EIB), the mandate provided by the board of directors to Operations Evaluation states that the evaluation of public and private sector operations as well as related policies and strategies is aimed at “identifying aspects which could improve operational performance, accountability, and transparency.” Interestingly, the term “learning” does not appear in this definition. This reflects, next to accountability, the emphasis put on change and improvement rather than on learning in itself. If there is tension, then it is between accountability and institutional improvement rather than between accountability and learning. Otherwise, learning would be assumed to *always* induce change, which it does not. By carefully structuring evaluations, this issue can be overcome.

In our view, the body of knowledge produced by an evaluation can be used for different purposes: to learn and introduce change or to inform and contribute to the process of accountability. The approach used for a given evaluation and the knowledge produced

may be different according to the aim of the evaluation, as may be the users of this knowledge. Through evaluation at the EIB, our board of directors and the wider public are better informed about the extent to which the EIB's overall mission has been achieved, in particular through its lending activities. Thus, evaluation contributes to increased accountability. And through evaluation, the EIB's management and its operational services learn what works and why, allowing improvements to future activities. In this way, evaluation can contribute to institutional change and improvement.

Historically, accountability was probably the most important component of evaluation at the EIB. Evaluations focused on how well projects were implemented and were operated with respect to project plans. This could even be called operational accountability, as little emphasis was given to the achievement of higher-level objectives or to outcomes and impacts. In recent years, however, the emphasis has shifted toward higher-level, more strategic evaluations that develop the learning aspects more. The mandate to achieve both the accountability and improvement objectives remains, but the balance between the two has evolved. This has introduced tensions, because even if they are expected to reinforce each other, limited resources often imply that working more toward one objective is done at the expense of the other.

As an example, the EIB's board approves operations evaluation studies using a wider array of information collection methods when carrying out thematic evaluations. This is a departure from the old approach in which thematic evaluation reports simply summarized the information collected through project evaluations. The use of wider information collection methods, along with a more analytical approach, has resulted in more forward-looking and learning-gearred reports but involved a reduced sample of projects analyzed. For accountability reasons, however, the board does not welcome a decrease in the number of projects evaluated per year. Another example is the board's strong preference for random samples, which are fundamental for accountability but not for learning purposes. With fixed resources, a decision has to be made for each thematic evaluation to favor either learning or accountability.

Tensions between accountability and learning for institutional change have also been intensified by the faster pace of development change,

particularly in the context of crises. Accountability has traditionally been associated with looking at how previous activities have worked. The knowledge stemming from those evaluations was useful to draw lessons for the future, but that was as long as the social and institutional context was stable. If institutional objectives and the type of activities to achieve them change significantly, the relevance of lessons from past activities is reduced. One of our current challenges is how to draw useful lessons for the future using knowledge stemming from evaluations geared toward accountability.

Viewed from this perspective, the question becomes not so much what trade-offs exist between accountability and improvement, but how the evaluation process can be optimized given these objectives. In other words, how can an evaluation be structured to deal with both objectives? Finding the answer can be guided by the main audiences for the different types of knowledge resulting from evaluation. Of course, the interest of a certain audience is not necessarily limited to either one of the two objectives. For instance, the EIB's board might also want to be assured that operational improvements stemming from evaluation activities are being implemented.

The increased use of evaluation for change and improvement has meant that at Operations Evaluation, while safeguarding our independence, we increasingly engage with operational services at different stages of the evaluation process. For instance, aside from the formal consultation of finalized reports, we have introduced a more interactive process to discuss approach papers as well as our emerging findings and recommendations. The latter is done after the data collection process has been completed but before the report is drafted (during the analytical stage of the evaluation process).

Operations Evaluation is trying to make the body of knowledge produced by its evaluations more accessible to EIB services by preparing new activities and departing from the current approach of simply making evaluation reports available on the web and hoping that the relevant staff will use them as needed—an approach that has not proved very effective. Indeed, knowledge can induce improvement, but only when the right institutional setting is in place and when there is a corporate culture of change. As part of this effort, Operations Evaluation is in the process of designing an improved follow-up mechanism to track the implementation of agreed recommendations from evaluation reports.

A concrete example of the interaction between the accountability and improvement aims of evaluation at the EIB is the workshops organized just ahead of board meetings at which thematic evaluations are discussed in more detail than during the plenary board meetings. The workshops bring together a wider audience—members of the EIB’s management and operational services, board members, and advisors—and allow staff to discuss their work at a high-level forum.

The intensified emphasis on the multifaceted aspects of the learning process related to evaluations has brought about a beneficial development, the importance of which, in our view, can hardly be overestimated. Instead of being seen by management and operational services as a kind of external “rating agency” giving judgments on the quality of the activities of the EIB, Operations Evaluation is now much more recognized as a player that contributes to the improvement of the institution’s activities. And this has increased cooperation between the EIB’s operational services and Operations Evaluation, which, in turn, has helped improve evaluations.

### **Note**

<sup>1</sup> The magnitude and scope of this demand/interest is not directly controllable by the evaluation unit. It can perhaps be influenced but not determined.





**Part 3**

## **Evaluation Approaches**

Part 3 examines evaluations from several perspectives. These include methodological approaches to corporate evaluations in the World Bank, influential evaluations at ADB, and developing a gender-responsive monitoring and evaluation system for international development agencies.

In Chapter 7, Anis Dani, lead evaluator of country and corporate evaluations at the Independent Evaluation Group, uses evaluations of the World Bank Group's safeguards policies, matrix system, country-level engagement on governance and anticorruption, and assistance to fragile and conflict-affected states to synthesize good evaluation practices and to illustrate essential elements of evaluation design. These elements include corporate objectives, theory of change, key questions, evaluation methodology, and evaluation matrix.

To achieve influence, meanwhile, Hemamala Hettige, senior advisor of Independent Evaluation at ADB, argues in Chapter 8 that an evaluation must connect with current and emerging development needs, engage actively with stakeholders, and communicate widely. She contends further that evaluators have many roles—as educator, advisor, facilitator, change agent, accountability advocate, and innovator—which shape development thinking and help bring about necessary changes. Influential evaluations at ADB included a 2012 evaluation that sped up the organization's efforts to become a stronger knowledge institution, as well as the evaluation of ADB's response to natural disasters and disaster risks, which led to a policy decision to establish a disaster response facility and informed ADB's operational plan for integrated disaster risk management.

The section wraps up with a chapter from Michael Bamberger, development and gender evaluation consultant and former World Bank senior sociologist, who argues for the integration of gender into development evaluation, given compelling evidence on persisting gender equalities. He offers practical steps in designing and implementing a gender-responsive monitoring and evaluation system.



## Chapter 7

# Corporate Evaluations in the World Bank Group: Approaches to Enhance Rigor and Evaluation Impact

*Anis A. Dani*

Corporate evaluations in multilateral development banks (MDBs) are undertaken to assess the achievement of corporate policies<sup>1</sup> and strategic priorities.<sup>2</sup> These objectives tend to be long-term, cross-cutting, and relevant to the entire or a significant part of an organization. Successfully achieving corporate objectives depends on actions by multiple organizational units at different levels requiring coordination, commitment, and ownership across an organization. Corporate objectives tend to be relevant to the clients of MDBs and their effectiveness in delivering products and services to clients has immediate or downstream effects on these institutions. At the same time, the effectiveness and efficiency of achieving these corporate objectives tends to depend on ownership and actions by client governments and other stakeholders. So the achievement of corporate objectives is intricately linked to development effectiveness.

This chapter synthesizes methodological lessons from evaluations of corporate policies and strategies at the World Bank Group assessed by the Independent Evaluation Group (IEG) in the last 5 years. It draws in particular on insights from IEG's evaluations on the World Bank Group's safeguards policies (IEG 2010), the effectiveness of its matrix system (IEG 2012), country-level engagement on governance and anticorruption (IEG 2011), and assistance to fragile and conflict-affected states (IEG 2014). The chapter draws attention to five essential elements of evaluation design, encapsulated in the approach paper for each evaluation that lays out its scope and methodology. This is followed by five distinct sets of actions to undertake evaluations. In addition to the evaluation content, the evaluation process plays an important role in achieving a successful evaluation, as this chapter aims to show.

Within the World Bank Group, corporate evaluations face two major constraints. Corporate policies and strategies often lack clearly formulated objectives supported by a results framework and performance indicators, and they often lack counterfactual and baseline data to measure impact. For example, while there is broad understanding of the objectives of the World Bank's safeguard and procurement policies, at the corporate level they do not have a results framework or indicators to monitor their performance. To the extent that there is monitoring of these policies, it is done at the level of individual lending operations. Strategic priorities for the matrix system and fragile and conflict-affected states both lack a clear formulation of objectives and a results framework. And corporate policies, priorities, and processes for all these evaluation areas often lack baseline data against which progress can be measured. Furthermore, they do not lend themselves to establishment of a counterfactual.

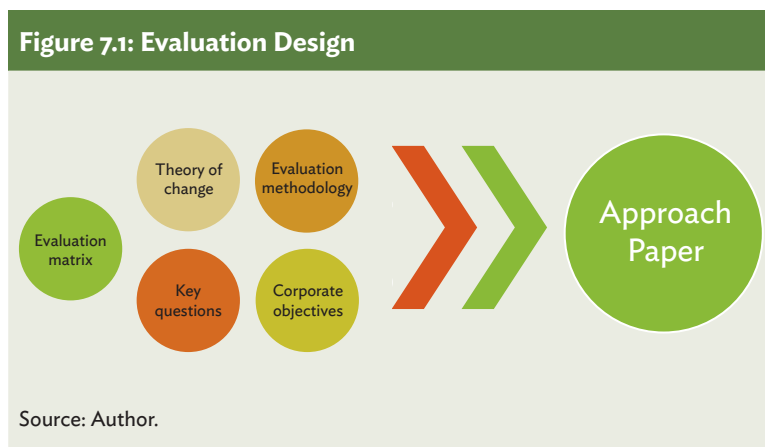
While these constraints are not unique, they are more prevalent at the corporate level than at project, country, and sectoral levels, at which considerable effort over the past decade has gone into improving the formulation of strategic or project development objectives and the development of results frameworks at the World Bank. As long-term, cross-cutting, and relevant to all sectors and country programs—or a significant part of an MDB's programs—corporate policies and strategies have a much longer life cycle and greater inertia, and are consequently much more difficult to change. Their organizational footprint makes them powerful catalysts for action but also, in some cases, inhibitors of innovation and risk taking.

The introduction of environmental and social safeguard policies in the 1980s positioned the World Bank as the global leader in preventing or mitigating adverse environmental and social impacts in MDB-financed projects. However, some clients and staff increasingly perceive excessive attention to fiduciary and safeguard requirements as raising the cost of doing business with the World Bank, sometimes crowding out other aspects of technical quality. The 2007 strategy for governance and anticorruption gave new impetus to innovative work at the World Bank and legitimized and mandated more systematic attention to these issues in country assistance strategies and lending operations. Yet it also gave rise to concerns over whether the enhanced attention to corruption was crowding out attention to other aspects of governance and institution building.

To balance these competing interests, corporate evaluations assess both the extent to which their intended objectives are being met and their effect on the ability of MDBs to engage with their clients and deliver the needed products and services. When evaluating corporate policies, these constraints should be discussed and an appropriate methodology designed and articulated in the approach paper.

### The approach paper

The approach paper is the primary vehicle through which evaluation design is developed to clarify the scope of work and the methodology of the evaluation. Careful attention to five essential elements in the approach paper can lead to robust evaluation design (Figure 7.1).



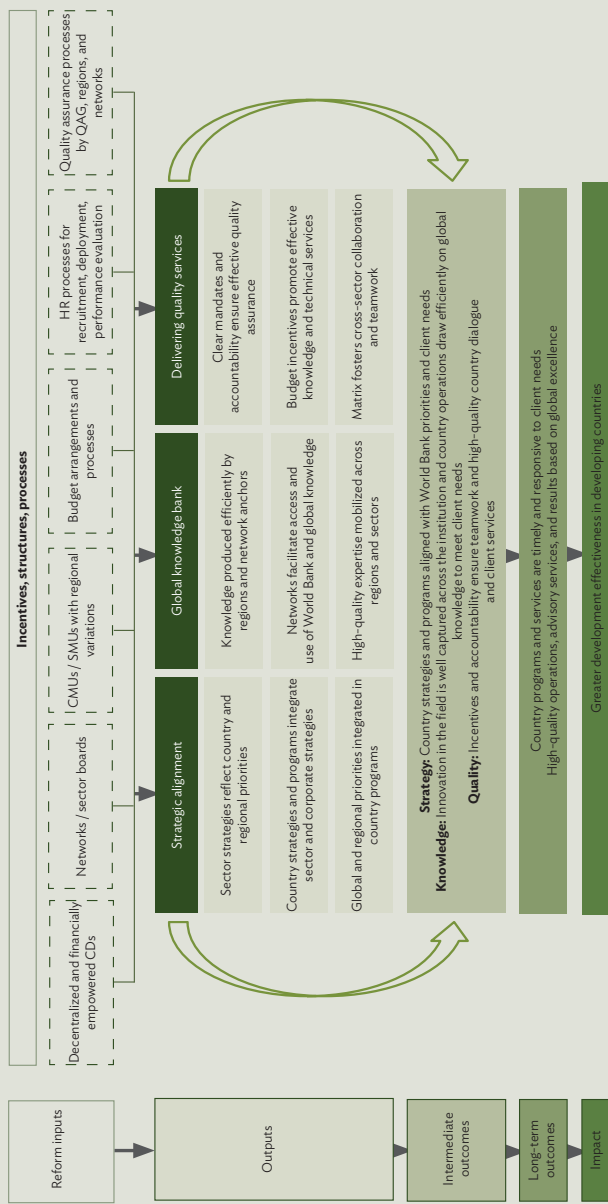
**1. Corporate objectives:** These need to be explicitly identified to conduct an objectives-based evaluation. In some instances, such as for the evaluations of safeguard or procurement policies, they are specified or may be easier to identify from the policies themselves. In other instances, as in matrix and fragile and conflict-affected states evaluations, they have to be extrapolated from background documents. In such cases, it is essential to consult with stakeholders to understand their perceptions of what the corporate objectives are and whether these have evolved over time. Understanding the significance and relevance of the corporate objectives presupposes a thorough understanding of the organizational context that led to those objectives. This context should be described clearly in the approach paper.

**2. Theory of change:** More often than not, corporate policies and priorities lack an explicit theory of change. Their results framework has not been formulated and indicators are not specified. For any evaluation to be credible, it is imperative that a theory of change be constructed (Imas and Rist 2009).

Typically this involves three steps: (i) the formulation of a results chain that maintains a clear line from inputs and outputs to intended outcomes and impacts, (ii) identifying comparators or a counterfactual where feasible, and (iii) articulating explicit or implicit assumptions and hypotheses. Although this can be derived from a review of background documents to reach a common understanding and ownership, it is best to develop these in consultation with stakeholders. Figure 7.2 presents an example of a results chain that also served as the evaluation framework for the matrix system evaluation. In the safeguards evaluation, where it was not possible to construct a counterfactual, the performance standards the International Finance Corporation adopted were used as a comparator for the environmental and social safeguard policies the World Bank was using. Similarly, for the evaluation of fragile and conflict-affected states, the performance of 33 such states was compared with that of 31 nonfragile states.

**3. Key evaluation questions:** The specification of the evaluation questions is vital to reach agreement with boards and management on the broad scope and specific focus of evaluations. The questions need to be sharpened iteratively during the process of stakeholder consultation, but they also need to be consistent with the theory of change to enable evaluative judgment. The questions asked in the evaluation of the matrix system made it clear that the scope of the evaluation would be limited to strategy, knowledge, and quality (IEG 2012). Two related aspects of the matrix system that were part of overall reform—decentralization and human resources—were deliberately excluded from the evaluation. Similarly, in an evaluation on fragile and conflict-affected states, fragile situations due to subnational conflict and violence in countries that were not as a whole classified as these types of states, were excluded. Sub-questions help to clarify the methodological approach by identifying what the evaluation team believes will be feasible within the time and resource constraints of the evaluation. These sub-questions listed in the approach paper for the evaluation on fragile and conflict-affected

Figure 7.2: Matrix management evaluative framework



CD = country director, CMU = country management unit, HR = human resources, QAG = Quality Assurance Group, SMU = sector management unit. Source: Independent Evaluation Group. 2012. *The Matrix System at Work: An Evaluation of the World Bank's Operational Effectiveness*. Washington, DC: World Bank.

states clarified the priority themes to be focused on: building capacity of the state, building human and social development, and promoting inclusive growth and jobs. The approach paper also conveyed the importance the evaluation gave to gender and partnerships.

**4. Evaluation methodology:** The evaluation methodology is developed by considering (i) findings that can be drawn from existing evaluations, (ii) findings that can be obtained by analyzing existing databases, and (iii) what new data need to be collected.

Typically, evaluation units like IEG and the operational or research departments of MDBs have evaluation material from micro (project level) evaluations and macro (sector or thematic level) evaluations that can be tapped for corporate evaluations. However, unless they were undertaken as deliberate inputs to the corporate evaluation, they may not address the evaluation questions directly. In the preceding years leading up to three of the four evaluations discussed in this chapter, IEG undertook several project performance assessment reviews to inform the safeguards evaluation and, prior to the evaluation on fragile and conflict-affected states, undertook several country program evaluations in postconflict countries. In MDBs generally, project-level results might also exist from assessments management has conducted. Compiling and synthesizing these findings is essential for efficiency and to compare the findings from self-evaluations with those from independent evaluations.

In addition to project, sector, and corporate level evaluation material in evaluation units or within the operational complex of MDBs, there may also be a repository of data on operations, budgets, and human resources. Relevant databases maintained by agencies such as the Organisation for Economic Co-operation and Development and specialized agencies of the United Nations are also useful for corporate evaluations of MDBs. Such databases should be identified in approach papers, and resources for analyzing them should be budgeted appropriately.

Despite access to evaluations and databases, new data collection is often necessary to fill data gaps and to validate or cross-check the validity of data or information reported by management. Consideration needs to be given to what new data are needed and how these will be collected within the time and resource constraints of an evaluation. Evaluations tend to rely heavily on quantitative data

to assess performance, but qualitative data are equally important for understanding the underlying drivers of performance. Here, mixed methods are gaining purchase (Mertens and Hesse-Biber 2013). New data collection often took the form of staff surveys and key informant interviews. But with the advent of new technologies, it is now more feasible to undertake beneficiary and client surveys within the time and resource constraints that are typically available for corporate evaluations.

**5. Evaluation matrix:** The evaluation matrix summarizes the logic and content of proposed evaluations. It typically lists the most essential questions along the vertical column, with the horizontal rows for each question summarizing the evaluation design listing. These include the information required, sources of information, data collection tools and techniques, the data analysis method to be used, and the strengths and limitations of the approach. An evaluation matrix is a valuable tool to develop cost and time estimates and for checking the feasibility of the evaluation and identify gaps in the approach. If sufficient information is available at the beginning of the evaluation process, the matrix can be used to organize the materials available to design the evaluation, as was the case for the evaluation on fragile and conflict-affected states. Alternatively, an evaluation matrix can be prepared when the approach paper is being completed to check the internal logic and realism of the evaluation design and methodological approach proposed, as was done for the matrix system evaluation. The approach paper should clearly state the limitations of the evaluation.

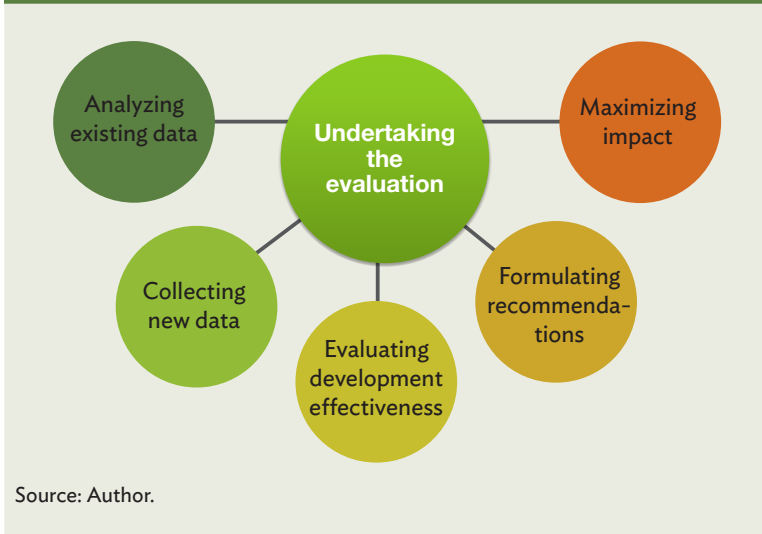
### **Undertaking corporate evaluations**

At a minimum, undertaking a successful corporate evaluation requires five distinct sets of actions, as shown in Figure 7.3.

**1. Analyzing existing data:** Like other large corporations, MDBs maintain multiple databases containing data on operations, budgets and expenditures, human resources, and, sometimes, on various aspects of operational quality. In IEG's experience, these databases can be a valuable resource for corporate evaluations.

However, more often than not, data generation is a bigger industry than data analysis and utilization. The good news is this means that evaluators often have access to a trove of data. The bad news is that

Figure 7.3: Evaluation actions



with the exception of data produced by the research departments of MDBs, such data may be collected for administrative purposes and may not be organized in a manner that is easily evaluable. Frequent limitations include (i) a lack of data consistency over time or across regions or countries; (ii) timeliness and incompleteness of data; (iii) constraints on the interoperability of databases, such as between budget and human resources data; (iv) frequent absence of data on results; and (v) lack of transparency in the assumptions behind the data.

Since 2001, for example, the World Bank has maintained a computerized database of all financial transactions and a time recording system to manage all internal tasks. The underlying assumption is that staff will charge costs and report time accurately against the charge codes assigned to individual lending, nonlending, and institutional tasks. In practice, tasks across organizational units may not be recorded accurately since one unit may hold the budget code while another that does not control the budget may provide the cross-support. Furthermore, since many tasks involve inputs by staff (whose costs include overheads) and consultants (whose costs exclude overheads), personnel inputs are not easily comparable. Unraveling and correcting for these underlying assumptions is essential for this type of analysis to be useful.



Once these constraints are addressed, analyses of internal or external databases can generate insights that are relevant both for evaluation and subsequent operational use. For example, the merger of data from the budget and human resources databases generated empirical findings for the matrix system evaluation on the extent of cross-support across organizational boundaries. The evaluation found that less than 1% of staff time was spent on collaboration across regions and sector networks, despite the rhetoric about collaboration, and so helped identify the disincentives for collaboration. Such insights informed the design and operating procedures of the Global Practices at the World Bank Group during the subsequent reorganization.<sup>3</sup>

Similarly, the analysis of databases of the Organisation for Economic Co-operation and Development (OECD) on overseas development assistance, foreign direct investments, and remittances helped to understand the rapid growth in overseas development assistance to fragile and conflict-affected states in the past decade (with the World Bank being a laggard). Analysis of OECD data on migration also identified the increasing significance of migration from fragile and conflict-affected states, which has provided jobs and valuable foreign exchange, but has not received much support from development partners. These findings informed the World Bank's dialogue with International Development Association deputies.<sup>4</sup> Evaluations are likely to find greater ownership when they present new analytical insights that have practical implications, in addition to assessing development effectiveness.

**2. Collecting new data:** Clarity on evaluation questions and existing data and information sources helps to identify what new data will be needed, how it can be collected, and what sort of analysis can be undertaken in a timely and cost-effective manner. Whenever new data are needed, it is important to start the process of designing and pretesting research instruments early. Structured instruments and quality assurance processes for data collection in evaluations are essential to ensure consistency and to understand the limitations of evaluations. Typically, data collection involves partnering with local research firms, think tanks, universities, or nongovernment organizations. The pressure of time often makes these partnerships extractive. Local partners are often asked to collect and hand over data, but are not always involved in its subsequent analysis or use. It is good practice to invest more time in building local evaluation capacity and using local experts as continuing resource persons to

help interpret the results as an integral part of an evaluation. It is also good practice to disseminate findings jointly with local evaluation partners after an evaluation is completed.

There is an increasing trend toward including the voices and opinions of stakeholders, including country clients and beneficiaries, to obtain feedback on evaluation impacts—and this practice is yielding better understanding of results. Careful selection of respondents is essential to avoid selection bias, and it is important to be aware of the limitations and generalizability of the data. Client and staff surveys are useful to understand their perceptions, but these also have limitations because they are subjective and partial views of reality. Their feedback needs to be subjected to the same degree of rigor and triangulation as other forms of data. Questions about empirical experience should supplement perception surveys and subjective, interpretive questions. But care must also be taken not to take stakeholder feedback and beneficiary opinions as the final word. Triangulation of evidence from multiple sources is vital to data validation and robust analysis.

**3. Evaluation metrics to assess development effectiveness:** The ultimate purpose of corporate policies and strategies is to increase development effectiveness by enhancing the ability of MDBs to provide better services to clients. To assess the development effectiveness of these corporate policies and strategies, corporate evaluations assess their relevance, effectiveness, efficiency, and sustainability, which are now discussed.

**Relevance:** Organizational inertia characterizes most MDBs, which means corporate policies and strategies tend to change slowly. An important question for corporate evaluations is whether the policies and strategies being evaluated are still relevant or if client demand and internal or external priorities have changed. Moreover, the portfolio itself may have evolved since those policies and strategies were formulated, as happened in the safeguards evaluation. Here, the portfolio evolved from domination by investment projects to the increasing prevalence of budget support operations and sector-wide program loans, but the safeguard policies were ill-adapted to those lending instruments. The evaluation on fragile and conflict-affected states found that the definition of fragility had evolved, but the criteria for identifying fragile countries had not kept pace. As a result, many countries facing fragile and conflict-affected situations, such as

those affected by the Arab Spring, had hitherto been excluded from the official list of fragile and conflict-affected states and fragility risks were ignored. In such cases, operational practice has sometimes gone beyond official policy to deal with emerging challenges—and corporate evaluations need to go beyond the letter of the policy to account for such changes.

**Effectiveness:** Once the corporate objectives have been clearly defined, it should be feasible to assess the extent to which those objectives have been achieved. The absence of a baseline and agreed corporate indicators renders evaluation more challenging than might have been expected, since it is rarely possible to establish a counterfactual. Construction of a results framework *ex post* (discussed earlier) offers a solution to overcome this constraint. In the absence of baseline and pre-existing monitoring indicators, alternative methods to assess the achievement of corporate objectives can include testing of hypotheses formulated during evaluation design or establishing comparators.

The overall objective of the safeguard policies was to ensure that projects proposed for World Bank financing were environmentally and socially sound and sustainable. However, the manner in which the policies were written and implemented emphasized strong up-front analysis and compliance with procedures and standards during project appraisal. The evaluation tested the implicit hypothesis that that strong procedures and compliance during appraisal would suffice to achieve the safeguard policy goals. The evaluation found that most project teams paid relatively little attention to environmental and social outcomes during implementation, which is now leading to a reconsideration of the policy framework and the incentive structure for safeguards work.

The matrix evaluation found that one of the three key objectives of the matrix system—the transformation of the World Bank into a knowledge bank that is mobilizing global knowledge and providing cutting-edge knowledge services to clients—had not been achieved. Although the number of knowledge products produced by the World Bank had multiplied rapidly, they were neither easily accessible nor being used effectively. Lending volume remained the primary measure of success. Since the matrix system evaluation, however, the World Bank has ratcheted up attention to knowledge services and strengthened incentives to leverage its capacity for providing them.

In the evaluation of fragile and conflict-affected states, the results from 33 low-income fragile countries were compared with those from 31 low-income countries that were not fragile. The evaluation demonstrated that over time World Bank support to fragile countries had improved and, by the end of the evaluation period, project outcomes were substantially better in fragile than in nonfragile low-income states (IEG 2014).

**Efficiency:** Assessing efficiency is more difficult and is sometimes not feasible in corporate evaluations. One of the major challenges posed in the safeguards evaluation was that the benefits of the safeguard policies were neither documented nor quantified. And the actual costs of safeguard implementation, which would have to include client costs and project costs, were extremely difficult to find. Internal staff costs on appraisal and supervision were somewhat more readily available. The evaluation undertook two separate activities. For projects whose costs on appraising and supervising environmental and social safeguard issues could be distinctly identified, the data were regressed to test the effects of these expenditures on environmental and social outcomes. These findings were inconclusive, suggesting that administrative costs incurred by the World Bank were not the major driver of outcomes. The evaluation also undertook a benefit-cost modeling exercise to impute benefits from the environmental and social management plan for a few randomly selected projects whose characteristics fell within the normal range of typical projects. These benefits were then compared with the preparation and implementation costs to the World Bank and the costs to the client. While the numbers lacked precision, they did provide orders of magnitude to demonstrate that safeguard benefits significantly outweigh costs.

In the evaluation of fragile and conflict-affected states, assessing efficiency posed a dual challenge—data constraints and the recognition that costs are likely to be significantly higher than in countries where fragility and conflict risks are lower. The high cost of security measures in Afghanistan and Yemen raised the cost of delivering development interventions astronomically. Nonetheless, access to country-level budget and portfolio data made it feasible to analyze trends in expenditure and unit costs for project preparation and supervision and compare the trends for the 33 countries classified as low-income fragile and conflict-affected states with the 31 low-income countries that were not on that list. While unit costs were

notably higher, by 12%–18%, project outcomes in fragile and conflict-affected states significantly improved and were better in these countries than in the other cohort. These improvements suggest that the additional investment in fragile and conflict-affected states was yielding value for money.

**Sustainability:** With a little bit of creative thinking, sustainability is quite easy to examine in corporate evaluations. As illustrated by the following examples, the knack is to ask the right questions:

- **Safeguards evaluation:** Is the delivery model required by the safeguards policies consistent with the resources available and adequate to achieve desired environmental and social outcomes?
- **Matrix system evaluation:** Can knowledge activities be sustained and integrated in operational work if they continue to be heavily dependent on donor-financed trust funds?
- **Fragile and conflict-affected states evaluation:** Can the higher administrative costs of delivery be sustained if concessional financing from the International Development Association to these countries is significantly enhanced?

Such questions make it possible to assess the sustainability of corporate policies and strategies. But formulating these questions requires a thorough understanding of the priorities and constraints that managements and boards have to grapple with. Consultations with both are essential as part of the data gathering process for evaluations.

**4. Formulating recommendations:** Recommendations are often the most contentious part of evaluations since they have consequences for management and, by extension, for clients and development partners. The lessons emerging from good practice suggest that while evaluators should maintain complete autonomy over their evaluation findings, they should not be wedded to the language of a recommendation as long as the spirit of their recommendations is maintained. Furthermore:

- Recommendations should always be based on evaluation findings and focus on a few of the most important ones. IEG

now makes a practice of listing, next to each recommendation, the specific findings that give rise to recommendations in the management action record.

- Recommendations should be clear, realistic, and actionable; exhortative statements tend to be ignored by management.
- Recommendations should take account of management actions. If recent actions or policy and strategic changes have been announced, it is highly unlikely that they will be reversed. But if management is considering several options, a timely recommendation might inform the decision making process.
- Evaluators should be more open to obtaining feedback from management and clients on the language of recommendations. Evaluation findings should be non-negotiable, but if feedback from management or clients indicates that a slight rephrasing of a recommendation would make it more feasible or acceptable—and as long as the spirit of the recommendation does not change, revising the language can enhance stakeholder ownership.

These practices are fully consistent with the reform of the management action record process IEG undertook in consultation with World Bank Group management. Under the revised procedures, after receiving comments on a draft evaluation report, IEG convenes a separate consultative meeting to obtain feedback on the language of IEG's recommendations from management and the organizational unit being evaluated. While the final decision on what to include or exclude rests with IEG, this mechanism often helps to clarify language and focus attention on issues that are more likely to be implementable. After the board has discussed and approved the evaluation's findings and recommendations, management prepares a more detailed action plan to implement the recommendations. IEG subsequently tracks implementation performance over the next 4 years as an integral part of its annual results and performance report.

**5. Maximizing development impact:** The impact of evaluations heavily depends on the manner in which the findings are shared, disseminated, and used. Evaluators, particularly independent evaluators, need to be realistic. Having significant evaluation

findings and recommendations is often insufficient to bring about corporate change. The likelihood of impact is much greater when an organization itself is contemplating change and the evaluation can feed into that process. The key variables that can influence the impact of evaluations are its timing, the alignment of evaluative questions and recommendations with the issues management is grappling with, and the manner of engagement with internal and external clients. The optimal time to undertake an evaluation is when some stakeholders within an organization feel that certain corporate policies and strategies need to be reconsidered, and when early signals suggest that there is appetite for internal reform in the area being evaluated. Early engagement with these stakeholders and with management can provide insights into the questions they are asking, and that the evaluation could examine. Perhaps even more important, however, is the manner of engagement with internal and external clients.

Evaluations have to tread a fine line between maintaining independence and ensuring the relevance of the evaluation itself. For corporate evaluations this cannot be achieved if evaluations are undertaken behind a veil of secrecy. It is crucial to consult with stakeholders early to inform the evaluation design. To ensure that the evaluation addresses current reality, it is equally essential to maintain dialogue with the main actors in the MDB and its clients to stay abreast of any discussions or activities being undertaken. As discussed above, while the evaluation has to arrive at its findings independently, the recommendations can benefit from management feedback to formulate them in a manner that is more likely to be actionable and adopted. Postevaluation activities should go well beyond disseminating evaluation findings through the launch of a report. Besides releasing evaluation findings to the media, when evaluations address policies, strategies, processes, and activities that are relevant to organizational management and change, engaging in those change processes after the evaluation has been completed can be an effective way of increasing development impact.

The safeguards evaluation provided the analytical underpinnings that enabled the World Bank's board to authorize a comprehensive review and update of the safeguard policies (IEG 2010). Continuous dialogue with multiple organizational units at the World Bank and with civil society organizations helped build consensus that safeguard reform is long overdue and helped them understand the finer details of the

evaluation's findings. That said, developing a new policy framework is the responsibility of management.

The launch of the fragile and conflict-affected states evaluation report was organized in collaboration with the G7+ group of these states and was followed by a 2-day retreat with World Bank country directors and managers. The retreat, organized by IEG in collaboration with the World Bank's Center for Conflict, Security and Development, allowed in-depth discussion on seven different topics covered by the evaluation that the managers were grappling with. While the evaluation itself was looking backward at past performance, the discussions were more forward-looking. One of the benefits of this collaborative experience is that the evaluation team was invited back to present its findings and recommendations at several training events organized by management for staff of the World Bank Group and other development partners. The point is, for corporate evaluations to have more lasting impact, evaluation teams need to be more proactive in helping their intended audience interpret and adopt its lessons.

## **Conclusion**

Corporate evaluations present more complex challenges than evaluations of projects or country assistance programs, which use more standardized instruments and evaluation methods. The evaluation design needs a fair degree of customization to the corporate policy or strategy it is evaluating. This chapter has summarized good practices from recent IEG evaluations of World Bank Group performance. It is worth reiterating a few final, implicit messages:

- Evaluations that combine accountability with an attempt to derive lessons for the future of what has worked and what has not tend to have greater ownership. The focus of accountability is primarily retrospective. Identifying forward-looking lessons increases the likelihood of impact.
- A relatively cost-effective way of generating new evidence is to invest resources in analyzing existing management databases and other knowledge products. This can generate new findings that the proprietors of the database may not be aware of and may be keen to cite and use.



- Evaluations that use mixed methods are more credible. An iterative use of quantitative and qualitative methods benefits both. Qualitative methods can help to design quantitative survey instruments and explore the underlying factors leading to the quantitative results.
- The challenge of collecting new data should not be underestimated. It is rarely possible to conduct a corporate evaluation without additional data. It is therefore essential to plan ahead, invest in evaluation instruments, and describe the methodology and limitations of the data used.

## Notes

<sup>1</sup> Examples include safeguard or procurement policies.

<sup>2</sup> Examples include strategic priorities for gender, governance and anticorruption, and fragile and conflict-affected states.

<sup>3</sup> In fiscal year 2014, the World Bank undertook a major organizational restructuring wherein 5,700 sector staff were regrouped into 14 Global Practices and 5 cross-cutting areas (such as gender, climate change, and others) to connect global and local expertise in these practices and areas.

<sup>4</sup> The International Development Association was established in 1960 at the World Bank to provide soft loans to developing countries. Member countries replenish its resources every 3 years. Its credits and grants have totaled \$255 billion, averaging \$15 billion a year in recent years, with Africa receiving 50% of the share.



## Chapter 8

# Influential Evaluations at ADB

*Hemamala Hettige*

**E**valuation *influence* has gained traction as a term for describing the effect of an evaluation. It is broader than the term *use* (Kirkhart 2000; Henry and Melvin 2003). Influence occurs at multiple levels—individual, interpersonal, and collective—and through various channels. At the collective level, evaluation influence manifests itself in organizational practices and decisions. This occurs through routes such as policy and strategy, operational planning, business processes, tools and indicators, and project design and implementation.

Influential evaluations are the first step in managing change. To be influential, an evaluation must be topical and timely—connecting with current and emerging development needs—and it must justify the rationale for the change. The evaluator must engage with affected parties to craft practical and implementable recommendations and communicate evaluation findings to the target audience. Evaluation roles are best fulfilled when influential.

A focus on the factors underlying evaluation influence is crucial for enabling evaluators to comprehend it (Mark and Henry 2004). By reviewing various factors that contribute to influential evaluation, evaluators can be better equipped to work out an influence plan.

### **Evaluator roles in influential evaluation**

How evaluators engage with intended users around evaluation recommendations affects their implementation (Donaldson et al. 2009). Analytical skills matter, but the interpersonal relationships of evaluators, clients, and other users remain crucial to strengthen the influence of evaluations.

Evaluators can serve in varied roles—educator, advisor, facilitator, change agent, accountability advocate, and innovator—to inform an organization how to shape development thinking and bring about changes.

As an educator, the evaluator engages others in dialogue and reflection to clarify assumptions, beliefs, and knowledge. As an advisor, the evaluator guides the stakeholders and decision makers by shedding light on issues, offering options, and proposing future directions. With a diverse skill set and knowledge, evaluators guide the change process. Where the process is owned by stakeholders, it can empower and improve the influence of evaluation (Morabito 2002).

As facilitators, evaluators ease the progress of a process. Facilitators seek understanding and encourage others to share what is important to them (Morabito 2002). As change agents, evaluators focus on tasks that promote discussion, debate, and transformation. To this end, evaluators need to engage participants meaningfully and motivate them to take next steps.

As accountability advocates, evaluators assess if organizations are doing the right things, properly allocating and using resources, and achieving intended outcomes. This role requires feedback on performance and evaluation lessons to boost development effectiveness.

As innovators, evaluators offer fresh ideas. These include proposing new tools and indicators to better assess a situation. Here, evaluators need to understand that risks accompany innovation. Evaluators must encourage clients by not being too harsh on innovation. This can be done by crediting the strategic relevance of interventions.

Evaluation objectives and organizational needs may require evaluators to combine roles to foster influential evaluations. Where development issues are generally misunderstood and action is imperative, evaluators may need multiple hats—educator, facilitator, advisor, and change agent, for example—to reach their target audience, deepen understanding of development imperatives, and initiate required action.

### **Influential evaluations at ADB**

Independent evaluation at ADB recognizes the need for change, timely evaluation results, and stakeholder ownership of influential evaluations. Influential evaluations work through various channels—policy and strategy, operational planning, business processes, tools and indicators, and project design. And this chapter discusses the

types of evaluations that have influenced ADB operations recently, along with the processes that promoted evaluation influence. It centers on (i) channels that work, (ii) processes that stand out, and (iii) engagements with stakeholders that boost evaluation influence.

### **Policy and strategy**

A 2012 evaluation of ADB's knowledge products and services was influential in expediting the organization's efforts to become a stronger knowledge institution (Box 8.1) (IED 2012c). Management was keen to see immediate action, so the timing of the evaluation was favorable. It wanted to engage with the evaluation team in an iterative approach—learning from the ongoing evaluation and contributing to the evaluation. Among the interaction areas were realigning the knowledge management group, creating a knowledge sharing and services center, and preparing the knowledge management directions and an action plan for 2013–2015. Here, the evaluators performed a valuable facilitator role in making sure emerging evidence informed new directions.

#### **Box 8.1: Evaluation of knowledge products and services**

This evaluation identified lessons to help the Asian Development Bank (ADB) become a stronger knowledge institution. It had a twofold approach: looking backward to assess past accomplishments against expectations, and looking forward to determine what features are essential to make ADB more effective as a knowledge institution. The evaluation recommended the following:

- Improve the structure of incentives to better reward staff for doing knowledge work.
- Improve information technology systems.
- Strengthen the identification of knowledge needs.
- Strengthen knowledge sharing by better capturing and sharing tacit knowledge through the increased use of the community of practice network, workshops, and social media.
- Strengthen knowledge use by more extensive distribution of knowledge products and services through the media, seminars, and other forums.
- Prepare a strategic directions document on knowledge management to identify areas in which ADB can add most value through its knowledge work.

*Continued*

### Evaluation influence

- The evaluation sped up the strengthening of ADB as a knowledge institution, as seen in the realignment of the group under the Vice-President for Knowledge Management, creation of the Knowledge Sharing and Services Center within the Regional and Sustainable Development Department, greater focus on knowledge in some regional departments, and monitoring of knowledge issues among staff and stakeholders through surveys. Support from ADB Management for immediate results favored these efforts.
- The evaluation guided the preparation of the knowledge management directions and action plan (2013–2015), which came out in March 2013, and a knowledge forum in July 2013. It also helped shape the new publication guidelines by ADB's Department of External Relations.
- Within the Independent Evaluation Department, other reports (e.g., a governance special evaluation study) adopted this innovative methodology. Moreover, its evaluation findings informed the annual evaluation review and other presentations by evaluators.
- Active engagement with ADB Management and staff during and after the evaluation and communication of evaluation results within and outside ADB helped reinforce evaluation influence. The role of evaluators as facilitators was also very important.

Sources: Evaluation team; Independent Evaluation Department. 2012. *Special Evaluation Study on ADB's Knowledge Products and Services*. Manila: ADB.

Communicating evaluation results took various forms. Evaluators discussed the findings at a meeting hosted by the Agence Française de Développement and made a presentation to the Evaluation Cooperation Group on knowledge and advisory services. A synthesis on knowledge evaluation is under preparation. An article was published in *eVALUatiOn Matters*, the magazine of the African Development Bank's Operations Evaluation Department. These efforts broadened the reach of evaluation findings.

The evaluation of ADB's response to natural disasters and disaster risks was another influential evaluation (IED 2012d). Timely engagement with policy makers and effective interaction were crucial. The change agent role of evaluators helped promote discussions of important issues and urged decision makers to act. For example, discussions with the ADB Board underscored how natural disasters impede development. This led to a policy decision to establish a disaster response facility. The Board also required the pilot facility implementation to adopt the evaluation's recommendations (Box 8.2).

### Box 8.2: Evaluation of response to natural disasters and disaster risks

This evaluation assessed (i) the relevance of the disaster management policy and operations of the Asian Development Bank (ADB), (ii) the responsiveness of disaster risk management and disaster recovery management practiced during 1995–2011, and (iii) the results achieved and anticipated to be achieved in the same period. The evaluation provided important lessons for mainstreaming disaster risk reduction and for disaster recovery projects. It proposed several recommendations:

- Country partnership strategies for all developing member countries at risk from natural disasters must include a natural disaster vulnerability assessment. Appropriate investment programs need to deal with the assessed risks.
- ADB must apply an integrated approach to disaster recovery operations. The primary focus on infrastructure restoration needs to be complemented by activities directed at livelihood restoration and improved resilience of both infrastructure and economic activity.
- ADB needs to coordinate more regularly with other development partners, take more of a leading role in countries where it provides major development support, and undertake more technical assistance jointly.
- ADB needs to integrate climate change and natural disaster activities, and to improve capacity in both areas.
- ADB should review risk finance models and products developed by other disaster risk financing institutions, including those that work with private sector operators.
- ADB's 2008 Action Plan, which had earlier listed appropriate actions to be taken, needs to be updated and refined.

#### Evaluation influence

- The evaluation helped the Board decide to establish the disaster response facility. The Board required ADB Management to adopt the evaluation's recommendations in implementing the facility. The Board also tripled the minimum resource allocation for piloting the facility (from \$1 million to \$3 million) and called for a review of the facility after 3 years.
- It helped in formulating ADB's Operational Plan for Integrated Disaster Risk Management 2014–2020, which seeks to strengthen disaster resilience.
- It heightened public awareness of issues on natural disasters and climate change through extensive media coverage and through evaluators' participation in major global events.
- The evaluation helped implement change by engaging with decision makers to discuss evaluation findings.

Sources: Evaluation team; Independent Evaluation Department. 2012. *Special Evaluation Study on ADB's Response to Natural Disasters and Disaster Risks*. Manila: ADB; ADB. 2014. *Operational Plan for Integrated Disaster Risk Management 2014–2020*. Manila.

Media coverage of the Independent Evaluation Department's work and the participation of evaluation experts have brought our work to a wider audience, particularly an influential evaluation study on disaster risk management. A November 2012 press release on an evaluation study on this issue was picked up by some 40 media organizations, including the New York Times, Los Angeles Times, Nikkei News of Japan, and the Xinhua news agency of the People's Republic of China. ADB evaluators served as speakers and panelists in a number of international conferences and seminars on tackling natural disasters and climate change.<sup>1</sup> These efforts helped draw attention to the growing frequency and intensity of natural disasters in Asia and the Pacific and to the need to take action to build greater resilience to these events.

Another notable example of an influential evaluation was the 2012 annual evaluation review of ADB's operational performance (IED 2012a). It found that ADB's financing for core operational areas in 2011 had far exceeded the 80% target of total financing under Strategy 2020, and had hit 96%. The review caught the attention of the DEC, which concurred with the evaluators' view that the remaining 20% allocation for noncore areas calls for greater compliance. This prompted a Management decision to conduct a midterm review of Strategy 2020 to assess how it has been implemented to date and to offer strategic directions for the future. The independent evaluation report on *Inclusion, Resilience, Change* that was prepared in parallel to the management's midterm review is another example of a very influential evaluation, which benefited from good analysis, timely recommendations, and stakeholder engagement. Much of the evaluation recommendations were absorbed into Management's revised midterm review and presented by the ADB President at the 2014 Annual Meeting.

### **Operational planning**

Evaluations that connect with development needs can influence decision making and also take them to the next step—planning for their operation. For example, the evaluation of support to fragile and conflict-affected situations (FCAS) helped define changes in ADB's engagement in affected countries (Box 8.3) (IED 2010b).



### Box 8.3: Evaluation of support to fragile and conflict-affected situations

The study assessed Asian Development Bank (ADB) support for fragile and conflict-affected situations (FCAS) and drew up recommendations for future engagement in this area. Several actions were recommended to integrate FCAS in ADB operations:

- Recognize the fragility and conflict characteristics of a country, including any that might exist at the subnational levels, during the country partnership strategy preparation.
- Take a different approach in fragile countries from those facing conflict.
- Provide flexibility in FCAS projects and build long-term programmatic approaches.
- Concentrate on capacity development and institution building in FCAS countries through a systematic needs assessment.
- Pay attention to the volatility and any insufficient availability of financial resources for FCAS operations.
- Improve ADB's institutional effectiveness by enhancing staff capacity and skills to make FCAS operations more effective.

#### Evaluation influence

- The evaluation guided the preparation of country partnership strategies.
- As a result of the evaluation recommendation, ADB set up development coordination offices in Kiribati, the Federated States of Micronesia, Palau, the Marshall Islands, Solomon Islands, and Vanuatu to increase its field presence. Joint development coordination offices were also established with the World Bank in Kiribati, Solomon Islands, and Vanuatu to allow closer alignment with governments, development partners, and other stakeholders at the country level.
- Extensive communication of evaluation findings to officials and stakeholders in developing member countries partly explain the evaluation's influence.

Sources: Evaluation team; Independent Evaluation Department. 2010. *Special Evaluation Study on ADB's Support to Fragile and Conflict-Affected Situations*. Manila: ADB; ADB. 2013. *Operational Plan for Enhancing ADB's Effectiveness in Fragile and Conflict-Affected Situations*. Manila.

What made this evaluation influential were its timeliness, good communication, and active evaluator engagement as advisors to operational staff. The evaluation led to various papers calling for a closer look at FCAS countries, informed the handbook that guides FCAS operations, and helped expand ADB's network of FCAS experts. In addition, it shaped the FCAS operational plan, which identified actions to help ADB operations be more FCAS aware (ADB 2013b).

It also led ADB Management to strengthen staff support for FCAS operations under the 3-year workforce plan (2010–2012), in response to the evaluation recommendation to improve field presence and supervision in FCAS countries.

### **Business processes**

Evaluations can shape changes in business processes. The multitranche financing facility (MFF) evaluation completed in 2012 was a real time evaluation of MFF design and implementation (IED 2012b). It provided an opportunity for evaluative feedback on the modality before any completed MFFs had been independently evaluated, and provoked considerable interest from the Board. As a result, the new MFFs submitted for Board approval in September 2013 faced greater scrutiny (Box 8.4). This led to a decision to conduct an informal Board seminar in 2014 on how the evaluation recommendations were implemented. Key topics proposed for discussion include the need for a comparative matrix to justify MFF, fiduciary issues, and a management action plan.

The evaluation was influential because its theme was aligned with current concerns in ADB, and evaluators served as change agents and advisors to improve business procedures. Moreover, findings were communicated to the ADB Board, Management, and staff in a timely manner.

### **Tools and indicators**

In various evaluations of greenhouse gas (GHG) emissions, evaluators served as advisers and change agents in supporting clean energy as well as innovators in developing new tools and indicators. ADB's efforts to mainstream GHG quantification and reduction prior to project approval began well before 2012 when an internationally accepted GHG accounting mechanism was agreed upon (Box 8.5). To reduce the GHG emissions intensity of its portfolio, ADB scaled up the development of appropriate and affordable renewable energy technologies, and set up and managed several climate change mitigation-focused trust funds. ADB also worked with other multilateral development banks to access externally managed climate funds that eased the transition of economies to a low-carbon development path and to improve energy access and security.

#### Box 8.4: Real time evaluation of the multitranche financing facility

Introduced in 2005, the multitranche financing facility (MFF) aims to strengthen the capacity of the Asian Development Bank (ADB) to mobilize development finance and knowledge for developing member countries. Flexibility and client orientation of its financial products and making ADB more compatible with existing and evolving market practices are central to this effort. The MFF modality requires Board approval for the total investment program, but not its individual tranches, which are processed separately. After a 3-year period, the Board of Directors mainstreamed the MFF modality. However, it was concerned about possible lapses in due diligence and risks to implementation and accountability. Subsequently, it requested the Independent Evaluation Department to assess the MFF modality in relation to its development effectiveness.

The evaluation recommended measures that spanned the entire MFF cycle:

- Apply the standards for the needed quality of MFF prerequisites for MFF investment programs as designed at the time of mainstreaming.
- Manage the use of flexibility during implementation without compromising the benefits of the MFF modality.
- Conduct facility-wide midterm reviews of ongoing MFF programs and formal reviews at any time deemed appropriate.
- Regularly monitor MFF portions not converted into tranches and ensure prudent financial planning.
- Ask for regular submission of necessary documentation from clients and make all relevant due diligence information accessible within ADB.

#### Evaluation influence

- Management adopted all recommendations except the one requesting facility-wide midterm reviews.
- Management proposed to establish a set of MFF tranche cancellation criteria and amend Staff Instructions.
- Training sessions on MFF requirements and processing started in September 2013.
- Evaluators' engagement with Management and staff before and after the evaluation in advisory and change agent roles contributed to evaluation influence.

Sources: Independent Evaluation Department (IED). 2013. *Learning Curves: Real Time Evaluation of ADB's Multitranche Financing Facility*. Manila: ADB; IED. 2012. *Real Time Evaluation of the Multitranche Financing Facility*. Manila: ADB.

### Box 8.5: Evaluation of greenhouse gas implications of ADB's sector operations

The evaluation estimated greenhouse gas (GHG) emissions of energy projects approved from 2001 to 2008 for six countries that accounted for 80% of the Asian Development Bank's (ADB) energy portfolio during the period. These were based largely on available project documents. The evaluation developed methodologies for GHG intensity computations or adapted them from methodologies accepted by the Clean Development Mechanism Executive Board. The analysis showed a distinct shift to GHG-efficient investments in the energy portfolio following the launch of the energy efficiency initiative and carbon market initiative in 2006.

To enable ADB to monitor and eventually reduce the GHG emission intensity of its energy operations, the evaluation recommended the following:

- Prepare a consistent framework to identify projects with significant GHG impacts or savings at the concept clearance stage, and assess their GHG implications at appraisal.
- Promote GHG efficient investments by establishing a financing mechanism that buys down in part or full the incremental cost of clean coal technologies.
- Promote scaling up of appropriate and affordable renewable energy technologies.
- Assess opportunities for methane destruction by using methane emissions from coal mining and municipal landfill for power generation or as burner fuel.
- Innovate with lending modalities to facilitate industrial energy efficiency improvements.

#### Evaluation influence

- ADB Management accepted all recommendations.
- ADB is now particularly successful with scaled-up development of appropriate and affordable renewable energy technologies in its developing member countries, which has contributed to the mainstreaming of clean energy in its energy sector operations.
- A follow-up study, which focused on energy efficiency improvements in major energy using sectors such as industry and buildings, enhanced awareness within ADB of innovative modalities such as energy service companies and guarantee funds.
- Continued engagement by the evaluators with the energy community of practice during and after the evaluation in advisory and change agent roles also contributed to evaluation influence.

Sources: Evaluation team; Independent Evaluation Department. 2009. *Evaluation Knowledge Brief on Greenhouse Gas Implications of ADB's Energy Sector Operations*. Manila: ADB.

The evaluation on reducing GHG recommended the use of carbon intensity indicators. International climate change organizations have accepted this concept as the most appropriate method for monitoring carbon emissions for transport projects. ADB continues to consult with other development banks working on related activities to explore the feasibility of including specified indicators for monitoring carbon emissions from transport investments (IED 2013a).

The subsequent evaluation on reducing carbon emissions from transport projects (IED 2010a) developed a tool called the Transport Emissions Evaluation Model for Projects, which measures carbon emissions during the project construction and implementation stages.<sup>2</sup> It has been adopted by other international financial institutions, including the Global Environment Facility and the World Bank, and the Inter-American Development Bank recently started using it to quantify emissions from its transport investment projects. Other positive outcomes relate to its application by Clean Air Asia, an initiative to promote better air quality in the region's cities, to measure black carbon emissions.<sup>3</sup> The Clean Technology Fund projects have adopted it as well, after ADB consultants piloted it in a Ho Chi Minh City metro project.

The evaluation of ADB's social protection strategy was pivotal in calling attention to underinvestment in social protection in Asia. The strategy emphasized the role of social protection in reducing poverty and promoting inclusive growth (IED 2012e), with the evaluators drawing attention to the accountability of ADB in this area (Box 8.6).

As a result of this evaluation, the Board was keen to track ADB's input to social protection and pushed for the prioritization of social protection indicators in the ADB results framework. Processes that helped promote this evaluation included timing, communication of evaluation results, and the active participation by evaluators in various discussions on social protection.

### **Project design**

Evaluation influence was evident in evaluation findings and recommendations being incorporated into improved project designs. The impact evaluation of rural water supply and sanitation in Punjab, Pakistan, drew attention to the almost exclusive focus on improving access to water supply. The evaluation recommended that

### Box 8.6: Evaluation of social protection strategy

The study evaluated the Asian Development Bank's (ADB) experience in social protection in Asia and the Pacific. It reviewed how ADB responded to the 2001-approved Social Protection Strategy and what needs to be done to take this agenda forward. It concluded that social protection needs to be an integral part of ADB's corporate strategy to meet the challenges posed by growing inequality, demographic and social changes, and crises. It urged ADB to connect social protection with its core areas of investment to increase its presence in building social protection systems, and it emphasized the need for partnerships and capacity development.

#### Evaluation influence

- The evaluation emphasized underinvestment in social protection in the region and highlighted ADB's accountability in this area.
- An important evaluation result was the prioritization of social protection indicators.
- The design of support for safety nets has already started in Pakistan, in partnership with the World Bank and the Department for International Development of the United Kingdom.
- The active participation of evaluators in discussions on ADB's results framework for inclusive growth and social protection and the communication of evaluation results contributed to evaluation influence.

Sources: Evaluation team; Independent Evaluation Department. 2012. *Special Evaluation Study on ADB's Social Protection Strategy 2001*. Manila: ADB.

future investments in water devote more attention to sanitation. It highlighted that the project design must demonstrate strong synergy between water supply systems and improvements of drainage and street pavements, minimize sanitation hazards in and around water supply systems, and include solid waste and wastewater management (IED 2009). The evaluation proposed that a baseline study be undertaken during project preparation with benchmark indicators for results monitoring and evaluation. To track sustainability, the study recommended postproject monitoring of project benefits.

The priorities of ADB's Water Operational Plan 2011–2020 and of other recent water projects have resonated with evaluation recommendations on sanitation.<sup>4</sup> For example, the project for improving community-based rural water supply and sanitation in Sri Lanka's Jaffna and Kilinochchi districts supported a phased sanitation system comprising a sewage collection system, sewage treatment plant, treated effluent sea outfall, and maintenance equipment. It

also included baseline studies and awareness campaigns for hygiene and sanitation practices (ADB 2011a).

Similarly, in the Lao People's Democratic Republic, sanitation-related features were integrated into the water supply and sanitation project (ADB 2013c). These comprised facilities for public sanitation and septage disposal, improved household sanitation, and hygiene awareness. The project's design also incorporated lessons from ADB's independent evaluations, including those relating to good practices in water supply and sanitation.

## **Conclusion**

Influential evaluations are not accidents. They require careful planning and execution. Understanding how evaluations lead to better decisions and their implementation is essential. Evaluators must reflect on the channels that shape changes in policy and practice. The routes to influential evaluation include directing changes to policy and strategy, proposing measures to improve business processes, and offering fresh insights to solve existing issues.

Even the highest quality evaluations do not automatically guarantee action on findings and recommendations. Three key factors contribute to influential evaluations: connecting with current and emerging development imperatives, engaging actively with stakeholders, and interacting and communicating widely.

Key evaluation messages must provide timely and credible results that decision makers can draw on, and encourage interaction on development imperatives. Opening communication channels on evaluation findings boosts transformative and innovative processes. Such channels shape critical thinking that guides timely decisions and actions and also nurture change processes to improve development effectiveness. Influential evaluations matter.

## Notes

<sup>1</sup> These included (i) International Seminar on Lessons from Mega-Disaster hosted by the World Bank and the International Monetary Fund in Tokyo on 14 October 2012, (ii) the United Nations Conference on Climate Change in Doha on 29 November 2012, (iii) the Asia Regional Press Conference of the Centre for Research on the Epidemiology of Disasters on 11 December 2012 in Bangkok, and (iv) Asian Congress on Citizen and Environment Safety, co-hosted by the National University of Singapore, 5–7 June 2013.

<sup>2</sup> The Transport Emissions Evaluation Model for Projects considers passenger and freight travel activity, the shares of trips by different modes and vehicle types (structure), fuel carbon dioxide efficiency (intensity), and fuel type, validated by more detailed emission factor models. The models consider induced traffic demand generated by changes in the generalized time and money cost of travel by different modes, building on best practice analysis techniques.

<sup>2</sup> Clean Air Asia advances innovative ways to improve air quality in cities by sharing knowledge and experiences through partnerships. It brings together cross-cutting expertise in urban development, transport, energy reform, environmental management, and environmental health. Its network covers the People's Republic of China, India, Indonesia, Nepal, Pakistan, the Philippines, Sri Lanka, and Viet Nam.

<sup>4</sup> The Water Operational Plan states that “ADB will pursue investments in sanitation infrastructure and services wherever required to secure higher gains to public health . . . The scaling-up of sanitation will be a key activity in the surge to deliver progress against the MDG targets and will be central to the success of future urban and rural water operations” (ADB 2011b).



## Chapter 9

# Developing a Gender-Responsive Monitoring and Evaluation System for International Development Agencies

*Michael Bamberger*

In Asia, as in other regions, every society has rules governing appropriate behavior for men and women and boys and girls in the home, the community, the labor market, schools, and in politics. Some of these rules are regulated by social customs, others by laws or the operation of the labor market. The forms of control can be subtle, enforced by legal sanctions, or even the threat of violence. A particularly worrying example in several Asian countries is the preference for male children and female infanticide. While some sectors of society believe these rules to be based on “natural” differences between men and women, they are in fact socially constructed and vary from one society to another and over time. However, despite differences across societies, these rules place women at a disadvantage in key dimensions of development. This chapter draws on the international experience of governments, donor agencies, and nongovernment organizations to outline the main steps in the design and implementation of gender-responsive monitoring and evaluation systems.

The persistence of significant gender inequalities in all regions negates fundamental human rights and the expansion of human freedoms. Furthermore, gender inequalities are very significant barriers to the achievement of development objectives (Box 9.1).

The influential 2011 *World Development Report on Gender Equality and Development* argues that promoting gender equity can make a major contribution to development in three major areas: by fully utilizing the capacities of both women and men, through improved development outcomes for the next generation, and by broadening policy choices by making institutions more representative (World Bank 2012).

### Box 9.1: The economic cost of gender inequality: examples from Africa, the Middle East, and Asia

- Asia is losing about \$44 billion per year due to limited access to employment opportunities for women.
- Asia is losing \$16 billion–\$30 billion from limited access to education for girls.
- If India increases the ratio of female to male workers by 10%, gross domestic product would increase 8%.
- In the Middle East and North Africa, if the labor force participation of women had increased in the 1990s at the same rate as women's education, the average household income would have been 25% higher.
- Tanzania could increase growth by 1% by removing barriers to female entrepreneurs.
- Total agricultural output in sub-Saharan Africa could increase by 6%–20% if women's access to agricultural inputs was equal to men's.

Sources: Department for International Development of the United Kingdom. 2008. *The Gender Manual: A Practical Guide*. London; United Nations Economic and Social Commission for Asia and the Pacific. 2007. *Economic and Social Survey of Asia and the Pacific*. Bangkok.

In light of this compelling argument, many governments and international development agencies now recognize gender equality as a priority development objective. Achieving gender equality requires integrating gender into all aspects of programming, budgeting, implementation, and monitoring and evaluation (M&E). Many organizations have operationalized their gender equality strategies through gender action plans that stress the critical role of appropriate gender tools for data collection and analysis for M&E. While a gender action plan provides a useful framework for an integrated approach to gender equality, it is nevertheless not essential, and many agencies begin by building gender into existing M&E systems. After that, having gained experience with gender M&E, they can develop a broader gender framework.

### **Increasing the international community's commitment to gender equality**

Over the past few years, the international community has increasingly recognized the importance of gender equality as a development objective in its own right and as an essential element to achieve other development goals, such as sustained and equitable growth, more efficient program implementation, and better governance.

This is reflected in the implementation by the United Nations, multilateral and bilateral development agencies, and international nongovernment organizations of new or significantly updated gender policies and gender action plans.<sup>1</sup> The rationale for the greater focus on gender equality was spelled out in the 2011 *World Development Report on Gender Equality and Development*, which made both the business case—gender as smart economics—and in terms of fundamental human rights and social justice. This report has been cited widely by other agencies and there is now a broad consensus on both the rationale for gender equality and the elements of a gender equality strategy. A recent example of the growing commitment of development institutions to gender equality is the fact that the March 2014 edition of the African Development Bank’s *eVALUatiOn Matters* quarterly journal, published by its Operations Evaluation Department, was titled “Gender Inequality and You.”

### **ADB is among the pioneers in promoting gender equality**

The Asian Development Bank (ADB), in promoting gender equality as one of its central development objectives, has been in the vanguard. Following a systematic assessment of the lessons learned from its earlier and more limited gender strategies, the organization in 2007 launched its Gender and Development Plan of Action. It has periodically updated the plan and approved it to guide actions in three core areas for 2011–2012: country partnership strategies and lending operations, policy dialogue and gender capacity support to developing member countries, and organizational effectiveness.

ADB’s Gender Equality and Women’s Empowerment Operational Plan, 2013–2020 continues to move the agenda forward, with gender equality defined as one of the five drivers of change in the institution’s long-term Strategy 2020:

*ADB recognizes that without harnessing the talents, human capital, and economic potential of women, the region risks significant economic and social costs. Gender equality is critical in its own right and essential for better development outcomes in terms of inclusive growth, faster poverty reduction, and attaining the Millennium Development Goals (ADB 2013a).*

ADB's revised corporate results framework also includes gender mainstreaming targets of 45% in all its operations (ADB 2013a). ADB also pioneered promoting gender equality through its rapid gender assessment methodology, which is considered a very cost-effective way to assess gender outcomes at the project and sector level.

Despite the wide recognition of gender equality as a development objective, however, many M&E systems do not adequately measure differences in development outcomes for women and men and girls and boys. Important issues that are often not captured include the following:

- how a woman's multiple productive, reproductive, and community maintenance roles limit her ability to participate in and benefit from development initiatives;
- women's access to and control of productive resources;
- constraints on women's access to program benefits;
- gender-based violence; and
- women's participation in decision making at the household, community, and national levels.

M&E systems fail to capture gender differences for many reasons. Managers and staff in some sectors believe that men and women benefit equally from well-designed programs and, consequently, gender analysis is not required. Some gender issues are also considered culturally sensitive, so some agencies may be reluctant to address them. A number of methodological issues can also be particularly challenging for gender analysis:

- Gender roles, processes, and outcomes are affected by a wide range of economic, social, political, legal, and psychological factors—all of which must be considered in the M&E system.
- Gender processes and outcomes are often difficult to measure. Many processes concern sensitive issues such as domestic violence, sexual harassment in public space, power relationships and ownership and control of household or community resources, sexual behavior, and mechanisms

for the spread of HIV/AIDS. These are difficult to study with conventional quantitative surveys as many people are reluctant to discuss or to report honestly. Furthermore, many of these behaviors take place in contexts, such as the household, where it is difficult for the researcher to be present.

- A blind spot for many M&E systems is that they are only designed to assess the extent to which the intended outcomes of government programs are achieved, but do not look for unintended outcomes (Bamberger 2012). This is a serious problem for gender analysis as many interventions can have serious negative consequences for some groups of women or men. For example, when women obtain credit to start a business, some husbands resent their spouses' greater economic independence, which may lead to increased domestic violence.

### **Constructing a gender-responsive monitoring and evaluation strategy**

To fully implement an agency-wide gender equality policy, it is necessary to develop an integrated gender equality M&E strategy (an engendered M&E strategy). This will ensure that gender objectives are built into all stages of global, country, sector, and project planning; that appropriate performance measurement indicators are defined; that gender is built into results frameworks; and that engendered M&E systems are in place.

Figure 9.1 illustrates the typical steps in setting up a gender-responsive M&E system. For reasons of simplicity, the example only illustrates how such a system would be created for a country program. An additional level of complexity is required to create an organization-wide system, as each organization would need to adapt the structure and implementation process to its own policy and planning framework. The main steps for launching a gender-responsive M&E system are the following:

#### **Create a country gender strategy and gender action plan.**

This reviews the development agency, the government's national and sectoral development objectives, and the gender dimensions of each objective. It also provides a background on the country gender context and prioritizes intended gender outcomes.

The action plan defines (i) gender outcomes to be achieved, (ii) gender-related analytical work, (iii) gender issues for policy dialogue, (iv) targets for gender mainstreaming in specific sectors and operations, and (v) areas where development agencies or their national partners require additional gender training and capacity building. Ideally, for each intended output the gender action plan should include an implementation plan that identifies the gender output indicators, baseline reference points, targets, budgets, and timelines, as well as assigns responsibilities for data collection. The plan also outlines the proposed gender M&E strategy and the key impacts, outcomes, and output indicators to be tracked. In some cases, it may not be possible or appropriate to develop a full gender action plan and in these cases a gender-responsive M&E strategy should only be developed for priority sectors or programs. In these cases, the goal would be to gradually broaden the focus until a full gender action plan can be developed.

**Define gender objectives at different levels of country programs.** Gender objectives are identified for different levels, which typically include the following:

- contribution to global development goals, such as the Millennium Development Goals;
- contributions to the objectives of a country program and/or budget priorities (including strategies of other donor), policy dialogue, sector programs and projects, and institutional building and capacity development); and
- organizational effectiveness; that is, how well gender interventions are planned and monitored, the adequacy of budget allocations, staff awareness, and capacity development.

**Develop a gender analysis framework.** This draws on international gender and feminist literature to identify the main dimensions and issues that the gender analysis must address. These are then translated into performance indicators used to assess the gender-responsiveness of country program activities. Some of the broad issues built into the framework are the following:

- The multiple productive, reproductive, and social maintenance roles of women and how these limit their ability to fully participate in and benefit from development programs.
- Closely related to the above is women's time use and time poverty.
- Gender differences in access to and control over productive resources, including financial resources at the household, community, and wider levels.
- Gender differences in control over decision making at the household and wider levels, including participation in the community and wider political processes.
- Gender differences in access to labor markets, including labor market segmentation, wages and salaries, and the promotion of economic opportunities.
- Legal, political, economic, cultural, and other constraints on the full participation of both women and men.
- The causes, magnitude, and consequences of domestic and gender-based violence. This includes issues such as human trafficking and the sex trade, child marriage, and, in some countries, female genital mutilation. A full analysis should include psychological and well as physical violence.

**Develop a theory of change.** A gender theory of change identifies problems constraining the achievement of gender equality, and outlines processes and mechanisms through which interventions are expected to achieve their intended gender outputs, outcomes, and impacts. It also identifies the key assumptions that should be tested at each level. Many theories of change also identify economic, political, social, environmental, and psychological factors that can affect outcomes (positively and negatively). Recent literature has emphasized the importance of spelling out the mechanisms through which change is expected to be produced (Astbury and Leeuw 2010). For example, how are training programs for women on money management expected to help create female-owned businesses?

Some theories of change use a three-step process that includes a problem tree, a solutions tree, and the fully articulated theory.

Ideally, a gender theory of change should begin with a diagnostic study to provide a better understanding of the nature of gender relations and constraints in the target areas (Funnell and Rogers 2011). It should also include participatory consultations with both female and male stakeholders, including the most vulnerable groups, on their concerns and priorities. The three guidance notes of the Department for International Development of the United Kingdom on the design of programs to tackle violence against women illustrate how the theory of change can be used to inform program design as well as for designing gender-responsive M&E systems. It is useful to simplify the theory of change into a logical framework that represents in graphical form program activities, outputs, outcomes (sometimes called program objectives), and the goals or final outcomes. The logical framework provides a useful link to the development of the program results framework.

**Identify gender-responsive indicators.** Indicators define what needs to be measured in a way that is economical and technically sound and that adequately describes the constructs being studied. The main kinds of indicators for M&E include: inputs, processes, outputs, outcomes, and contextual factors affecting outcomes. Examples of the latter could include whether an economy is growing or declining, levels of conflict or violence in a community, and whether local authorities and political groups are supportive of the gender program.

Gender analysis requires that standard indicators be adapted to capture differences between women and men for each kind of indicator and for other relevant categories such as youth, the elderly, religious and ethnic minorities, and so on. Most gender indicators will be obtained by disaggregating standard indicators by sex (for example, the number of girls and boys enrolled in school, or female and male farmers visited by extension workers). However, additional indicators are usually required to address specific gender dimensions, such as time use and control of productive resources that are not included in conventional M&E systems.

Assessing the quality of services is particularly important for gender analysis because the design and delivery of program services are



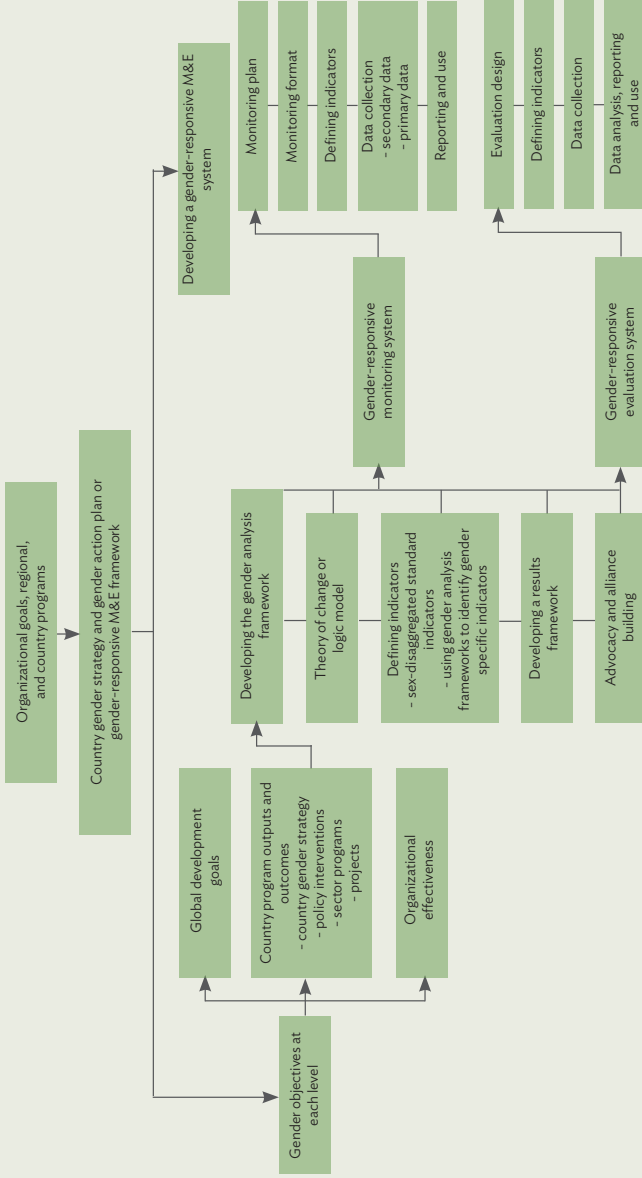
often not adapted to the special needs of women. For example, many programs do not make provision for child care for women attending program-related meetings or project workdays, or selecting locales that women can easily reach or where they feel comfortable.

While sex disaggregation of standard indicators is an important and economical way to begin to examine gender differences in program implementation and outcomes, there are many important gender dimensions that dig deeper to examine how gender relations affect development outcomes. Gender analysis provides tools to help understand the underlying causes of gender inequalities and examines how gender rules determine the economic and social roles and opportunities of different groups of women and men, and how these affect their ability to participate in and contribute to development. A gender analysis framework helps define the gender-responsive indicators and measures to be included in a monitoring system. There are many different gender analysis frameworks,<sup>2</sup> each focusing on different aspects of women's (and sometimes men's) socially defined roles and how these affect their participation in household, economic, and political activities, including development programs.

The *Harvard Gender Analysis Framework* is one of the most widely used.<sup>3</sup> It argues that women's participation in and enjoyment of the benefits of development projects is constrained by their heavy time burdens, resulting from their multiple roles in the production of goods and services, reproduction and maintenance of human resources (reproduction and care for family members), responsibilities for maintaining community resources, and by unequal access to and control of productive resources. These are measured through two instruments: an "activity profile" and an "access and control profile." These quantify the number of hours a week household members spend on different activities and rate the level of access and control household members have over different resources (land, equipment, labor, capital, animals, and so on). This framework can be used to measure changes in women's time use and control of resources at the start and end of projects.

**Build gender into the results framework.** Most development agencies now use results-based M&E systems to monitor implementation and outputs and to evaluate program effectiveness. According to Kusek and Rist (2004), the key elements of a results-based M&E system include the following:

**Figure 9.1: Stages in the development of a national gender-responsive M&E framework**



M&E = monitoring and evaluation.  
 Source: Adapted from Bambergler. 2013. Engendering M&E: Nuts and Bolts of Monitoring and Evaluation Systems. Series No. 27. World Bank.

- baseline data to describe the problem or situation before the intervention,
- indicators for outcomes,
- data collection on outputs and how and whether they contribute to achievement of outcomes,
- more focus on perceptions of change among stakeholders,
- systematic reporting with more quantitative and qualitative information on progress toward outcomes, and
- cooperation with strategic partners and information on the success or failure of a partnership strategy in achieving desired outcomes.

For the gender analysis framework all appropriate indicators must be sex-disaggregated, and additional gender-specific indicators included together with information on how gender data will be collected (as, for example, in the earlier *Harvard Gender Analysis Framework*).

**Use advocacy and alliance building.** Implementing gender-responsive M&E systems will frequently encounter challenges and even opposition. Some challenges will be technical (lack of gender analysis expertise) and some will be logistical (problems in convincing agencies to change procedures or add resources to collect sex-disaggregated data). Some challenges will stem from opposition to promoting women's rights or reluctance to collect data on sensitive issues, such as domestic violence, labor market inequities, or a woman's control over her reproductive choices. Dealing with these challenges will often require developing alliances with agencies that have technical expertise in gender analysis and enlisting the support of women's organizations and other groups that can help convince governments and other partner agencies to adopt a gender-responsive data collection and analysis approach.<sup>4</sup>

**Develop a specifically gender-responsive M&E system.** For most development programs, a gender monitoring system looks very much like a conventional monitoring system with some additional questions to measure differences in how women and men participate in and are affected by projects. Much of the additional information is obtained

by ensuring that standard information on program participation or outputs is sex-disaggregated. However, it will often be necessary to collect additional information not included in conventional monitoring. For example, there may be questions on special gender needs, barriers to participation in a particular program, and the services or resources provided to address these differences. Some of this information, for example, on family resistance to the participation of women or girls, women's lack of control over economic resources, or threats of gender-based violence against women who attend meetings, may require special data collection methods, such as focus groups, in-depth interviews, using informants, or observation.

In cases where a program's primary objective is to achieve changes in gender relations or to address a problem that specifically affects women, such as gender-based violence, the monitoring system may rely more on qualitative methods, such as case studies, in-depth interviews, participant observation, or focus groups.

Gender outcomes and impacts can be evaluated in several ways:

- by including gender indicators in a standard impact evaluation design;
- by adding a gender-specific module or data collection method to an impact evaluation design (for example, administering a special module to women in a sub-sample of households when only the "household head," in most cases male, is interviewed);
- by implementing a special gender impact evaluation; this option may be appropriate where gender outcomes are the primary program goal.

The whole range of conventional impact evaluation designs can be used for the first two options. However, until recently, many gender impact evaluations used predominantly qualitative or mixed methods in their design, partly because many of these projects were quite small-scale. But now that gender is becoming a central policy objective of many international donors and nongovernment organizations, and the scale of gender-related interventions has increased, there is increasing use of experimental and quasi-experimental designs for evaluating gender outcomes. For a discussion of the steps for designing gender-responsive M&E systems, see Bamberger (2013).

## Conclusion

Women and men are affected by and respond to development interventions in significantly different ways. But despite compelling evidence of persistent gender inequalities, and the consequences for development programs, conventional M&E systems frequently fail to address gender differences. Consequently, organizational commitment to achieving gender equality requires developing gender-responsive monitoring and evaluation systems. ADB should be congratulated for being among the pioneers in designing and implementing a gender policy and gender action plan and in developing rapid gender evaluation methods for providing rapid and cost-effective ways for evaluation to help achieve gender objectives in projects and programs.

Despite significant progress in recent years, many challenges remain before such systems are fully implemented and used. Some relate to limited financial resources, others to lack of technical expertise or to the logistics of collecting quality data on gender indicators. One of the biggest challenges is the reluctance of development agencies and their partners to fully commit to the often sensitive and controversial issues relating to the political, legal, economic, sociocultural, and psychological factors constraining women's (and often men's) full participation in and enjoyment of the benefits of development. Significant progress has been made but much remains to be done.

## Notes

<sup>1</sup> Examples of institutions and other bodies that have done this include the African Development Bank, Asian Development Bank, Inter-American Development Bank, United Nations Development Programme, World Bank, Canadian International Development Agency, Department for International Development of the United Kingdom, United States Agency for International Development, Oxfam USA, Oxfam Canada, and Care International.

<sup>2</sup> For a review of gender analysis frameworks see Williams et al. (1994) (see especially Module C.7 on Analytical Frameworks. The Gender in Agriculture Sourcebook (World Bank, Food and Agriculture Organization, and International Fund for Agricultural Development 2009) illustrates how more specific gender analysis frameworks can be designed for areas such as rural finance, land use, fisheries, livestock, and others. For a review of more academic feminist frameworks, see Hesse-Biber (2012).

<sup>3</sup> See Williams et al. (1994), Overholt et al. (1985), and Rao et al. (1995).

<sup>4</sup> For a detailed discussion of advocacy strategies to promote evaluation (in general, not just for gender responsive approaches), see Karkara (2013).





**Part 4**

**Strengthening  
Evaluation Capacity  
and Country Evaluation  
Systems**

As attention to development results and effectiveness rises, so too have calls for helping establish the skills for carrying out the needed functions. The common challenge for almost all developing countries in the region is the lack of evaluation and management capacity in the public sector.

In response, several countries have launched results-based and evaluation functions in recent years and have begun to gather experience. Among them, India created the Independent Evaluation Office in 2011, while Malaysia began applying results-based management approaches to budgeting in the same year. The People's Republic of China (PRC) has given considerable attention to performance-based management since the 1990s and has joined forces with several organizations to provide training in Shanghai to raise capacity in the region. This section explores these experiences.

As detailed in Chapter 10, improving the effectiveness of government policies and programs for socioeconomic development is one of the objectives of India's new Independent Evaluation Office. With full autonomy, it began undertaking impartial assessments of programs that affect public resources in 2013.

India's inability to implement and deliver public services is well known, notes Ajay Chhibber, director general of the office, "but exactly why this is so—be it in education, health care, water, roads, and electricity—is harder to explain." The country has many national flagship programs with ambitious goals, but their delivery remains uneven.

Chapter 11 examines Malaysia's experience in adopting managing for development results through its participation in the Asia-Pacific Community of Practice on Managing for Development Results, a peer learning initiative promoted by the Asian Development Bank. It was one of the first countries in the region to adopt the practice using an integrated planning, budgeting, and monitoring and evaluation model.

To build capacity in performance evaluation, promote effective use of public funds, and learn from international experience in development evaluation, meanwhile, the Ministry of Finance in the PRC joined forces in 2007 with the World Bank, Asian Development Bank, and the Asia-Pacific Finance and Development Center to launch the Shanghai International Program for Development Evaluation Training (SHIPDET). Over time, SHIPDET has built its reputation as a reliable training provider in performance evaluation, which augurs well for capacity development in the region.



## Chapter 10

# India's Independent Evaluation Office: A New Approach

*Ajay Chhibber*

India suffers from an inability to deliver public services effectively. Implementation bottlenecks and a controlled top-down approach to public governance characterize the country's pattern of development. It is now widely accepted that the state needs to be modernized and made fit to tackle the demands and emerging challenges of the 21st century.

Improving the effectiveness of government policies and programs for socioeconomic development is crucial to this effort—and this is one of the objectives of India's new Independent Evaluation Office. Created by the Cabinet in 2011, the office is attached to the Planning Commission and enjoys full autonomy in discharging its functions. It began doing so in 2013 to carry out its mandate of undertaking impartial assessments of programs that affect public resources.<sup>1</sup> It also aims at encouraging a culture of openness and learning in government and adopting good international practices in improving India's evaluation system (Planning Commission of India 2013).

Before establishing an evaluation agenda, it would be useful to step back and briefly consider what should be the relevant framework for an evaluation system that meets India's needs and the right approach to developing a work program for evaluation. Here, an important issue will be how topics are selected for evaluation that would be both tractable and most relevant for the issues that policy makers and systems face today.

More broadly, it may be useful to ask: what should the role of the state be at this stage of India's development; what should be the appropriate division between central, state, and local government; what should be the appropriate separation between the judiciary and the executive; how can government be made more capable to deliver public services through improved institutional capacity; and how can

the state provide better protection and safety nets for its citizens in the most cost-effective manner?

In setting up an evaluation agenda for the Independent Evaluation Office, it is my belief that it should focus, at least initially, on the following development challenges facing the Indian state and that are topical for evaluation: delivering public services and flagship programs; social protection, inequity, and disaster management; and regulation and economic management and public-private interface. Connecting with these challenges is of utmost importance in making the evaluation agenda relevant and meaningful. It sets the stage for tackling priority areas where evaluation can provide evidence-based findings on outcomes and impacts and offer forward-looking evaluation recommendations. Looking for fresh solutions to persistent development problems and promoting learning, accountability, and innovation can help make a difference in achieving development effectiveness.

### **Delivering public services and flagship programs**

India's inability to implement and deliver public services is well known, but exactly why this is so—be it in education, health care, water, roads, and electricity—is harder to explain. India has many national flagship programs with ambitious goals, but their delivery remains uneven. It is very often said that the problems of delivery lie at the state or local (district) level.

Decentralization and devolution are helping bring government closer to the people, but state and local capacity remains very varied and local governance is uneven. In the 2013–2014 fiscal year, around \$30 billion was budgeted for flagship programs—a significant amount intended to improve a range of basic services, such as education, health, sanitation, child support, maternal health, rural roads, access to electricity, housing, and guaranteed employment.<sup>2</sup> But the results have been quite varied across states and districts and, overall, mixed at best.

Various evaluations and assessments have suggested how to improve the effectiveness of these programs. Interestingly, some states, including Tamil Nadu, Gujarat, Karnataka, Kerala, and Andhra Pradesh, perform well on almost all the programs, whereas others, including Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh, and

Assam, perform poorly. That some states are able to take nationally designed schemes and make them work and others unable to do so raises a fundamental question for evaluators striving to improve the delivery of public services—why should we expect consistent performance on national programs across states with big differences in implementation capability as well as standards of governance and geographical and cultural differences? Surely nationally designed programs with standard implementation templates would give very different results across states, especially in the social sectors where the services being delivered have a high number of transactions with a lot of discretion. Providing education falls into this category because delivering quality education requires a number of factors to come together, including classrooms with toilets, qualified teachers, teaching material and an appropriate curriculum, transport, and midday meals. Problems in any of these inputs could compromise the quality of education.

Similarly on health delivery, trained nurses, medical equipment and medicines, health awareness, accessibility to health centers, knowledge of basic hygiene, and access to water and electricity are vital for good health outcomes. Coordinated and depoliticized decision making on a number of issues is needed to ensure proper service delivery. If health workers, for example, with political connections are hired to provide jobs to party workers then, even if everything else is done right, the system will fail.

Social services not only have high transaction and coordination needs, but they also have low specificity; in this sense, it is not easy to immediately measure their output. It takes time to measure whether a health or an education system is delivering a quality service. It is hard to trace back what is missing even if it is identifiable because it is difficult to specify. In other words, it has low specificity—and the impact of the program takes time to identify and measure.

By contrast, building rural roads or electrification has lower transactions but high specificity, making it easier to measure and monitor output for these types of projects. Indeed, precise specifications are required for rural roads or electricity lines, and monitors can come in quickly once a road is built to check if it meets the required specification (even before final payment has been made). The same cannot be done for education or health, other than sending observers into classrooms, but even then the learning outcomes may not be known for some time.

So we need to think more smartly about how to deliver education and health in a way quite different from roads, electricity, and water infrastructure. National schemes with very standardized specifications have very little modifiability in the way they are implemented and will lead to very different results. It is therefore not surprising that the Pradhan Mantri Gram Sadak Yojana has a better record of implementation across the country than any of the social schemes.

The Mid-day Meal Scheme is an interesting mixture of very high transactions and lack of specificity. It requires the meal to be cooked fresh on site, at a cheap price but not by a commercial company. So it requires considerable local capacity to ensure that it can be delivered consistently and safely across thousands of schools. What is surprising is that we have not seen more tragedies, such as the Bihar food poisoning incident. Where civil society is strong and parent-teacher associations are involved, the scheme does well, but in remote areas where these conditions often do not exist, there is a lot of corruption in the program and misuse is the outcome. The midday meal scheme is clearly ripe for new solutions.

So where do we go from here? One solution that has been suggested is to do away with national schemes altogether and transfer the funds to provide services to the states or to the *sarpanch* (local elected official) at the local government level. But this by itself will not solve the issue of variability of performance. The same issues which bedevil national schemes—weak institutional capability, local political capture, and lack of accountability—would likely come into play and affect how well the disbursed funds will be used.

Instead, we need to think more constructively about how to design national programs with a more imaginative combination of standardization, flexibility, and incentives that would bring back the focus on better outcomes, rather than a rigid set of national procedures that do not allow innovation and do not deliver services to people in many parts of the country. So, for the Mid-day Meal Scheme, if it is not easy to provide a cooked midday meal safely to children, it may be better to allow precooked, nutritious meals to be distributed to children. And where social workers (*anganwadi*) are being hired on a political rather than technical basis, we must ask whether a different model for delivering health services to children and their mothers is needed.

Local governance capacity is critical for implementing national programs and, as noted earlier, there appears to be considerable variability within states in how programs are delivered. In Bihar, which did not do well in the past in delivering national programs, Nalanda district performed much better than others. Kullu district in Himachal Pradesh is one of India's best performing districts on program implementation, not only because of local administration, but also because the elected district councils (*Zila Parishads*) play a very proactive role in anticipating funds from national flagship programs and planning their use with community involvement.

In states with weak institutions and implementation capacity, capacity building programs and much stronger monitoring need to build into the schemes. A more enlightened debate among the center and the states and local government is also needed on how to change the focus from expenditure to achieving better outcomes. There are also grounds to consider rationalizing further the number of schemes to avoid duplication and overlapping mandates. The current top-down model of national schemes with prespecified and rigid requirements and procedures clearly needs rethinking and a major reboot if India is to achieve its national goals and the internationally accepted Millennium Development Goals.

Instead of assessing scheme by scheme, we need to think of evaluations by outcome. Progress on infant and maternal mortality, nutrition, and school learning is affected by socioeconomic variables and the impact of various schemes. However, it is not easy to construct a model that will unravel these impacts. Simpler evaluations of assessing variability in birth attendance or teacher attendance should also be considered.

### **Social protection, inequity, and disaster management**

Social protection is a key objective of the Indian state, and comprises social assistance and social insurance. The latter is still being developed and therefore not yet ready for evaluation, but the country has had a large social assistance program for a long time, largely built around product-based subsidies.

Over the last 5 years, the explicit subsidy bill in the budget has increased by almost 1% of gross domestic product and now stands at almost 2.4%. Implicit subsidies could be even higher. Subsidies have evolved into a rights-based approach, but there is very little

focus on how to deliver social protection in an efficient and cost-effective manner. As a result, it is claimed that there is a great deal of leakage in India's social assistance programs and a lack of targeting that causes market distortions. In the case of food, for example, the public distribution system, which has large leakage, will be tested further with a food security bill. But it is not clear how much of the food subsidy bill comes from the subsidy itself and how much comes from the public handling and storage of large reserves of cereals by the Food Corporation of India. A careful evaluation of India's food procurement and transport policy must also be made to rationalize costs and look for alternative solutions.

The country's fertilizer subsidy policy is also costing huge resources and some say that it is not helping farmers, and that, particularly, many small farmers do not benefit from it as they do not use much fertilizer. The partial decontrol of fertilizer prices is distorting the fertilizer mix used in the country and thereby could even do long-term damage to soil productivity. The production and import of fertilizer is also not competitive and therefore adds to the overall subsidy bill.

The Mahatma Gandhi National Rural Employment Guarantee Act scheme, which guarantees rural workers 100 days of minimum wage employment annually, needs careful evaluation to assess its benefits, costs, and sustainability. Critics claim that it distorts the labor market and creates labor shortages in public projects and in the agricultural surplus of states. Its proponents, meanwhile, argue that it benefits the poorest, especially women, and that it has raised the rural wage without raising rural productivity.

Learning from international experience could be useful as India reforms its social assistance programs. In the developed world, welfare programs are largely built around individuals, not products. As India moves to a more market-based economy, it must also reconsider its approach to subsidy programs. Most middle-income countries, such as Brazil, Mexico, the Philippines, and Turkey are moving toward individual-based transfer programs instead of product-based programs. This has taken place to ensure that the markets for these products can operate and develop effectively. With the aim of making recommendations to modernize India's social assistance programs, the Independent Evaluation Office will begin an evaluation of the public distribution system and assessments of other parts of the social assistance system will follow.

India has a large number of programs that aim to address inequity in the system—against women, castes, and tribes. How effective these have been in dealing with these inequities requires careful evaluation, as evidence indicates mixed progress in achieving objectives on this front. Several programs have been designed to promote greater equality. One example is the National Scheme of Incentives to Girls for Secondary Education, but whether this objective has been achieved still needs to be assessed. And other factors that might hold back school attendance by girls, such as lack of toilets and transport, must also be reviewed. The attendance of girls at higher levels of schooling and labor force participation are also interesting issues for the office's evaluation agenda. The male–female population ratio is starkly indicative of discrimination against girls, but statistics show the ratio of girls to boys falls as incomes rise across the country.

India is also repeatedly hit by natural disasters and there is concern that with climate change, the frequency and intensity of disasters will increase. The deadly floods in the state of Uttarakhand in 2013 also highlighted the link between development and disasters. India has developed over time disaster management capacity and institutional infrastructure, including the creation of the National Disaster Management Authority. The effectiveness of India's disaster management and response capability should be assessed. The extent to which India has built disaster management into its planning systems and the disaster proofing of infrastructure investments are other important areas for evaluation.

### **Regulation and economic management and public-private interface**

In the early 1990s, India liberalized its license control regime and opened its trade account, which attracted substantial increases in investment. For a while the economy grew rapidly, but lately India has suffered a decline in foreign direct investment. Moreover, Indian businesses have started investing outside the country. There is a growing sense that doing business in India is difficult. India ranks 134 out of 189 countries in the latest World Bank's Doing Business index, with its performance due to heavy bureaucracy and an unclear regulatory environment that has led to project delays and uncertainty. As a result, India affords huge opportunities for corruption and state capture by vested interests. This has lowered India's competitiveness, made the business environment very difficult, and allowed powerful

interests to prosper. Furthermore, transparency and competition have been eroded even though the share of the private sector in the economy has increased.

Civil society plays an active role in exposing corruption by trying to increase transparency and stop uncompetitive resource allocation in land and minerals through the Right to Information Act, pursuing court cases, and attempts to strengthen the Ombudsman function.

India has huge infrastructure needs, but does not have the public resources to meet them. Similar to other countries, India uses public–private partnership agreements to bring private capital and management expertise to help build roads, airports, ports and for the power sector. The country’s telecommunications sector has largely been built up with a public sector regulator and private operators. However, in other sectors, public–private partnership arrangements have yielded decidedly mixed results.

The issues discussed in this section offer opportunities for several areas of evaluation. These include (i) the evolution of regulation and regulatory bodies; (ii) the conflicting and multiple regulations that delay project implementation; (iii) the role of judicial activism, some of which is arising from lack of clarity in multiple and conflicting legislation and its confusing implementation; (iv) the lack of clarity on the role of central versus state rules and regulations, with implementing agencies and investors having to sort through the maze; and (v) the record of public–private partnership programs, particularly how India’s record compares with other countries, which sectors have done better, and whether some states have performed better than others.

India’s privatization program needs a boost and its focus needs a rethink. The privatization program is seen largely as a mechanism to raise fiscal revenues to meet fiscal targets. The potential benefits of privatization to improve productivity, infuse new technology, and improve competitiveness have not been sufficiently assessed. An evaluation of India’s disinvestment program and its benchmarking with other countries could be a useful exercise for a better understanding of what needs to be done and what the pitfalls are, and for identifying the interests of various stakeholders in India’s disinvestment program.



With this approach, we hope to open a virtual discussion with research institutes, development institutions, and civil society organizations on what should be the Independent Evaluation Office's work program. We hope this debate will not only open the space for discussion, but will become an open forum for ideas and solutions as well.

## Notes

<sup>1</sup> For a discussion of a new stage in the evolution of India's evaluation function, see Mehrotra (2013).

<sup>2</sup> The 10 largest programs by outlay for 2013–2014 are Mahatma Gandhi National Rural Employment Guarantee, Sarva Shiksha Abhiyan, Pradhan Mantri Gram Sadak Yojana, National Health Mission, Integrated Child Development Scheme, Indira Awas Yojana, Jawaharlal Nehru National Urban Renewal Mission, Mid-day Meal Scheme, National Rural Drinking Water Programme, and Rajiv Gandhi Grameen Vidyutikaran Yojana.



## Chapter 11

# Results-Based Evaluation in Malaysia

*Koshy Thomas*

*“However beautiful the strategy, you should occasionally  
look at the results.”*

*Winston Churchill*

**W**ithin Asia, Malaysia has played a leading role in adopting managing for development results through its participation in the Asia-Pacific Community of Practice on Managing for Development Results, a peer learning initiative promoted by the Asian Development Bank. Malaysia was one of the first countries in the region to adopt the practice using an integrated planning, budgeting, and monitoring and evaluation (M&E) model. But M&E infrastructure, which is crucial for results-based evaluation, for the most part did not exist in the Malaysian public sector until the implementation of outcome-based budgeting in 2011, which forms part of this chapter’s story.

Results-based management emphasizes the importance of systematic results setting and clear statements on how these should be achieved using a comprehensive results framework. In 2011, the Ministry of Finance shifted to an integrated results-based management system. This is made up of three primary components: integrated development planning, a results-based budgeting system, and a results-based personnel performance system. The system also has two complementary components: a results-based management information system and a results-based monitoring and evaluation framework. The primary components provide the necessary framework for planning, implementing, monitoring, and reporting on organizational performance and linking organizational performance to personnel performance. The complementary components provide the dynamics to the entire framework. They bring to life static information by establishing some level of causality. This is especially important for the National Budget Office in arriving at resource allocation decisions.

The system's integration was made possible with the use of a single integrated performance management framework for planning, budgeting, and M&E (Figure 11.1).<sup>1</sup> This framework basically requires top management within ministries and departments to be involved in strategic performance planning through consultation and consensus building with the lower levels. This is in line with the top-down and bottom-up approach of integrated results-based management.

The alignment of Malaysia's national priorities down to implementation levels starts with the 5-year planning cycle. This medium-term planning framework spells out the national priorities that need to be met and sets high-level objectives in the form of national outcomes. This, in turn, guides ministries in developing their own results framework and subsequently guides budget preparation with direct linkages to the results areas. During implementation, the same framework serves as the tool in monitoring performance from the bottom up, since implementation takes place at the activity level.

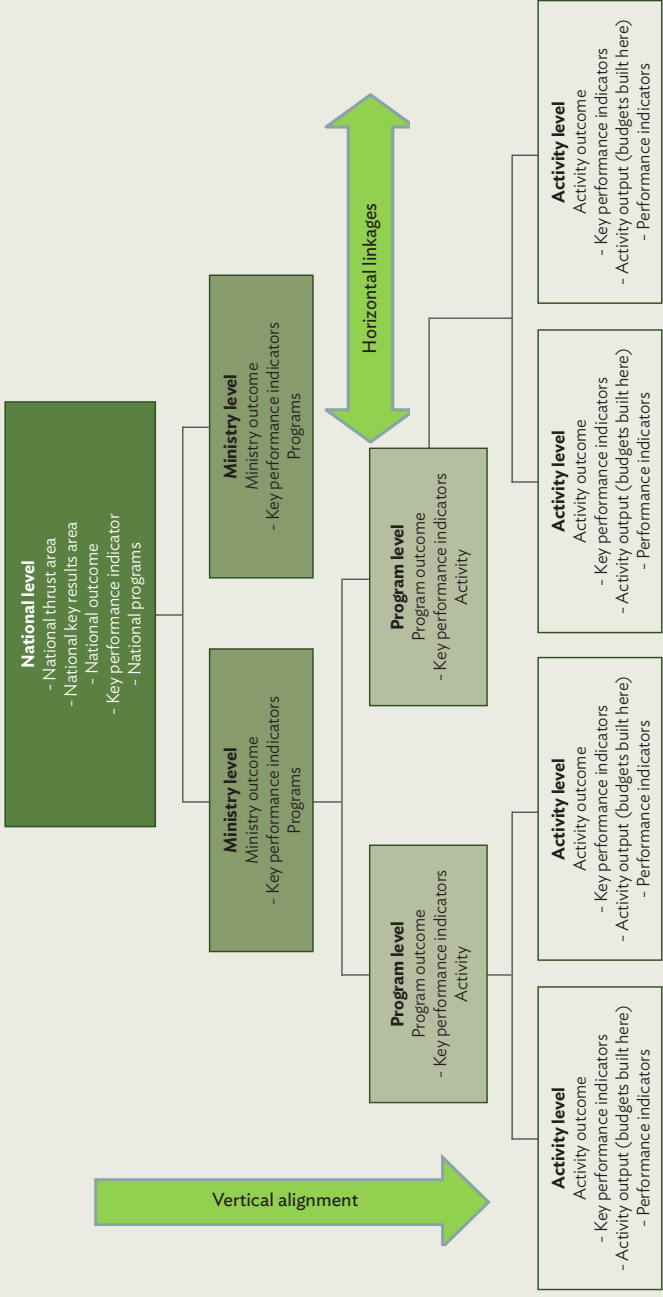
Under the outcome-based budgeting that began in 2011, the authority to manage resources is delegated to the activity level where outputs are generated under the concept of "let managers manage." Managers, in turn, are held accountable for results by way of a performance agreement that clearly sets out their obligations to deliver results based on the resources provided.<sup>2</sup> The performance agreement captures not just the results areas, but also the program theory incorporating the theory of change to determine why and what we are attempting to change. Assumptions and risks for each identified outcome are also reflected, along with risk mitigation measures.

Due to high data volume and the establishment of vertical and horizontal linkages to an outcome area, outcome-based budgeting is facilitated through an online application called MyResults. This is a government system that captures the performance of all ministries while linking to national priorities. The ministries are required to prepare their results framework and budget requirements in the application and submit it to the National Budget Office online.

### **Strategies in implementing the evaluation component**

Malaysia adopted performance budgeting back in 1970 with the introduction of the program performance budget system and

**Figure 11.1: Vertical and horizontal linkages in a program/activity environment**



Source: Outcome-based Budgeting Project Team Training Program. 2011. Ministry of Finance, Malaysia.

subsequently the modified budgeting system was implemented in 1990. However, limited success was achieved with both, despite the recognition that evaluation is critical to program improvement. Over this period, evaluations were done on an ad hoc basis and, in most cases, outsourced to third parties such as consultants and universities. Lack of data for decision making and establishing causality aggravated the problem. To rectify the situation, the government decided in 1999 that ministries should conduct program evaluations under the modified budgeting system. But a review of that system in 2010, aimed at making it compliant with managing for development results (OECD 2004), revealed that it had achieved only limited success. Conceptually, it promoted sector-based planning and program-based budgeting. But in practice its implementation had reverted to traditional budgeting focusing on input requirements at the institutional level rather than program performance. In 2011, the Ministry of Finance commissioned consultants to conduct an independent review of the modified budgeting system. The consultants concluded that the lack of an integrated M&E system was a key contributory factor to poor program implementation (KPMG 2011).

**Internalized self-evaluation model.** In developing a sustainable model for evaluation, a number of modalities were explored, which led to the adoption in 2010 of the internalized self-evaluation model in which program managers were directly involved in evaluation. Here, development evaluation became an ongoing learning process integral to the development cycle. Internal evaluation is evaluation conducted by a person or group directly associated with the organization or program that supports organizational development and intentional change. A number of developed countries have adopted the internalized self-evaluation model in part or holistically, including France, Japan, the Republic of Korea, the United Kingdom, and the United States. Traditionally, evaluation was perceived as something that could only be undertaken by external, highly trained consultants. Evaluations generally conducted in Malaysia tend to focus a great deal on methodology resulting in program managers shying away from evaluation. While it is true that methodology is important and critical in evaluation, the goal under the internalized self-evaluation model is to make it a simple management tool for managers at all levels to allow them to be involved directly. To this end, evaluation must be utilization focused (Patton 1997), and in the case of internalized self-evaluation, the focus is on stakeholders. In developing the model, the development team realized that better

buy-in and commitment to program modification and rectification resulted from direct involvement by management in the evaluation process. Moreover, performance information is critical for program managers in making decisions.

The internalized self-evaluation model was piloted in Malaysia from 2001 to 2006, and required program and activity managers to undertake evaluation internally. Under the pilot scheme, managers were trained in skills to manage and guide the evaluation. An experienced moderator guided the initial process to ensure adherence to the requirements that make the evaluation rigorous and valid. The strength of internalized self-evaluation lies in formats that have been extensively tested and used in a series of workshops to generate the following results:

- (i) The evaluation proposal is prepared through a standardized action plan.
- (ii) The program logic, outcomes, and key performance indicators are revisited and aligned for appropriateness and validity.
- (iii) Evaluation questions are generated after extensive consultation with stakeholders on four focus areas: appropriateness, effectiveness, efficiency, and economy.
- (iv) All significant data gaps in the existing system are identified for ratification for future evaluations.
- (v) The premise for a more effective monitoring system is identified and developed.
- (vi) Critical linkages that contribute to results are monitored.

Several reasons support the adoption of internalized self-evaluation other than the need to build internal capacity and empower program managers to make informed decisions. Among the most obvious ones are that stakeholders and managers have ownership of the evaluation process, evaluation is ingrained in the day-to-day operations of an organization as part of the developmental process, and there is

improved communication and understanding between groups within and between programs.

Over the pilot years, internalized self-evaluation was developed and tested locally and verified by local and international experts as an effective management tool. A study on its effectiveness as a capacity building tool jointly carried out by the Ministry of Finance with a local university showed encouraging findings.

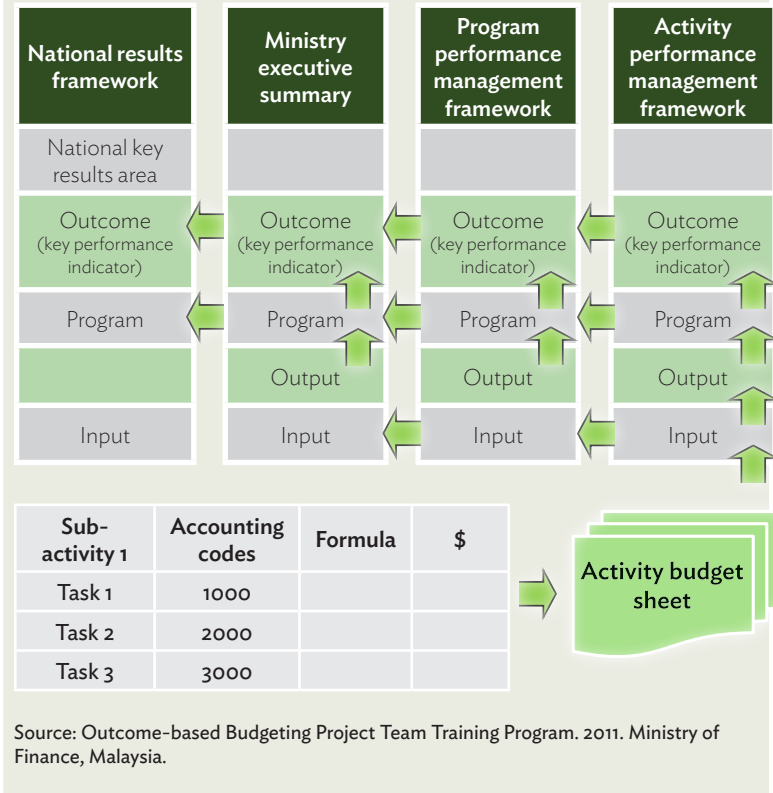
**Integrating the internalized self-evaluation model into outcome-based budgeting.** Within an organizational context, internalized self-evaluation is a tool for managers to guide organizational management and decision making that contributes to organizational learning. Therefore, it is essential to integrate M&E into organizational life. Under outcome-based budgeting, M&E is designed systematically as an integral part of a framework in which horizontal and vertical linkages are clearly defined (Figure 11.2). However, the success of internalized self-evaluation within the outcome-based budgeting framework depends, among other things, on top management support and commitment to planning and M&E, good data quality, and an organizational culture that supports continuous learning and critical program review for improved decision making.

The obvious question that comes out of implementing internalized self-evaluation relates to bias and nondisclosure of negative outcomes by affected program managers. In Malaysia's case, these were limited by conducting peer reviews through a predetermined institutional structure, conducting meta evaluations on evaluations driven by internalized self-evaluation, and verifying data quality through internal audits. In a worst-case scenario in which a program manager does not disclose critical findings, in all probability that manager will make program rectifications in the following year.

**Implementing the internalized self-evaluation model.** A pilot test was undertaken through a series of workshops guided by trained moderators, which propel the evaluation through a four-stage program logic evaluation process: preliminary planning and assessment, evaluation design, data collection and analysis, and reporting and information utilization (Figure 11.3).

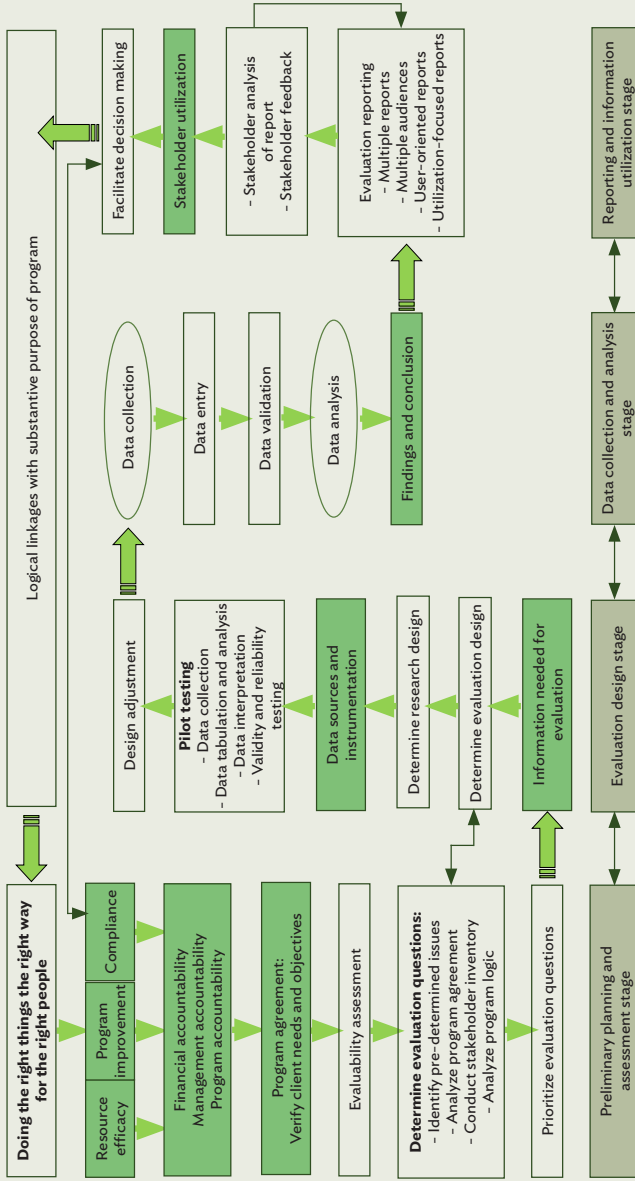


**Figure 11.2: Monitoring linkages in the strategic framework of outcome-based budgeting**



**Institutional arrangements to support M&E.** A national results framework linked to the ministry results framework has been implemented since 2013, which serves as the basis for identifying national and ministry level indicators of performance. This links broad national goals with key national strategies under the 10th Malaysia Plan, 2011–2015 and is used for measuring and reporting on the country's progress in achieving the plan's goals. It supports national planning and policy decision making, both for long-term planning and annual budgetary priorities. The national goals are cascaded into the sector priorities, which serve as the basis for identifying indicators of performance relevant to specific sectors and subsectors, ministries, and departments. The framework links the programs of government ministries with sector and national strategies.

Figure 11.3: Program logic and linkages evaluation flowchart



Source: A. Rassapan and J. Winston. 1999. *Evaluation Manual*. Ministry of Finance, Malaysia.

Without a system of ongoing data collection, it would not be possible to gather the right information from the right people at the right time. Such data should cover performance monitoring at the national, ministry, program, and activity levels for both operating and development expenditures. To this end, an Implementation Coordination Unit was set up under the Prime Minister's Office to monitor and coordinate all capital projects and report its progress to the Cabinet periodically. Since this unit is already involved in information collection specifically for capital projects, it is imperative that it supports the overall data collection strategy of the broader program that covers interventions funded by operating expenditure. That way, all data on interventions, whether funded by capital or operating budgets, are captured at the program level. The Ministry of Finance is a key player in this initiative by linking both capital expenditure and operating expenditure through the budgetary process to specific program outcomes.

Capacity building is an urgent requirement, whether it is for the interim measure of carrying out evaluation or for the longer-term initiative of setting up dedicated units in ministries. Capacity building also includes building high-level capacity among central agency officers, especially at the National Budget Office and the Economic Planning Unit, which can use the evaluative information for budgetary planning. A centrally located independent evaluation unit to serve as a policy reference point in support of developing and maintaining an effective and sustainable M&E system was set up in 2014. This unit is located in the National Budget Office and serves as the M&E "policy center" within government and works to establish evaluation policies and standards in the public sector, promote their implementation, and perform an oversight role in monitoring the use of evaluation across the public sector.

### **Evaluation challenges**

In piloting and implementing the internalized self-evaluation from 2001 to 2004, a number of challenges became evident:

- (i) Continuous capacity building is one of the biggest obstacles to ensuring sustainability when staff turnover is high. One urgent solution is to institutionalize knowledge and develop a decision support system internally so that staff turnover does not impact on the evaluation process.

- (ii) Evaluation is commonly perceived as an externally driven event although it is part of the program cycle. Traditionally, program managers have been involved only in program planning and implementation while evaluation was perceived as an externally driven initiative. The challenge of internalizing and institutionalizing to make evaluation part of the program manager's responsibility will require commitment from the highest level.
- (iii) Data quality is essential for ensuring effective evaluation. Understanding concepts in program logic, program theory, outputs, and outcomes and developing key performance indicators must be given priority—and program managers must be involved directly in ensuring data quality.
- (iv) Recommendations from evaluation findings have to be taken seriously and implemented systematically. If not, programs will not achieve their planned targets and results.

### **The way forward**

With the strategy for a comprehensive national monitoring system in place, short- and long-term initiatives to promote and implement evaluation using the internalized self-evaluation model are in the process of implementation. The adoption of an integrated performance management framework under outcome-based budgeting and the development of the MyResults application favor this integrated model. Systematic evaluation using internalized self-evaluation for program managers and complemented by external evaluation is scheduled to start in 2014, once data quality has improved.

Malaysia for a long time was a leader in public sector reforms, but has fallen short primarily due to lack of adherence to concepts and policy requirements. A framework is essential for developing strategies (Kaplan and Norton 2008), but tools are needed to translate strategies into operations. The framework is existent in outcome-based budgeting. Utilization must focus on performance to achieve objectives.

## Notes

<sup>1</sup> The integrated performance management framework is a strategic framework that links results from the national level down to implementation levels and embedded as a performance agreement at respective accountability levels.

<sup>2</sup> Performance agreements are signed at the ministry, program, and activity levels based on detailed performance targets that have been agreed. These include the Ministry Executive Summary signed by the controlling officer at the ministry level, the program performance management framework at the program level, and activity performance management framework at the activity level.



## Chapter 12

# Strengthening Evaluation Capacity in Asia and the Pacific

*Li Kouqing, Peng Runzhong, Zhao Min, and Wu Ningquin*

The People's Republic of China (PRC) established the Asia-Pacific Finance and Development Center (AFDC) in Shanghai in 2004 to strengthen institutional capacity building in finance and development and to promote communication and cooperation among economies in Asia and the Pacific. It is a public institution directly affiliated with the Ministry of Finance.

With increasing public attention on the performance of public expenditure and public management, the global consensus is that capacities for results-based management and evaluation need to be strengthened to improve development effectiveness. Many countries have been initiating a performance management framework to improve public service delivery and government efficiency. Weak capacity in both these areas is a bottleneck most developing countries in East Asia face. So it is essential to promote good practices in the performance evaluation of public funds in the region and to support research, knowledge sharing, and institutional and individual capacity development.

In the PRC, considerable attention has been given to performance evaluation since the 1990s. As former premier Wen Jiabao stated, "Performance evaluation, as the basis and prerequisite of administration accountability, can guide the efforts of the government and its staff to perform their tasks responsibly." Performance evaluation is considered an effective management tool for most government agencies. New evaluation campaigns focus on managing evaluations during the entire project cycle and emphasize the results of publicly funded projects. The PRC has also implemented monitoring and evaluation (M&E) pilots in almost all line ministries and at all levels of government. There has also been a notable shift to results-focused and performance-based budgeting and auditing systems. Since 2011, all central government-funded projects of over

5 million yuan (amounting to hundreds of billions yuan per fiscal year) have been monitored and evaluated by budgeting departments, evaluation professionals, or government officials. However, few Chinese are familiar with development M&E theories.

In 2012, the Ministry of Finance issued a plan for performance-based budgeting management (2012–2015), which aimed at building a budget management system in the PRC. For Asia and the Pacific, many developing countries consider performance-based management a very important tool in enhancing public management capacity.

### **Shanghai International Program for Development Evaluation Training**

The common challenge for almost all developing countries in the region, including the PRC, is the lack of evaluation and management capacity in the public sector. Results-based management is a relatively modern management concept and approach. To build capacity in performance evaluation, promote effective use of public funds, and learn from international experience in development evaluation, the PRC's Ministry of Finance, the World Bank, the Asian Development Bank (ADB), and AFDC jointly launched the Shanghai International Program for Development Evaluation Training (SHIPDET) in early 2007. SHIPDET is committed to promoting the development of evaluation capacity across the region through a range of activities, such as training workshops, research, and network building.

From the outset, ADB's Independent Evaluation Department (IED) has provided tremendous financial and human resources support to this collaboration. The participation of ADB staff and technical assistance consultants in the inaugural SHIPDET course contributed to the successful program launch. The experience also provided a very solid base from which to build a world-class program. IED has placed a lot of emphasis on developing evaluation capacity in the Greater Mekong Subregion and in Central Asia Regional Economic Cooperation countries. It has sponsored middle-level and senior officials to attend the SHIPDET courses, and when these officials return from their training, they often play a very important role in promoting and championing results-based management in their divisions or departments.



**Training offered to clients.** SHIPDET has grown rapidly since it was established. Initially, only two basic M&E training courses were offered—one for participants from the PRC and another for international participants. Participants mainly came from the finance and budgeting departments of government agencies. Over the years, SHIPDET has broadened its range of courses to include more advanced training for middle-level and senior government officials. In addition to the regular core courses, impact evaluation and results-based planning, budgeting, and evaluation are now part of the training. As most of the participants are from government, it is convenient for them to bring what they learned in class to their work and for them to promote management for development results there. SHIPDET not only conducts lectures, but also arranges visits to project sites with the help of IED.

Since 2008, the National Development and Reform Commission has listed SHIPDET's core courses in the training catalog for government officials from the PRC's underdeveloped western region. Here, internationally assisted and government-funded projects are being implemented to improve the welfare and standard of living of residents. It is crucial that local government officials learn about M&E and managing evaluations, particularly in cases in which a third-party agency is commissioned to conduct an evaluation. Indeed, as most of the government officials directly in charge of performance evaluations have now been trained at SHIPDET, the program's focus has shifted from finance bureaus to third-party intermediaries, mostly accounting firms that are hired by local governments to do the performance evaluation.

In recent years, several local provincial finance bureaus have commissioned AFDC to provide customized M&E training for their officials. Local governments have praised this initiative, which has been a meaningful start to promoting performance evaluation. AFDC's faculty also conducts research to apply international good practices to local situations and to provide training on international evaluation theories and PRC-specific evaluation approaches.

**Research and consultancy.** AFDC's research team focuses on results-based management, especially on evaluation theories, methodologies, case studies, guidelines, and policy recommendations. The purpose of the research is to solve the problem of matching international concepts to Asian and Chinese practices.

The center has also conducted dozens of research projects for various government agencies, including the Ministry of Finance, Ministry of Water Resources, and the Shanghai Urban Planning Bureau. This team has accumulated valuable experience in the process of providing consultancy services to projects of government agencies.

For consultancy and advisory services, AFDC teams have been invited to serve as external evaluation experts to the Ministry of Finance, Ministry of Agriculture, Ministry of Water Resources, Ministry of Education, Ministry of Health, Ministry of Environmental Protection, and several local governments to evaluate projects in finance, technology, education, health, poverty reduction, and infrastructure.

**Outputs.** SHIPDET has increasingly built its reputation at home and abroad as a reliable provider of training in performance evaluation. Under the leadership of the Ministry of Finance and with the support of its partners, it has conducted 38 training courses and seminars on performance evaluation for about 2,200 participants from 2007 to 2013. Almost half of them were international participants.

To meet the demand for evaluation capacity development in Central Asia, IED has conducted special topic courses on impact evaluations and performance budgeting under the SHIPDET framework. Middle-level and senior officials with basic knowledge of development evaluation are the target participants for these courses.

**Outcomes.** In 2010, a self-evaluation of the SHIPDET program indicated that most participants found the training useful to their work. As one participant pointed out: “SHIPDET opened the window to greater understanding of how development activities need to be managed. Results-based management and its process are key to successful development work. SHIPDET...helped me to incorporate these principles into our national development framework.”

In recognition of SHIPDET’s accomplishments, the World Bank’s Independent Evaluation Group designated AFDC in 2011 as the regional Center for Learning on Evaluation and Results (CLEAR) in East Asia. This opens a window of opportunity for more collaboration and communication with other partners in the region to further promote performance evaluation and results-based management in the public sector.

Partnerships established during the operation of SHIPDET have been valuable in promoting results-based management in Asia and the Pacific. SHIPDET has worked closely with international partners such as ADB, which has a powerful knowledge hub in the region. ADB has assisted SHIPDET in approaching foreign governments through its resident missions to screen and sponsor the most suitable officials to come to the PRC for training.

**Impact.** With the acceleration of performance-based management in the region in recent years, the demand for evaluation capacity building has grown rapidly. Many CLEAR-SHIPDET alumni have become important evaluation champions and practitioners in their own countries and in the region. In the PRC, for example, trainees have provided invaluable intellectual and organizational support to internationally-funded pilot projects.

### **Vision of the Center for Learning on Evaluation and Results in East Asia**

Since its launch in 2011, CLEAR-SHIPDET has aimed to be the region's evaluation capacity development center. To meet the demand for improved evaluation capacities, more international resources to conduct training programs, seminars, and workshops will be mobilized. In addition to government officials and evaluation practitioners from the PRC and other Asian developing countries, target clients include Chinese postgraduate students as they will be the future practitioners and drivers of M&E. In 2011, AFDC launched a diploma course on evaluation in the auditing master program of the Shanghai National Accounting Institute. Because of a severe shortage of evaluation professionals, the PRC considers accountants as the most important evaluation force. The institute is the government's training center for accountants, of which there are around 20 million in the country. AFDC has also started to work with Shanghai National Accounting Institute to train certified accountants through face-to-face training or video conferencing.

CLEAR-SHIPDET also aims to be a self-sustaining evaluation knowledge supplier. It has carried out considerable research on M&E under the leadership of the Ministry of Finance and AFDC. In March 2014, the ministry commissioned AFDC to do comparison research on the M&E systems of international finance institutions and the PRC's fiscal programs. As a result, AFDC will tender in the coming years

qualified think tanks to do at least one special research project on the application of international evaluation experience. CLEAR-SHIPDET staff members, however, will remain the core faculty, especially for lectures targeted at Chinese participants. CLEAR-SHIPDET's own professional staff will conduct almost 50% of lectures, consultant services, and professional guidance.

Furthermore, CLEAR-SHIPDET aims to be one of the best network centers for evaluation practitioners. With different kinds of practices, CLEAR-SHIPDET is on the way to becoming the M&E practice knowledge supplier for both the PRC and the region. In this sense, the continued support of think tanks of the World Bank, ADB, and other international institutions is vital. Evaluation will be promoted through website upgrading, case study seminars, and newsletter distribution with the aim that more people can share evaluation practices and experiences through this platform.

**Challenges ahead.** CLEAR's global program has provided strong technical support to CLEAR-SHIPDET and a small amount of financial support. From the perspective of sustainable financial resources, training programs face the risk of cutbacks unless there is sufficient support. How to generate sustainable revenues through training fees is one of the major challenges ahead for CLEAR-SHIPDET. Capacity development is needed for domestic trainers, including center staff, to reduce reliance on foreign experts. How to develop a domestic or regional network of qualified trainers in the future needs to be addressed. M&E training programs are available in the PRC and elsewhere in East Asia, although these are offered by different organizations. So building stronger marketing and branding of CLEAR in the region will be essential for the center's development, and part of this process will need to entail establishing further cooperation and collaboration with other regional M&E centers. It is our hope that IED, as a regional knowledge hub of performance evaluation, will continue to work with the Ministry of Finance and the World Bank to further promote results-based management in the public sector for Asia and the Pacific through innovations in the format, content, and outcomes of capacity development activities.

## References

- Asian Development Bank (ADB). 2000a. *Rising to the Challenge in Asia: A Study of Financial Markets*. (Vol. 1). Manila.
- . 2000b. *Seventh Replenishment of the Asian Development Fund—ADF VIII Donors’ Report*. Manila.
- . 2003a. *ADF VIII Midterm Review: Chairman’s Summary of Discussions*. Manila.
- . 2003b. *Enhancing the Independence and Effectiveness of the Operations Evaluation Department*. Manila.
- . 2003c. *Enhancing the Independence and Effectiveness of the Operations Evaluation Department (Preliminary Draft for Consultation)*.
- . 2008a. *Review of the Independence and Effectiveness of the Operations Evaluation Department*. Manila.
- . 2008b. *Strategy 2020: The Long-Term Strategic Framework of the Asian Development Bank 2008–2020*. Manila.
- . 2011a. *Proposed Grant Assistance to Sri Lanka: Improving Community-Based Rural Water and Sanitation in Postconflict Areas of Jaffna and Kilinochchi*. Manila.
- . 2011b. *Water Operational Plan: 2011–2020*. Manila.
- . 2013a. *Gender Equality and Women’s Empowerment Operational Plan, 2013–2020*. Manila.
- . 2013b. *Operational Plan for Enhancing ADB’s Effectiveness in Fragile and Conflict-Affected Situations*. Manila.
- . 2013c. *Proposed Loan and Administration of Grant to the Lao People’s Democratic Republic: Water Supply and Sanitation Sector Project*. Manila.

- Astbury, B., and F. L. Leeuw. 2010. Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation* 31(3): 363–381.
- Bamberger, M. 2012. *Unanticipated Consequences of Development Interventions: A Blind Spot for Evaluation Theory and Practice*. Guest lecture at the International Program for Development Evaluation Training (IPDET). 2 July 2012 (video). <http://ipdet.org/page.aspx?pagelD=videoMichaelBamberger2012>
- . 2013. *Engendering M&E. Nuts and Bolts of Monitoring and Evaluation Systems*. Series No. 27. World Bank.
- Donaldson, S. I., M. Q. Patton, D. M. Fetterman, and M. Scriven. 2009. Claremont Debates: The Promise and Pitfalls of Utilization-Focused and Empowerment Evaluation. *Journal of Multidisciplinary Evaluation* 6(13): 15–57.
- Duflo, E. 2011. *Balancing Growth with Equity: The View from Development*. Paper presented at Jackson Hole 2011 Symposium. 1 August.
- Funnell, S., and P. Rogers. 2011. *Purposeful Program Theory*. San Francisco, CA: Jossey-Bass.
- Guijt, I. 2010. Exploding the Myth of Incompatibility between Accountability and Learning. In J. Ubels, N-A Acquaye-Baddoo, and A. Fowler, eds. *Capacity Development in Practice*. London: Earthscan.
- Henry, G. and M. Melvin. 2003. Beyond Use: Understanding Influence on Attitudes and Actions. *American Journal of Evaluation* 24(3): 293–314.
- Hesse-Biber, S. (ed). 2012. *The Handbook of Feminist Research: Theory and Practice*. Thousand Oaks, CA: Sage Publications.
- Imas, L. G. M., and R. C. Rist. 2009. *The Road to Results: Designing and Conducting Effective Development*. Washington, DC: World Bank.

- Independent Evaluation Department (IED). 2009. *Impact Evaluation Study: Impact of Rural Water Supply and Sanitation in Punjab, Pakistan*. Manila: ADB.
- 2010a. *Learning Curves: Reducing Carbon Emissions from Transport Projects*. Manila: ADB.
- 2010b. *Special Evaluation Study on ADB's Support to Fragile and Conflict-Affected Situations*. Manila: ADB.
- 2011. *Special Evaluation Study on the Asian Development Fund Operations: A Decade of Supporting Poverty Reduction in the Asia and Pacific Region*. Manila: ADB.
- 2012a. *Annual Evaluation Review*. Manila: ADB.
- 2012b. *Real Time Evaluation of the Multitranche Financing Facility*. Manila: ADB.
- 2012c. *Special Evaluation Study on ADB's Knowledge Products and Services*. Manila: ADB.
- 2012d. *Special Evaluation Study on ADB's Response to Natural Disasters*. Manila: ADB.
- 2012e. *Special Evaluation Study on ADB's Social Protection Strategy*. Manila: ADB.
- 2013a. *Annual Evaluation Review*. Manila: ADB
- 2013b. *Learning Lessons: Making Infrastructure Disaster-Resilient*. Manila: ADB.
- 2014. *Inclusion, Resilience, Change: ADB's Strategy 2020 at Mid-Term*. Manila: ADB.
- Independent Evaluation Group (IEG). 2010. *Safeguards and Sustainability Policies: An Independent Evaluation of the World Bank Group*. Washington, DC: World Bank.

- \_. 2011. *Country Level Engagement on Governance and Anticorruption: An Evaluation of the 2007 Strategy and Implementation Plan*. Washington, DC: World Bank.
- \_. 2012. *The Matrix System at Work: An Evaluation of the World Bank's Organizational Effectiveness*. Washington, DC: World Bank.
- \_. 2014. *World Bank Group Assistance to Low-Income Fragile and Conflict-Affected States: An Independent Evaluation*. Washington, DC: World Bank.
- Jones, L., and I. Sakong. 1980. *Government, Business, and Entrepreneurship in Economic Development: The Korean Case*. Cambridge, MA: Harvard University Press.
- Kaplan, R., and D. Norton. 2008. Mastering the Management System. *Harvard Business Review*. January 2008.
- Karkara, N. 2013. *Advocating for Evaluation: A Toolkit to Develop Advocacy Strategies to Strengthen an Enabling Environment for Evaluation*. EvalPartners in cooperation with United Nations Women and International Organization for Cooperation in Evaluation.
- Kirkhart, K. E. 2000. Reconceptualizing Evaluation Use: An Integrated Theory of Influence. *New Directions for Evaluation* 88: 5–23.
- KPMG. 2011. Review of the Modified Budgeting System Framework. Inception Report submitted to the Ministry of Finance of Malaysia.
- Kusek, J., and R. Rist. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System*. Washington, DC: World Bank.
- Lin, J. Y., and Y. Wang. 2012. [The People's Republic of] China's Integration with the World: Development as a Process of Learning and Industrial Upgrading. *China Economic Policy Review* 01(1): 1–33.
- Manuel, T. 2013. *Independent Panel: Reviewing the World Bank Group's Doing Business Report*. Washington, DC: World Bank.



- Mark, M., and G. Henry. 2004. The Mechanisms and Outcomes of Evaluation Influence. *Evaluation* 10(1): 35-57. London: Sage Publications.
- Mehrotra, S. 2013. The Government Monitoring and Evaluation System in India: A Work in Progress. *Evaluation Capacity Development Working Paper Series*. No. 28. Washington, DC: World Bank.
- Mertens, D. M., and S. Hesse-Biber. 2013. Mixed Methods and Credibility of Evidence in Evaluation. *New Directions for Evaluation* 138 (Summer): 5-13.
- Morabito, S. 2002. Evaluator Roles and Strategies for Expanding Process Influence. *American Journal of Evaluation* 23(3): 321-330.
- Operations Evaluation Department. 2007. *Independent Evaluation at the Asian Development Bank*. Manila: ADB.
- Organisation for Economic Co-operation and Development-Development Assistance Committee. 2002. *Evaluation Glossary*. Paris.
- \_\_\_\_\_. 2004. *Policy Brief on Managing for Development Results*. Marrakesh Roundtable on Results. Paris.
- \_\_\_\_\_. n.d. *The Paris Declaration on Aid Effectiveness (2005) and the Accra Agenda for Action (2008)*. Paris.
- Overholt, C., M. Anderson, K. Cloud, and J. Austin. 1985. *Gender Roles in Development Planning: A Case Book*. West Hartford, CT: Kumarian Press.
- Patton, M. Q. 1997. *Utilization-Focused Evaluation*. The New Century Text. 3rd ed. Thousand Oaks, CA: Sage Publications.
- \_\_\_\_\_. 2011. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York and London: The Guilford Press.

- Pawson, R. 2006. *Evidence-Based Policy: A Realist Perspective*. London: Sage Publications.
- Petri, P. A. 1988. [Republic of] Korea's Export Niche: Origin and Prospects. *World Development* 16(1): 47–63.
- Picciotto, R. 2012. Experimentalism and Development Evaluation: Will the Bubble Burst? *Evaluation*: 1–17.
- Planning Commission of India. 2013. Independent Evaluation Office. Press Information Bureau.
- Rao, A., M. Anderson, and C. Overholt. 1995. *Gender Analysis in Development Planning*. West Hartford, CT: Kumarian Press.
- Sackett, D. L., W. M. Rosenberg, J. A. Gray, R. B. Haynes, and W. S. Richardson. 1996. Evidence-Based Medicine: What It Is and What It Isn't. *British Medical Journal* 312 (7023): 71–72.
- Thomas, V., and X. Luo. 2012. *Multilateral Banks and the Development Process: Vital Links in the Results Chain*. New Brunswick, NJ: Transaction Publishers.
- United Nations Development Programme. 2001. *Development Effectiveness Review of Evaluative Evidence*. New York.
- Williams, S., J. Seed, and A. Mwau. 1994. *The Oxfam Gender Training Manual*. United Kingdom and Ireland: Oxfam.
- World Bank. 2012. *World Development Report on Gender Equality and Development*. Washington, DC.
- World Bank, Food and Agriculture Organization, and International Fund for Agricultural Development. 2009. *Gender in Agriculture Sourcebook*. Module 16: Gender Issues in Monitoring and Evaluation. Washington, DC.

*The book also draws on the background work and discussions with N. Fostvedt, B. Murray, P. Petri, R. Picciotto, H.S. Rao, and V. Thomas.*

## About the Contributors

**Michael Bamberger** worked for 9 years with the Gender and Development Department of the World Bank and, since 2001, as an independent consultant specializing in development and gender evaluation. He has a PhD in sociology from the London School of Economics. He has assisted in the development or evaluation of gender policies and strategies for the United Nations Development Programme, World Food Programme, UN Women, the African Development Bank, the World Bank, AusAid, and EvalPartners. Recent publications include *RealWorld Evaluation: Working under Budget, Time, Data and Political Constraints* (2012, with Jim Rugh and Linda Mabry); *Engendering M&E* (2013); and *How to Design and Manage Equity-Focused Evaluations* (2011, with Marco Segone).

**Ajay Chhibber** is the first director general of Independent Evaluation in India, with the status of minister of state (appointed in August 2013). Prior to this, he was based in New York as assistant secretary-general and assistant administrator at the United Nations Development Programme and head of its Asia-Pacific Bureau. Previously, he worked at the World Bank for almost 25 years on a range of development issues, managing its programs in Indonesia, the Pacific, Turkey, and Viet Nam. He was the lead author of the seminal work on governance at the World Bank, the 1997 *World Development Report on the Role of the State*. His publications include five books on economic development and numerous articles in international journals and major newspapers. He has a PhD from Stanford University and a master's degree from the Delhi School of Economics. He has worked at the Indian Planning Commission and at the International Food Policy Research Institute. He also taught economics at Georgetown University and the University of Delhi.

**Luigi Cuna** worked for a number of years as an independent consultant in monitoring and evaluation of agricultural and rural development programs. Afterward, he worked for 5 years at the International Fund for Agricultural Development, where he contributed to strengthening the self-evaluation system. Subsequently, he joined its Independent Office of Evaluation. Since 2011, he has been an evaluator at the

Evaluation Department of the Council of Europe Development Bank. He has a PhD from the University of Bologna.

**Anis Dani** worked at the World Bank from 1995 to 2014 as lead social scientist, social policy adviser, operations adviser, and lead evaluator, Country and Corporate Evaluations at the Independent Evaluation Group. He has a PhD in anthropology from the University of Pennsylvania. He had worked for the Aga Khan Foundation, United Nations Development Programme, German Technical Cooperation, Swiss Agency for Development and Cooperation, Novib, and the Asia Foundation. At the Independent Evaluation Group, he led evaluations of the World Bank Group's Safeguards and Sustainability Policies, the Matrix System evaluating the World Bank's organizational effectiveness, the Afghanistan country program, and Bank Group support to Low-income Fragile and Conflict-affected States. Besides operational and evaluation work, he coordinated the World Bank's work on social analysis and published books on *Poverty and Social Impact Analysis of Reforms: Lessons and Examples from Implementation* (2006); *Institutional Pathways to Equity: Addressing Inequality Traps* (2007); *Inclusive States: Social Policy and Structural Inequalities* (2008); and *Assets, Livelihoods, and Social Policy* (2008).

**Bastiaan de Laat** is team leader/evaluation expert at the European Investment Bank (EIB). He has longstanding experience in evaluation. Founder-director of the French subsidiary of the Technopolis Group (1998–2006), he led many evaluations for and provided policy advice to various local, national, and international public bodies. He trained several hundreds of European Commission staff and national government officials in evaluation and was involved in the design of monitoring and evaluation systems for various public organizations. He has a PhD from the University of Amsterdam. Before joining EIB in 2009, he worked as evaluator at the Council of Europe Development Bank. Independently of his position at EIB, he has been secretary general of the European Evaluation Society since 2011 and has co-organized the society's conferences since 2010. He has published on a wide variety of topics, most recently on enhancing evaluation use.

**Cheryl W. Gray** is director of the Office of Evaluation and Oversight at the Inter-American Development Bank. She joined in June 2011 after working 25 years at the World Bank, most recently as director of Independent Evaluation. Prior to that position, she served as a director in the World Bank's Europe and Central Asia Vice-Presidency,

managing the World Bank's work on economic policy, governance, judicial reform, gender, and poverty analysis in 30 countries. From 1997 to 2002, she served as director for public sector reform, co-authoring the World Bank's anticorruption and public sector reform strategies. Before joining the World Bank, she was an advisor in the Ministry of Finance in Indonesia for 3 years. She has a PhD and law degree from Harvard University and an undergraduate degree in economics from Stanford University.

**Samer Hachem** has served since 2013 as manager of Independent Development Evaluation at the African Development Bank (AfDB). Before this, he worked at AfDB's Quality Assurance and Results Department, monitoring AfDB's performance on aid effectiveness and contributing to the introduction of results tools such as development effectiveness reviews and a new results management system. He has more than 15 years of experience in monitoring and evaluation. He started his career in management consulting, focusing on the financial services sector and leading monitoring and evaluation in large-scale change management programs. He also worked at the United Nations Development Programme, where he helped roll out new results-focused management policies and systems. As a consultant in West African countries, he evaluated capacity in managing for development for results and developing improvement plans. He has degrees in business administration and engineering from HEC Montreal, Canada, and École Centrale Paris, France.

**Caroline Heider** joined the Independent Evaluation Group of the World Bank Group in October 2011 as director general and senior vice president. She reports directly to the World Bank Group's board of executive directors through the Committee on Development Effectiveness. She has more than 25 years of international experience in evaluation. Before joining the Independent Evaluation Group, she headed the Office of Evaluation at the World Food Programme. She has extensive evaluation experience in international development and humanitarian assistance and experience in related areas of strategic planning, management audit, and program design. She has worked with six multilateral organizations. She is a lifetime member of the International Development Evaluation Association and a member of the American Evaluation Association. She is currently chair of the Global Evaluation Advisory Committee, UN Women. She has been a member of the Australasian Evaluation Society and served a 2-year term as vice-chair of the United Nations Evaluation Group.

**Hemamala Hettige** is senior advisor of Independent Evaluation at the Asian Development Bank (ADB). Drawing on her 17 years of evaluation experience, she focuses on supporting the director general of Independent Evaluation in his long-term vision and acts as primary focal point for communications with the Board, the Development Effectiveness Committee, Management, and the evaluation community. From 2008 to 2013, she was the director, Independent Evaluation Division 2. She joined ADB in 1997 and has led a variety of corporate, country, thematic, and project evaluations. She has a PhD from Boston University and a bachelor's degree in economics from the University of Sri Lanka. After her studies, she worked for 9 years at the World Bank (Washington, DC) on operational matters and research related to financial sector and environmental issues.

**Haruhiko Kuroda** is the governor of the Bank of Japan. He was the president of the Asian Development Bank (ADB) and the chair of ADB's Board of Directors from 2004 to 2013. Before joining ADB, he was special advisor to the Cabinet of Japanese Prime Minister Junichiro Koizumi and a professor at the Graduate School of Economics at Hitotsubashi University in Tokyo. He holds a bachelor's degree in law from the University of Tokyo and a master of philosophy in economics from the University of Oxford. He has represented Japan's Ministry of Finance at a number of international monetary conferences as vice minister of finance for international affairs.

**Kouqing Li** is the president of the Shanghai National Accounting Institute (SNAI). He is also the director-general of the Asia-Pacific Finance and Development Center (AFDC), which promotes evaluation capacity building for the international community through training workshops, academic research, and network building. Prior to this, he became vice president of SNAI in 2003, and executive deputy director-general of AFDC in 2007. He started teaching in 1988 at the Department of Economics and Trade of the Shanghai University of Finance and Economics and earned his professorship in 2000. He received his PhD in business administration and a master's degree in economics from this university in 2001 and 1988, respectively. Among his numerous publications are *Customer Value Advantage Theory and Economic History of New China*. In 2012, Li translated the *Roads to Results* into Chinese (with the AFDC team)—the first book on development evaluation in the People's Republic of China.

**Rachel Meghir** was appointed director of the Evaluation Department of the Council of Europe Development Bank in 2009. She joined the bank in 1992 as country manager, gaining operational and field experience in project design, appraisal, and supervision. Prior to this, she worked for 5 years at the Development Centre of the Organisation for Economic Co-operation and Development, where she was in charge of the Research Programme “Financing of Development.” She holds a graduate degree in development economics and a master’s degree in sociology from the University of Paris. She contributed to international workshops and conferences, published several articles, and co-authored the general synthesis of the case studies (*Financial Systems and Development: What Role for the Formal and Informal Financial Sectors?*, OECD Development Centre Studies, Paris 1991).

**Bruce Murray** was the first director general after the evaluation function at the Asian Development Bank (ADB) became independent in 2004. Prior to this, he had a long career in ADB operations, including as ADB’s first resident representative in the People’s Republic of China. After ADB, he chaired independent panels that reviewed the independent and self-evaluation systems for the boards of the International Fund for Agricultural Development and the Inter-American Development Bank, undertook a self-evaluation of the World Bank Group’s Independent Evaluation Group, and engaged in evaluation consulting assignments for the European Bank for Reconstruction and Development, African Development Bank, International Monetary Fund, and the Small Country Financial Management Centre. He peer reviewed various evaluations and commented on draft papers for the Transport Research Bureau. He has also undertaken volunteer evaluations of anticorruption projects in the Philippines for the Partnership for Transparency Fund and has given sessions on the People’s Republic of China and evaluation at the Asian Institute of Management and Texas A&M University. He has a master of arts in economics and a master of science in statistics from the University of Toronto.

**Ashwani Muthoo** is a professional evaluator and currently deputy director of the Independent Office of Evaluation of the International Fund for Agricultural Development (IFAD). He is responsible for managing numerous staff and consultants, as well as for conducting corporate level evaluations that contribute to organizational change

and improved development effectiveness. Examples of these evaluations relate to the institutional and operational efficiency of IFAD, agriculture in Africa jointly with the African Development Bank, the field presence pilot program, gender, private sector engagement, and innovation and scaling up. He is the principal author of the evaluation manual of the Independent Office of Evaluation and is responsible for methodology development and oversight. He also prepares the *Annual Report on Results and Impact of IFAD Operations*, the annual flagship report of the evaluation office.

**Takehiko Nakao** is the President of the Asian Development Bank (ADB) and the chair of ADB's Board of Directors. He assumed office in April 2013. Before joining ADB, he was the vice minister of finance for international affairs at the Ministry of Finance of Japan. In a career spanning more than 3 decades, he has gained extensive experience in international finance and development. He holds a master's degree in business administration from the University of California, Berkeley. He joined Japan's Ministry of Finance in 1978. From 1994 to 1997, he served as economist and advisor at the International Monetary Fund. He was assigned as minister at the Embassy of Japan in Washington, DC between 2005 and 2007. In 2010 and 2011, he was a visiting professor at the University of Tokyo. He has published books and numerous papers on financial and economic issues.

**Rakesh Nangia** is evaluator general and vice president of the African Development Bank (AfDB). Since his appointment in January 2012 as evaluator general, he has led the formulation and implementation of the independent evaluation strategy, which seeks to strengthen the evaluation function at AfDB, enhance outreach to stakeholders, and promote the use of evaluation to inform policy development and programming. He has degrees in business administration and engineering from the Indian Institute of Technology in Delhi and Harvard University. He has more than 25 years of experience in development work. Prior to joining AfDB, he was the director of Strategy and Operations for the Human Development Network at the World Bank. At the World Bank, he served in many different capacities, including acting vice president for the World Bank Institute (2007–2009) and manager, Portfolio and Country Operations in Viet Nam (2001–2006). Before moving to Viet Nam, he was based in Tanzania (1998–2001).



**Runzhong Peng** is director and professor of the Asia-Pacific Finance and Development Center under the Ministry of Finance of the People's Republic of China. He graduated from the Institute of World Economy of the Shanghai-based Fudan University and holds a PhD in economics. He also has a bachelor's degree in chemistry from Changchun Institute of Geosciences. Since 2002, his research area has covered Asia-Pacific Economic Cooperation economies, focusing on finance and regional economic cooperation and project evaluation. He also worked as senior expert for the Ministry of Finance in performance evaluation of agriculture and infrastructure projects funded by international financial organizations in the country. He is either an author or co-author of more than 20 papers and books, in English and Chinese.

**Peter Petri** is the Carl J. Shapiro professor of international finance at the Brandeis International Business School, a senior fellow of the East-West Center, and a visiting fellow at the Peterson Institute for International Economics. He was founding dean of the International Business School and has held visiting appointments at Fudan University, Keio University, Peking University, Asian Development Bank Institute, the Organisation for Economic Co-operation and Development, and the World Bank. He received AB and PhD degrees from Harvard University. His recent publications include *The Trans-Pacific Partnership and Asia-Pacific Integration: A Quantitative Assessment* (2012, with Michael Plummer and Fan Zhai), *ASEAN Centrality and the ASEAN-US Economic Relationship* (2013, with Michael Plummer), and *Navigating a Changing World Economy: ASEAN, the People's Republic of China, and India* (2013, with Fan Zhai).

**Robert Picciotto**, former vice president, Corporate Planning and Budgeting and former director general of the Independent Evaluation Group at the World Bank, is currently visiting professor at the International Development Institute (King's College London). He is a member of the Academy of Social Sciences. A former member of the United Kingdom Independent Advisory Committee on Development Impact (2006–2010), he currently serves on the boards of the United Kingdom Evaluation Society and the European Evaluation Society. He is a graduate of the Woodrow Wilson School (Princeton University) and of the Ecole Nationale Supérieure de l'Aéronautique (France). Among his numerous publications are *Global Development and Human Security* (2010, with Funmi Olonisakin and Michael Clarke) and *Impact of Rich Countries' Policies on Poor Countries* (2004, with Rachel Weaving).

**Halady Satish Rao** started his career in 1970 as an industrial engineer in Associated Cement Companies Limited. He has a degree in engineering from the Indian Institute of Technology Kharagpur and an MBA from the Indian Institute of Management Calcutta. After working for 4 years in various cement plants and doing his MBA, he joined the Planning Commission of India as deputy director of the Project Appraisal Division. In 1977, he moved to the Industrial Credit and Investment Corporation of India, then a development bank engaged in long-term financing of Indian private industry. After 7 years in project financing and merchant banking in this corporation, he joined the Asian Development Bank (ADB) in 1984. He retired from ADB in 2011 as director general of the Independent Evaluation Department. Postretirement, he consulted with the African Development Bank and is currently a visiting professor at MBA schools in India teaching infrastructure project finance.

**Moisés Schwartz** is director of the Independent Evaluation Office of the International Monetary Fund. Before this, he was president of the National Commission for Retirement Savings in Mexico. He earned a bachelor's degree in economics from the Tecnológico Autónomo de México and a PhD in economics from the University of California, Los Angeles. Between 2004 and 2006, he served as an executive director of the International Monetary Fund, representing Mexico and seven other countries. Prior to that, he held several senior positions in Mexico's public administration, including as the finance minister's chief of staff and director general of International Financial Affairs within the Ministry of Finance. He has also served as director of Macroeconomic Analysis and director of Economic Studies in Mexico's Central Bank.

**Koshy Thomas** is currently deputy director of the National Budget Office. He is the head of the Outcome-Based Budgeting Project Team under the Ministry of Finance, which is responsible for designing, planning, and implementing the outcome-based budgeting system in the Malaysian public sector. He has been trained in economics and human services in Malaysia and Australia. In the past 28 years, his work at the finance ministry covered finance, fiscal policy, and public investments. He has extensive international experience working collaboratively with the Asian Development Bank, United Nations Development Programme, and the United Nations Children's Fund in East Asia, Africa, and Central Asia on scoping missions, training, and advisory assistance. He is currently the chair of the Coordinating

Committee of the Community of Practice on Managing for Development Results for Asia and the Pacific, and the vice president of the Malaysia Evaluation Society.

**Vinod Thomas** is director general of independent evaluation at the Asian Development Bank (ADB). Reporting to ADB's Board of Directors, he leads the team that assesses ADB's development effectiveness and provides lessons to inform actions. He is responsible for evaluating policies, programs, projects, and technical assistance as well as thematic issues at the regional, country, and sector levels. Prior to this, he was the director general and senior vice-president of the Independent Evaluation Group at the World Bank. He has a PhD in economics from the University of Chicago. Among his numerous publications are *Multilateral Banks and the Development Process: Vital Links in the Results Chain* (2012, with Xubei Luo), *Development Imperatives for the Asian Century* (2013, with Peter Petri), and *Climate-Related Disasters in Asia and the Pacific* (2013, with Jose Ramon Albert and Rosa Perez).

**Jan Willem van der Kaaij** is inspector general of the European Investment Bank. He became a member of the board of directors of the bank in 2002 and inspector general in 2007. He has 25 years of professional experience that include 7 years in Washington, DC as executive director of the World Bank; executive director, Inter-American Investment Corporation; financial counsellor at the Royal Netherlands Embassy; and advisor to the executive director at the International Monetary Fund. Previous posts include deputy director, Foreign Financial Relations Directorate at the Ministry of Finance, The Hague; member of the Administrative Council for the Netherlands of the Council of Europe Development Bank, Paris; head of Multilateral Banks Division of the Ministry of Finance, The Hague; and economist at De Nederlandsche Bank (central bank of the Netherlands), Amsterdam. He obtained his master's degree in history from the University of Groningen in 1985, specializing in contemporary and economic history. He has published articles on the International Monetary Fund and De Nederlandsche Bank.

**Zhongjing Wang** is executive director for the People's Republic of China at the Asian Development Bank (ADB) and chair of the Development Effectiveness Committee since October 2012. Under his leadership, the committee assists the Board of Directors to carry out its responsibility of ensuring that ADB's programs and activities achieve development effectiveness. He has a master's

degree in public administration and a certificate of international tax program from Harvard University. He obtained his bachelor of law in international politics from Peking University. Prior to joining ADB's Board of Directors, he was deputy director general at the International Department of the Ministry of Finance and worked in the State Administration Taxation of the People's Republic of China.

**Horst Wattenbach** is senior evaluator at the Council of Europe Development Bank. He has worked for almost 25 years in development cooperation throughout the project cycle for development banks such as KfW and Council of Europe Development Bank, for the Food and Agriculture Organization, and for consultancy agencies. He is an agricultural economist from the Technical University Berlin. His secondary specialization is in environmental and resource economics (Harvard Institute for International Development and Göttingen University). Over the last decade, he worked mostly in independent evaluation units and was responsible for management and implementation of evaluations in various intervention sectors, as well as development of evaluation methodologies and strategies.

**Ningqin Wu** is project coordinator of the Asia-Pacific Finance and Development Center and the lead project coordinator of the Center for Learning on Evaluation and Results–Shanghai International Program for Development Evaluation Training. Since 2007, she has been involved in the operation and management of the Shanghai International Program for Development Evaluation Training and other monitoring and evaluation related training programs. She has also helped translate monitoring and evaluation-related papers and books such as the *Road to Results* (2012).

**Ivory Yong-Prötzel** is head of Operations Evaluation at the European Investment Bank Group. She joined in 2006 with a solid background in ex post evaluation and a particular interest in evaluation methods. Her expertise covers the evaluation of public policy programs, with a particular focus on productive sectors financed through public and private operators, as well as the design, testing, and dissemination of evaluation methodologies and related quality assurance processes. Before joining the European Investment Bank, she was head of the evaluation department in a European consulting firm, where she led or participated in numerous high-level evaluations on European cooperation program. Before getting involved in evaluation, she worked as a consultant in economic

policy. She began her professional activity in academia, specializing in macroeconomics and econometrics, focusing on applied economics and with an interest in hyperinflation and public debt.

**Min Zhao** is associate professor of the Shanghai National Accounting Institute and deputy director of the Asia-Pacific Finance and Development Center. She earned a PhD in economics from Shanghai Jiaotong University in 2007. At present, she is the team leader of the Center for Learning on Evaluation and Results-Shanghai International Program for Development Evaluation Training. Her research interests are in public finance and performance management, particularly subnational debts management. She has published almost 20 articles in international journals such as *Public Finance Research*, *Journal of Asian Innovation and Entrepreneurship*, and *Asia and Pacific Economics*.

## Evaluation for Better Results

This book marks 10 years of independent evaluation at the Asian Development Bank. It shares the journey to organizational and financial independence of evaluation and the transformation in evaluation approaches to support effectively accountability and learning. The book captures how evaluation is advancing from largely ex post assessments of programs and projects to a varied set of forward-looking and real time tools and innovations in evaluation methods.

The book offers a collection of papers from distinguished development practitioners and evaluators around the world who share their experiences in capacity development for evaluation at the country level. Evaluation for better results requires connecting with development challenges in a changing development context, focusing on and measuring outcomes, and proposing lessons and solutions to improve results. Tried and tested approaches may not be enough by themselves for dealing with unfamiliar situations as well as with past development problems that remain intractable. Innovative approaches are crucial in evaluation methods and outreach to enhance rigor and influence.

## About the Asian Development Bank

ADB's vision is an Asia and Pacific region free of poverty. Its mission is to help its developing member countries reduce poverty and improve the quality of life of their people. Despite the region's many successes, it remains home to approximately two-thirds of the world's poor: 1.6 billion people who live on less than \$2 a day, with 733 million struggling on less than \$1.25 a day. ADB is committed to reducing poverty through inclusive economic growth, environmentally sustainable growth, and regional integration.

Based in Manila, ADB is owned by 67 members, including 48 from the region. Its main instruments for helping its developing member countries are policy dialogue, loans, equity investments, guarantees, grants, and technical assistance.

Asian Development Bank  
6 ADB Avenue, Mandaluyong City  
1550 Metro Manila, Philippines  
[www.adb.org](http://www.adb.org)

 Printed on recycled paper



Printed in the Philippines