



# 2004 Annual Report on Operations Evaluation



© 2005 The International Bank for Reconstruction and Development / The World Bank  
1818 H Street, NW  
Washington, DC 20433  
Telephone: 202-473-1000  
Internet: [www.worldbank.org](http://www.worldbank.org)  
E-mail: [feedback@worldbank.org](mailto:feedback@worldbank.org)

All rights reserved  
Manufactured in the United States of America  
First edition July 2004; 2005

The findings, interpretations, and conclusions expressed here are those of the author(s) and do not necessarily reflect the views of the Board of Executive Directors of the World Bank or the governments they represent.

The World Bank cannot guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in the work do not imply on the part of the World Bank any judgment of the legal status of any territory or the endorsement or acceptance of such boundaries.

#### **Rights and Permissions**

The material in this publication is copyrighted. Copying and/or transmitting portions or all of this work without permission may be a violation of applicable law. The World Bank encourages dissemination of its work and will normally grant permission promptly.

For permission to photocopy or reprint any part of this work, please send a request with complete information to the Operations Evaluation Department, World Bank, 1818 H St., N.W., Washington, D.C. 20433.

Photo courtesy of the World Bank Photo Library.

World Bank InfoShop  
E-mail: [pic@worldbank.org](mailto:pic@worldbank.org)  
Telephone: 202-458-5454  
Facsimilie: 202-522-1500

Operations Evaluation Department  
Knowledge Programs and Evaluation Capacity  
Development (OEDKE)  
E-mail: [eline@worldbank.org](mailto:eline@worldbank.org)  
Telephone: 202-458-4497  
Facsimilie: 202-522-3125



Printed on Recycled Paper

---

# Contents

vii	<b>Acknowledgments</b>
ix	<b>Foreword</b>
xiii	<b>Acronyms and Abbreviations</b>
1	<b>1 Introduction</b>
	1 Objective of the Report
	1 Key Concepts and Assessment Framework for This AROE
	5 Organization and Sources
7	<b>2 The Bank's Corporate Approach to Results and M&amp;E</b>
	7 An Evolving Agenda
	8 Addressing Challenges
15	<b>3 M&amp;E in Bank Guidelines for Lending and Nonlending Products</b>
	15 For Country Products, Guidelines Are Going in the Right Direction
	16 For Sector Products, Guidelines Are Silent on Results
	18 For Lending, Guidelines Are a Work in Progress
	20 In the Bank's Analytical Products, M&E Is Still Rare
23	<b>4 Institutional and Organizational Factors Affecting the Bank's Focus on M&amp;E and Managing for Results</b>
	23 Traditional Approach Versus a Results-Oriented Approach
	23 Staff and Managerial Incentives in the Bank Do Not Fully Support the Adoption of the Results Agenda
	27 Decisionmaking Based on Outcome-Related M&E Information Is Weak
	29 Need for Clarity about Roles and Responsibilities
35	<b>5 Improving OED's Effectiveness</b>
	35 Results-Oriented M&E in OED Guidelines
	37 OED's Response to Recommendations for Itself Contained in Last Year's AROE
	38 Bank Staff Awareness about OED Studies and Their Perception of Quality
	39 OED's Country Assistance Evaluations
	40 OED's Support to the International Program for Development Evaluation Training

<b>41</b>	<b>6 Recommendations</b>
	41 For Bank Management
	42 For OED
<b>43</b>	<b>Annexes</b>
	45 A: Product Initiation, Implementation, and Completion: Guidelines, Templates, and Assessments Reviewed
	47 B: CAS Retrospective's Thematic Results Matrix
	49 C: Interviewee Matrix and Interview Guide
	51 D: OED Survey Results
	55 E: Management Response
	63 F: Chairman's Summary: Committee on Development Effectiveness (CODE)
<b>67</b>	<b>Endnotes</b>
<b>71</b>	<b>Bibliography</b>
	<b>Boxes</b>
	3 1.1 Measuring Impact
	3 1.2 Bangladesh Health and Population Program Project: Importance of Measuring All Key Links in the Causal Chain
	4 1.3 Three Criteria for Assessing the Results Orientation of the Bank's M&E System
	10 2.1 M&E: A Difficult Business Requiring Serious and Sustained Effort
	11 2.2 Linking Performance Information to Resource Allocation
	13 2.3 Changing Organizational Incentives: The Central Challenge
	16 3.1 M&E: Whose Responsibility?
	17 3.2 Building Client M&E Capacity
	20 3.3 ICRs Lack Information on Results
	21 3.4 Results-Oriented ESW/AAA: Some Examples
	25 4.1 Managing for Results in DFID: Linking Personal and Corporate Objectives
	27 4.2 Pay Based on Results
	28 4.3 Documenting the Results of Bank Projects in the Europe and Central Asia Region (ECA)
	29 4.4 Indonesia CAS: Building M&E Capacity through Hiring an Experienced M&E Specialist to Work Closely with Bank Staff and Clients
	30 4.5 Results Focus: Highlights of Regional Initiatives
	31 4.6 Producing and Maintaining Indicators Is Costly
	31 4.7 Getting Results in Pakistan's Petroleum Sector
	36 5.1 Brazil CAE: Good Assessment of the Results Framework of the Bank's Country Program
	37 5.2 Involving Borrowers in an Impact Evaluation: Basic Education in Ghana
	<b>Tables</b>
	2 1.1 The Causal Chain for a Child Health Program for the Poor
	9 2.1 Changing Focus of the Bank's Corporate Approach
	12 2.2 Selected Measures of Development and Organizational Effectiveness
	18 3.1 Results-Based CAS Adds Value

24 4.1 Changing to a Results-Oriented Bank  
26 4.2 An Incentive System for Results Would Reward the Later Stages More  
32 4.3 How Specific Organizational Approaches Can Foster or Inhibit  
Managing for Results

**Figures**

19 3.1 Insufficient Evidence on Outcome Is Leading Reason for Disconnect  
in ECA  
39 5.1 OED Studies Ranked on Timeliness and Relevance to Work



---

# Acknowledgments

This report was prepared by a team led by Soniya Carvalho. Manuel Penalver advised and Sue Berryman, Nils Boesen, Deepa Chakrapani, John Eriksson, Matthew Stephenson, Harold Wackman, and Howard White wrote background papers. Ruth Driscoll, John Greiner, Rupa Ranganathan, Arleen Rogan, and Susan Seling provided inputs. Gisu Mohadjer and Padmanabha Hari Prasad commented on a draft. The team was assisted

by Annisa Cline-Thomas, Julia Ooro, and Yezena Yimer. Document production support was provided by William Hurlbut. Caroline McEuen edited the report.

OED is grateful to the Swiss Agency for Development Cooperation for its review and financial assistance.

The Report was prepared under the direction of Victoria Elliott, Manager of OEDCM.

---

Director-General, Operations Evaluation:

*Gregory K. Ingram*

Director, Operations Evaluation Department:

*Ajay Chhibber*

Manager, Corporate Evaluation and Methods:

*Victoria Elliott*

Task Manager: *Soniya Carvalho*







# Foreword

The World Bank and the development community are under increasing pressure to demonstrate the results of development assistance. Monitoring and evaluation (M&E) can help meet this challenge by providing the evidence needed to demonstrate results and by giving program managers information they can use to improve results.

An M&E system that is results-oriented has three characteristics. It is based on a results framework that lays out the causal chain between inputs and intended outcomes. It includes measurable indicators that permit tracking of the key links in the causal chain. And it specifies the arrangements for collection, analysis, and use of the data.

This year's *Annual Report on Operations Evaluation* (AROE) assesses how results-oriented the Bank's M&E systems are, and the extent to which they contribute to managing for results.

International experience shows that organizations typically focus more attention on the technical and procedural aspects of M&E systems than on changing organizational cultures and incentives. This review finds that the Bank's most recent results initiative, launched in September 2002, is an important development that has raised staff awareness of results orientation. The Bank has improved the results orientation of a number of its product guidelines and has begun to collect better

information for decisionmaking, especially at the country-program level. But the Bank has not yet given adequate attention to creating an organizational culture that values lesson learning or an incentive framework that promotes managing for results. The recommendations in this AROE are intended to help the Bank move forward faster and more effectively toward results-oriented M&E and managing for results.

## Corporate Vision

The Bank's corporate documents have spelled out the Bank's approach and an implementation plan for better measuring, monitoring, and managing for results. The message that "results are important" has been reinforced by high-level statements and a Bankwide learning event. But the Bank has not yet articulated what a results-oriented Bank is, what kinds of M&E data are needed to inform decisionmaking at each level, and what kinds of incentives should prevail. Such an articulation would help the Bank to be selective in collecting M&E data—without creating a "measurement bureaucracy."

Furthermore, it would help ensure that the various initiatives ongoing across the Regions fit within the larger institutional effort and promote results-oriented staff behavior.

Management has correctly attributed the limited achievements of the 2000 M&E Improvement Program to inappropriate incentives, unclear roles and responsibilities, and capacity constraints. These three problems remain as issues that the Bank needs to address.

### **Operational Guidance**

The Bank has changed, or is changing, the guidelines and templates for many of its lending and nonlending products and needs to close the remaining gaps in guidelines for the initiation, implementation, and completion of all its major products. This is especially important for the majority of the Bank's analytical products, for which the systematic measurement of results has not yet been attempted, and at the sector level, where self-evaluation of sector strategy implementation remains particularly weak.

The Bank has also begun to collect better information about the results of its country programs, notably through the results-based Country Assistance Strategies (CASs). A handful of these have been prepared to date—their contribution to a results focus will depend on how well they are implemented. And although work at the sector level is less advanced than at the country level, improvements in information for decisionmaking at the sector-program level may result from the implementation of management's Implementation Follow Up (IFU-7) initiative. The IFU-7 initiative is also still in its early stages.

### **Accountabilities of Sectors, Networks, and Individual Staff**

To make further progress toward a results orientation, the Bank needs to address three issues of accountability. First, managers and staff do not have a common understanding of the roles of the anchors, networks, and sectors in managing for results. They observe that the sectoral dimension of the matrix is fragmented, leading to excessive management and transactions costs, task proliferation, and dilution of

the results focus. Second, the fragmentation of the sectors works against the cross-sectoral collaboration that is necessary for achieving development outcomes.

Finally, Bank staff consider that “getting results” is not yet part of the rewards system for individual staff. The Bank's incentives reward the work staff do at the early stages of the project cycle—that is, preparing new operations—more highly than their work at the later stages of supervision—evaluation and lesson learning.

### **OED Can Support the Results Agenda More Effectively**

OED guidelines for its country, sector, corporate, and project evaluations do not include specific criteria for assessing the results framework, design of M&E, and implementation of M&E, although some actual evaluations show greater attention to results and M&E than the guidelines suggest. OED findings are not sufficiently well known to operational staff: the most recent survey showed that just 27 percent of respondents had read, or participated in an event related to, recent evaluations in their area of responsibility.

## **Recommendations**

### **For Bank Management**

- Develop and maintain a sustained vision, expressed in consistent messages to staff, of a results-oriented Bank and how organizational culture, incentives, and structure need to change. In particular, management should address the sectoral fragmentation and unclear accountabilities that the matrix structure currently entails.
- Develop a phased and costed plan showing how, and by when, different parts of the Bank would implement results management. Revise the full set of product guidelines and Quality Assurance Group questionnaires, and provide other operational guidance for staff, to indicate how the content and process of operations need to reflect managing for results, especially with respect to collecting and using results-oriented M&E data.

- Review and strengthen incentives such as managerial recognition, career advancement, and reward systems to promote results-oriented M&E, managing for results, and lesson learning.

**For OED**

- Strengthen attention to results-oriented M&E in OED by requiring OED evaluations at the country, sector, and project levels to address

systematically the three key criteria for results-oriented M&E, and increase the weight that OED ratings give to the extent of results-oriented M&E and managing for results in Bank programs.

- Communicate and disseminate findings and lessons of OED evaluations more widely both within and outside the Bank, in line with the new communications strategy under preparation.



Gregory K. Ingram

Director-General, Operations Evaluation



---

# Acronyms and Abbreviations

AAA	Analytical and advisory activities
ACS	Activity Completion Summary
AFR	Africa Region (World Bank Region)
APL	Adaptable Program Loan
ARDE	Annual Review of Development Effectiveness
AROE	Annual Report on Operations Evaluation
ARPP	Annual Review of Portfolio Performance
CAE	Country Assistance Evaluation
CAS	Country Assistance Strategy
CASCR	CAS Completion Report
CDF	Comprehensive Development Framework
CPIA	Country Policy and Institutional Assessment
DEC	Development Economics and Chief Economist (World Bank vice presidency)
DFID	Department for International Development (U.K.)
EAP	East Asia and Pacific Region (World Bank Region)
ECA	Europe and Central Asia Region (World Bank Region)
ESSD	Environmentally and Socially Sustainable Development (World Bank Network)
ESW	Economic and sector work
IBRD	International Bank for Reconstruction and Development (World Bank)
ICR	Implementation Completion Report
IDA	International Development Association
IFU-7	Implementation Follow Up-7
IPR	Implementation Progress Report
LCR	Latin America and the Caribbean Region (World Bank Region)
LIL	Learning and Innovation Loan
MDGs	Millennium Development Goals
M&E	Monitoring and evaluation
MNA	Middle East and North Africa Region (World Bank Region)
OECD	Organization for Economic Cooperation and Development
OED	Operations Evaluation Department

OPCS	Operations Policy and Country Services (World Bank vice presidency)
PAD	Project Appraisal Document
PPAR	Project Performance Assessment Report
PREM	Poverty Reduction and Economic Management (World Bank Network)
PRSC	Poverty Reduction Support Credit
PRSP	Poverty Reduction Strategy Paper
PSD	Private sector development
PSR	Project Status Report
QAE	Quality at entry
QAG	Quality Assurance Group
SAR	South Asia Region (World Bank Region)
SSIU	Sector Strategy Implementation Update
SSP	Sector Strategy Paper
STATCAP	Statistical Capacity Building Program



# Introduction

For more than a decade, aid organizations have faced considerable external pressure to become more effective and results-oriented. The perception that aid has failed to produce significant results, combined with declining aid budgets and reforms in Organization for Economic Cooperation and Development (OECD) countries aimed at results-based management, have been driving bilateral and multilateral organizations to search for new ways to improve development effectiveness.

Many of them have launched agendas of results-orientation and results-based management (more recently known as “managing for results”). The international endorsement of the Millennium Development Goals (MDGs) has given additional impetus to the quest for results and for evidence demonstrating their achievement.

## Objective of the Report

*The Annual Report on Operations Evaluation* (AROE) fulfills the OED mandate to assess the monitoring and evaluation of development effectiveness of World Bank activities. The objective of this year’s AROE is to assess how results-oriented the Bank’s monitoring and evaluation (M&E) systems are, and the extent to which they contribute to managing for results in the Bank. It focuses on actions the Bank has taken to improve its own systems; it does not examine actions taken to improve country capacity or harmonization with other

development agencies.<sup>1</sup> The 2003 AROE identified the Bank’s frameworks for both monitoring and evaluation and commented on their overall effectiveness.<sup>2</sup> This AROE uses the same frameworks to focus on the results orientation of M&E, including clarity in the definition of intended results, the associated causal chains, and related performance measurement. It also discusses the institutional and organizational factors affecting the Bank’s focus on M&E and managing for results.

## Key Concepts and Assessment Framework for This AROE

This AROE adopts an assessment framework that derives from the logical framework (log frame) approach adopted by many aid organizations, including the Bank.<sup>3</sup> Key concepts—including the log frame approach—are discussed first, and the assessment framework is presented next.

### Key Concepts

In the log frame approach, the causal chain is as follows:<sup>4</sup>

- The objective of interventions is to improve the welfare **outcomes** of the target group, based upon the achievement of...
- **Intermediate outcomes**, such as increased consumption of goods and services, as a result of...
- **Outputs**, such as the goods produced or services delivered, as a result of ...
- **Inputs**, such as the financing, organizational capacities, and human capital provided by the intervention.

Table 1.1 illustrates the links in the causal chain for a program to improve child health among the poor. Periodic impact evaluation and related research are needed to verify the links in the causal chain underlying the logic of project design (see box 1.1).

#### *How is the log frame approach linked to M&E?*

Understanding the causal chain that underlies a program is key to its systematic monitoring and evaluation. Monitoring verifies whether each step in the chain is actually materializing. It helps to answer “what has happened?” Evaluation examines why each step may or may not be materializing. It helps to answer “why has change occurred?”

<b>Table 1.1</b>		<b>The Causal Chain for a Child Health Program for the Poor</b>	
<b>Causal chain</b>		<b>Performance indicators (to be disaggregated by poverty status and gender)</b>	
Outcomes <sup>a</sup>	↑	Infant and child mortality Morbidity Stunting (height/age) and wasting (weight/height) Total fertility rate	
Intermediate outcomes	↑	Access to potable water Percentage of women who know how to use oral rehydration salts correctly Contraceptive prevalence rate Number of children using supplements and oral rehydration salts Family planning services utilized	
Outputs	↑	Number of trained health staff Number and coverage of children in well-baby clinics Number and coverage of immunization campaigns/services Family planning services/contraceptives supplied or sold Vaccination campaigns/quantity vaccines delivered Nutritional supplements and oral rehydration salts distributed New institutions and strategies established	
Inputs		Spending on maternal and child health programs Private and donor spending on services Public policy/regulatory environment Strategic planning	

*Note:* Exogenous factors such as household income, food prices, and educational level of the mother will also be important determinants of program outcomes and should be accounted for.

a. Some specifications of the log frame include **impacts** in addition to **outcomes**. In most cases, the distinction is difficult and may not be necessary, although in a few instances it is helpful (for example, for an education project, learning achievement may be classified as an **outcome** and the resulting changes in earnings/incomes as an **impact**). This AROE uses the term **impact** to refer to change in **outcomes** attributable to the intervention.

*Source:* Background work, 2004 AROE.



**Box 1.1****Measuring Impact**

Impact evaluations examine the change in welfare outcomes that may be attributed to a project, program, or policy. Implicit in this definition is the notion of the counterfactual: what happened to welfare outcomes compared with how they would have fared in the absence of the intervention? Such evaluations, therefore, differ from many other evaluations that focus on inputs and the implementation process. Impact evaluation is demanding in its data requirements and can be costly compared with other sorts of evaluation.

For most M&E, the measurement system has to rely on a chain of plausible association. That is, project design should be based on a clearly specified model (log frame) of how the proposed inputs will, through the supported activities, produce the outputs to have the desired outcomes. The monitoring system needs to collect data on each of these links in the causal chain. However, periodic impact evaluations and related research are needed to verify the links in the causal chain underlying the logic of project design.

Source: White 2004a.

**How is M&E linked to results?** M&E is critical for achieving results: it provides managers with information they can use to verify if they are on course and to improve program performance. It is also critical for demonstrating results, because it provides the needed evidence. M&E promotes lesson learning and accountability for results.

**How do M&E and results come together in program design?** Results orientation means that M&E should be part of the program design process—a program designer should start by determining the intended results, figuring out the most efficient way to get them, and determining how he/she would know if they were (or were not) materializing and why. This process makes it obvious what is to be monitored, and the evaluation design becomes integral to program design.

**What are results and how should they be monitored and evaluated?** This AROE draws on the OECD/

Development Assistance Committee (DAC) definition of results: “the output, outcome, or impact<sup>5</sup> of a development intervention.”<sup>6</sup> A focus on outcomes and impacts does not imply that inputs, processes, and outputs should not be monitored or evaluated. Determining whether outcomes are materializing, and why they are—or are not—materializing, requires assessing all key links in the causal chain, including inputs, processes, outputs, and outcomes. The importance of assessing all key links in the causal chain is illustrated by the Bank-supported Bangladesh Health and Population Program Project (see box 1.2). A results orientation implies that outcomes are the starting point in conceptualizing program design and M&E. It requires that outcomes (or variables that are good proxies for outcomes) be regularly monitored.<sup>7</sup>

**Assessment Framework**

**How can the results orientation of the Bank’s M&E systems be assessed?** Based on the log frame

**Box 1.2****Bangladesh Health and Population Program Project: Importance of Measuring All Key Links in the Causal Chain**

This project supports the government’s efforts to improve the health status of the poor. The project’s M&E system includes the systematic collection and analysis of data relating to outcomes, intermediate outcomes, outputs, and inputs. Most outcome indicators (such as under-five mortality) have continued to improve, even though intermediate outcome indicators (for example, utilization of project

services) have not improved much. This means that exogenous factors lie behind improvements in the outcome indicators, and that problems related to utilization need to be tackled. By tracking intermediate outcome indicators (not just outcome indicators), the government and the Bank were able to identify weaknesses in the project, even though outcome indicators were improving.

Source: White, based on Bhuiyan and Hay 2003.

approach and the concepts discussed above, this AROE identifies three criteria for assessing the results orientation of the Bank's M&E systems:<sup>8</sup>

- Existence of a **results framework**, which lays out the objectives of the intervention and the assumed causal chain between funding and the intended outcomes
- **Design of M&E**, which includes measurable indicators that enable tracking of all key links in the causal chain
- **Implementation of M&E**, which includes arrangements and responsibilities for data collection, analysis, and utilization.

These criteria are discussed in more detail in box 1.3.

***How can the extent to which the Bank's M&E systems are contributing to managing for results be assessed?***

This AROE uses “managing for results” synony-

mously with “results-based management” and draws on the OECD/DAC definition of results-based management: “a management strategy focusing on performance and achievement of outputs, outcomes, and impacts.”<sup>9</sup> While results-oriented M&E refers to performance measurement, monitoring, and evaluation issues, “managing for results” refers to issues of performance management. These processes are linked, in that managing for results can occur effectively only in the presence of results-oriented M&E. Results-oriented M&E underpins, and is the main tool for, managing for results. Accordingly, an assessment of the extent of the results orientation of the Bank's M&E systems will also illuminate the Bank's ability to manage for results. Additional analysis of key institutional and organizational issues affecting the Bank's focus on M&E and managing for results—issues of organizational incentives, decisionmaking processes, and structure—can help to give a more complete

**Box 1.3**

**Three Criteria for Assessing the Results Orientation of the Bank's M&E System**

**Results Framework**

- Statement of Bank objectives and time-bound targets at both the intermediate and final outcome levels
- Specification of the causal chain between Bank funding and the intended outcomes
- Specification of the expected time lags between links in the causal chain
- Identification of exogenous factors (enabling conditions and threats) that affect achievement of expected outcomes
- Identification of potential unintended consequences of the intervention, whether positive or negative.

**Design of M&E**

- A performance indicator system that tracks all key links in the causal chain, monitors and evaluates the status of exogenous factors, and picks up unintended consequences
- A performance indicator system that meets three standards: simplicity, attribution/contribution, and a schedule for measuring each indicator that matches the time lags between links in the causal chain

- Individual indicators that are reliable, valid, comparable, specific, measurable, realistic, and targeted
- An analytic plan for reporting progress in relation to plan, analyzing shortcomings, determining why targets and outcomes are or are not being achieved, and determining whether and how result failures can be attributed to implementation failures.

**Implementation of M&E**

- For a new activity: Program design that describes organizational arrangements; allocation of responsibilities; and institutional capacity for collecting, analyzing, and utilizing M&E data.
- For an activity under way: M&E data being collected in a timely way. M&E data being utilized to make decisions about the intervention.
- For a completed activity: Availability of the right data to determine whether the intended outcomes occurred, to assess why or why not, and to identify lessons for future activities.

Sources: Background work, 2004 AROE; Mosse, Sontheimer, and Rogers 1996.

picture of the Bank's ability and willingness to manage for results and of the contribution of results-oriented M&E to managing for results.

### **Organization and Sources**

Chapter 2 assesses the Bank's corporate approach to results-oriented M&E and managing for results. Chapter 3 examines the results orientation of M&E in the Bank's guidelines and templates for lending and nonlending products at the country, sector, project, and analytical levels. Although guidelines and templates are necessary for securing a results orientation, they are by no means sufficient. Therefore, Chapter 4 addresses institutional and organizational factors that affect the Bank's focus on M&E and managing for results, and Chapter 5 addresses OED's contribution to it. Chapter 6 presents

recommendations for both Bank management and for OED.

This AROE draws on the March 2004 Operations Policy and Country Services (OPCS) report, *Implementation of the Agenda on Managing for Results: Progress Report* (hereafter, the 2004 OPCS Progress Report). It also draws on other reports from OPCS as well as on Quality Assurance Group (QAG) reports and OED evaluations. Semi-structured interviews were conducted with 34 Bank staff to assess issues of organizational culture, incentives, and structure (Annex C presents the Interviewee Matrix and the Interview Guide). In addition, a survey of 3,159 Bank staff was conducted to obtain their views on OED evaluations (Chapter 5 summarizes the results, and details are presented in Annex D).





# The Bank's Corporate Approach to Results and M&E

The Bank launched a renewed results initiative in September 2002, entitled “Better Measuring, Monitoring, and Managing for Development Results.” This important initiative has raised staff awareness of results orientation. The message that *results are important* has been reinforced by high-level statements and a Bankwide learning event. The Bank has started the process of change toward results-oriented M&E and managing for results, with an initial focus on improving its product guidelines and templates.

The Bank has also begun to collect better information about the results of its country programs, notably through the results-based Country Assistance Strategies (CASs). This chapter traces the evolution of the Bank's results agenda and discusses the challenges posed by a move to results-oriented M&E and managing for results.

## **An Evolving Agenda**

The Bank's emphasis on results and M&E is not new. The Bank has been working to increase its results orientation ever since the Wapenhans report of 1992, which concluded that “the Bank's success is determined by benefits ‘on-the-ground’—sustainable development impact—not by loan approvals, good reports, or disbursements” (World Bank 1992). The follow-on initiative to the Wapenhans report, *Getting Results: The World Bank's Agenda for Development Effectiveness* (World Bank 1993), aimed at shifting the emphasis from physical and financial

aspects of implementation to performance assessments and progress toward objectives.<sup>1</sup>

In response to these initiatives, M&E received increasing attention in the Bank. In 2000 a working group was formed to improve the quality of M&E in Bank-financed operations (hereinafter, the M&E Improvement Program). That group noted that M&E had remained a “persistent problem” throughout the 1990s (World Bank 2000), a finding echoed in OED's 2002 AROE: “The system of monitoring in projects remains weak, in spite of repeated OED and QAG recommendations and episodic efforts over the last decade to strengthen it” (OED 2002). The working group on the M&E Improvement Program sought to address the Bank's three main M&E problems: poor incentives for good M&E among Bank staff and borrowers, diffuse accountability and lack of clarity on M&E roles and responsibilities within the Bank and between Bank staff and borrowers, and weak Bank staff and borrower capacity

for M&E (OED 2002). In its report, the working group advocated a shift away from the traditional M&E model, which concentrated on tracking inputs and deliverables, and a move toward a results-based M&E approach focusing on “results, with particular emphasis on development outcomes.” It also introduced the concept of “managing for results,” stating that “monitoring that focuses on results becomes synonymous with management.” Recognizing that change would take time, the working group proposed an action plan for fiscal years 2001–04 that included piloting results-based M&E in selected countries and sectors, with the target that results-based M&E would begin to be mainstreamed in Bank operations in fiscal year 2004—a target that remains unmet.<sup>2</sup>

In September 2002, the M&E Improvement Program was subsumed under a broader initiative, “Better Measuring, Monitoring, and Managing for Development Results,” which shifted the focus toward M&E at the country-program level through the results-based CAS. An Implementation Action Plan, “Better Measuring, Monitoring, and Managing for Development Results: Implementation Action Plan” (hereafter, the 2003 Implementation Action Plan), spelled out the operational implications of the initiative.

The 2004 OPCS Progress Report assesses work toward implementing the 2003 Implementation Action Plan. It also provides an assessment of the M&E Improvement Program:

Although considerable efforts were undertaken to improve M&E, to date these efforts have produced only marginal improvements in the quality of M&E in operations . . . unfortunately, the three causes of poor M&E (inappropriate incentives, unclear roles and responsibilities, and capacity constraints) were valid predictors of the challenges the program faced. In addition, two critical program assumptions did not hold true during implementation (that resources would be readily available to support internal capacity building, and existing staff had sufficient skills and interest in M&E).

[World Bank 2004]

These causes of poor M&E still exist. The Bank needs to squarely address them.

Table 2.1 shows the changing focus of the Bank’s corporate approach, including the changing attention to incentives as reflected in the corporate documents on results. It also points to a number of improvements that are necessary in the Bank’s approach to results.

### Addressing Challenges

This section examines the extent to which the Bank’s corporate approach to results has addressed the three main challenges to results-oriented M&E and managing for results highlighted by international experience: (i) determining the right balance between measuring outputs versus outcomes; (ii) resolving the tension between attribution and accountability; and (iii) changing the organizational culture, incentives, and structure. These challenges are fully within the Bank’s powers to address and are discussed in some detail below. The Bibliography lists some articles from the general literature on results and M&E that discuss the international experience. Box 2.1 summarizes the problems faced by traditional M&E efforts—problems that may persist or be compounded when a results focus is adopted and that deserve continued attention.

The Bank also faces two other challenges that are beyond its ability to address alone. They are: (i) country capacity and will to manage for results and (ii) capacity for harmonized and coordinated action within the international community. To address these challenges, the Bank has supported statistical and evaluation capacity development in a number of countries and is participating in global partnerships through the Multilateral Development Bank (MDB) Working Group and, more recently, through the MDB/OECD-DAC Joint Venture on Managing for Results. The Bank is contributing to three areas of the global partnership: knowledge sharing, harmonization and coordination with other development agencies, and improving in-country statistical systems. As noted in the Introduction to this volume, this AROE focuses

Table 2.1

### Changing Focus of the Bank's Corporate Approach

Corporate document	Key components/pillars	Comments
2000 M&E Improvement Program	<ul style="list-style-type: none"> <li>• Improve Bank and borrower incentives for M&amp;E.</li> <li>• Clarify Bank and borrower roles and accountability for results-based M&amp;E within the Bank.</li> <li>• Build Bank and borrower capacity.</li> </ul>	<ul style="list-style-type: none"> <li>• Focused on Bank-supported projects and did not extend its work to the evaluation of sector strategies or Country Assistance Strategies (CASs).</li> <li>• Incentives and accountability were top component.</li> <li>• Adopted a target: results-based M&amp;E would begin to be mainstreamed in Bank operations in fiscal year 2004.</li> <li>• Focused on select sectors and pilot countries.</li> <li>• Included detailed costing of the program.</li> </ul>
2002 Better Measuring, Monitoring, and Managing for Development Results	<ul style="list-style-type: none"> <li>• In developing countries, focus on country-level outcomes, knowledge, information, and capacity.</li> <li>• In and across development agencies, focus on aid effectiveness and partnership.</li> <li>• Within the Bank, focus on defining and measuring operational results and on corporate issues, including reporting, incentives and training, and risk management.</li> </ul>	<ul style="list-style-type: none"> <li>• Shifted focus to country-program level through results-based CAS.</li> <li>• Incentive issue relegated to the third component.</li> <li>• Did not specify intermediate outcomes relating to the results the Bank was seeking.</li> <li>• Did not discuss program cost.</li> </ul>
2003 Implementation Action Plan	<ul style="list-style-type: none"> <li>• Country focus and ownership, focusing on country outcomes, including MDGs and other priorities, and coordinated donor support for enhancing country M&amp;E capacity.</li> <li>• Bank strategy and instruments, where design and piloting of results-based CASs is central.</li> <li>• Corporate reporting, staff learning and incentives.</li> </ul>	<ul style="list-style-type: none"> <li>• Objectives stated mostly as activities (e.g., help countries build statistical systems, report on results in corporate strategy and budget documents). Final or intermediate outcomes not specified.</li> <li>• Incentive issue continued as part of third component.</li> <li>• Silent on if and how the emerging results data will guide the Bank's allocation of lending and administrative resources.</li> <li>• Did not set measurable targets.</li> <li>• Did not include detailed costing of the program.</li> <li>• Promised that "the rationale and revised ["joined-up"] architecture will be detailed in the forthcoming Monitoring and Evaluation Progress Report to be reviewed by the Committee on Development Effectiveness (CODE) in the fourth quarter of fiscal year 2003," an unfulfilled promise.</li> </ul>
2004 OPCS Progress Report	<ul style="list-style-type: none"> <li>• In countries where development results are achieved, to strengthen both capacity and demand to manage for results</li> <li>• In the Bank, to enhance the relevance and effectiveness of Bank contribution to results</li> <li>• Across development agencies, to harmonize results-based approaches and better coordinate support to strengthen country capacity to manage for results.</li> </ul>	<ul style="list-style-type: none"> <li>• Candid reporting that the M&amp;E Improvement Program produced only marginal improvements. Also noted reasons for this poor performance. Did not, however, examine the extent to which each of these problems has now been addressed to ensure the success of the results agenda this time around.</li> <li>• Did not report back systematically on the 2003 Implementation Action Plan, but rather to a conceptually revised plan that puts partnership with aid agencies higher on the agenda. Should have dealt explicitly with the change of emphasis.</li> <li>• Included a framework for future actions and identified intermediate outcomes (e.g., "improvement in results orientation of PRSPs assessed in annual PRSP Progress Report"). But some intermediate outcomes are difficult to verify (e.g., "increased focus on performance in expenditure planning and management").</li> <li>• Did not provide details on cost of implementing the results agenda.</li> <li>• This report, which took the place of the Monitoring and Evaluation Progress Report mentioned in the 2003 Implementation Action Plan, did not detail the "joined-up" architecture as was promised in the 2003 Implementation Action Plan.</li> </ul>

Source: The "key components/pillars" column is based on corporate documents cited in the text; the "comments" column is based on background work, 2004 AROE.

**Box 2.1****M & E: A Difficult Business Requiring Serious and Sustained Effort**

The Bank's heavy investment in M&E in the 1980s met with little success. A key lesson from that experience was that M&E systems that are divorced from project management are unlikely to be used, even if they are designed perfectly. Care should be taken not to create a "measurement bureaucracy" and an M&E system that measures all that can be easily measured, but neglects what is important (but more difficult) to measure. Upfront clarity on the objective and on how the emerging data will be used can help avoid this. M&E is the tool. Results are the end.

The time and effort it takes to put M&E systems in place—including identifying a parsimonious set of key performance

Sources: OED 1985; background work, 2004 AROE.

indicators, determining qualitative data requirements, identifying the best modes of data collection sensitive to the type of data being collected, setting up control groups, collecting baseline data, ensuring ownership and buy-in from data producers as well as consumers, and determining feedback loops—should not be underestimated. There can also be tradeoffs: for example, between the speed with which information is desired and the depth of M&E. Continuous fine-tuning of the M&E system will be critical as objectives or context evolve and information needs change.

on actions the Bank has taken to improve its own systems, and does not examine actions taken to improve country-capacity or harmonization with development agencies.

#### **Determining the Right Balance Between Measuring Outputs Versus Outcomes**

As noted in Chapter 1, determining whether and why outcomes are (or are not) materializing requires an assessment of all key links in the causal chain. The Bank has yet to develop a complete list of input, output, and outcome indicators that it can monitor regularly and use as a tool for corporate decisionmaking. QAG's *Annual Report on Portfolio Performance* (ARPP) is an apex institutional report on the operational portfolio. It reports on portfolio size, composition, and quality. The 2003 ARPP made a first attempt to incorporate more results information, but transforming the ARPP into an Operational Performance and Results Review has been postponed to fiscal year 2007 because of the lack of relevant results-oriented M&E data in Bank operations. The absence of such reporting limits the scope of corporate decisionmaking to be grounded in results information.<sup>3</sup>

The 2003 Implementation Action Plan reported that the Bank was putting in place a "joined-up" M&E architecture that extended logically from the country-level CAS down through the products and services that are aligned with the CAS (World Bank 2003a, p. 11).

It noted that this architecture was designed to help close the gap that has often existed between, for example, project outputs (such as trained midwives) and higher-order country outcomes (for example, reduction in maternal mortality). Individual operations would articulate the link between outputs and expected outcomes directly influenced by project activities, while the CAS would articulate (and later validate) the posited causal links between these intermediate outcomes and higher-order country outcomes. It promised that "the rationale and revised architecture will be detailed in the forthcoming Monitoring and Evaluation Progress Report, to be reviewed by CODE in the fourth quarter of FY03" (World Bank 2003a, p. 11). Yet the 2004 OPCS Progress Report (which took the place of the Monitoring and Evaluation Progress Report) does not present such an architecture.

The Bank's yearly unit compact exercise informs the allocation of administrative resources across Bank units. Unit compacts have recently attempted—with varying degrees of success—to include outcome and intermediate outcome indicators and links between deliverables and outcomes. But the development of meaningful, comparable, and monitorable performance indicators and the systematic linking of deliverables and outcomes, remain works in progress. The link between the unit compacts and other processes that are part of the results agenda, such as the



results-based CAS, also needs strengthening. Linking results to resource allocation requires systematic results reporting that is not yet in place at the Bank. Bank management says that, in the future, it will move the unit compact exercise away from focusing corporate-level discussions on detailed plans for deliverables, and toward links among strategic objectives, performance, outcomes, and resource allocation. The experience of other agencies in linking performance information to resource allocation (see box 2.2) should be reviewed to draw lessons for the Bank.

### **Resolving the Tension among Attribution, Contribution, and Accountability**

The Bank has emphasized that improved country outcomes for sustainable growth and poverty reduction are its bottom-line measure of development effectiveness (World Bank 2002). While such outcomes emerge only gradually and are a result of multiple influences, aid organizations can contribute to them (World Bank 2002). International experi-

ence points out that as aid organizations shift toward measuring country or sectorwide outcomes, rather than project outcomes, and emphasize partnerships with other donors, it becomes increasingly difficult—theoretically and practically—to attribute development effects to individual actors and factors (White 2004a, b, c). The challenge is to ensure that the focus on contribution and shared responsibility does not compromise accountability. The Bank recognizes that even though narrow attribution is often not possible, the relevance and effectiveness of the contributions can and should be monitored and evaluated (World Bank 2002).

The Bank has attempted, but not yet succeeded, in identifying measures of Bank performance and of the Bank contribution to development outcomes for which staff can be held accountable. The Bank tried to develop a Corporate Scorecard, which identified measures of organizational effectiveness (mainly inputs) directly within the Bank's control and for which staff could be held accountable, and measures

#### **Box 2.2**

#### **Linking Performance Information to Resource Allocation**

**Chile's Ministry of Finance uses performance information (based on process and output indicators of efficiency and effectiveness and desk/impact evaluations and comprehensive spending reviews) in the budget cycle as a starting point for discussions with agencies. It does not mechanically cut funding to programs that fail to meet targets or increase the funding of programs that achieve results. But when results are poor, it pushes agencies to improve performance. New performance conditions are established and incorporated into the Budget Law, and/or agencies agree to modifications in their programs. These promises are reviewed in subsequent years to see if performance has improved. In cases of continuing failure to improve, funding to programs has been reduced. In a few instances, programs have been eliminated.**

**The Montgomery County, Maryland, program, Montgomery Measures Up!, is designed to help make performance meas-**

**ures and performance measurement an integral part of the way the County reports its progress, manages its programs, and allocates its resources. Performance data are an important—but not decisive—input into budgetary decisions. These data, including information on inputs, outputs, efficiency, service quality, and service outcomes, must be submitted with the budget. All requests for additional resources and for the restoration of budget cuts must identify the results being sought and how they will be measured. Budget analysts factor the performance data into their recommendations on resource allocation. Some departments are using performance data to manage services. For instance, the Department of Health and Human Services monitors program results/outcomes quarterly and includes results-based performance measures in its contracts with service providers.**

*Note:* Management notes that the governments of Chile and Montgomery County deliver services directly to users, while the Bank delivers support to country governments and not directly to service users.

*Sources:* Chile 2003. For Montgomery County: interviews with Montgomery County government staff, 2004 (John Greiner, Office of Management and Budget; Arleen Rogan and Susan Seling, Department of Health and Human Services).

of development effectiveness (outputs and outcomes) not fully within the Bank's control, but to which the Bank contributes. This effort has been ended, and no review has been undertaken to document and learn from the experience.

In the context of the IDA14 replenishment, Bank management has proposed a results measurement system that includes both indicators of country outcomes (Tier I) and indicators of IDA's performance (Tier II). Table 2.2 illustrates selected indicators from the Corporate Scorecard and the proposed IDA14 results measurement system. Some key issues remain to be addressed. It is difficult but critical to trace causal links relevant to the achievement of MDGs and related outcomes and to ensure that Bank policies and programs contribute effectively through these causal chains. The links between country outcome indicators and Bank performance indicators will need to be verified. Additional indicators, including qualitative measures, will need to be included to provide a more complete picture of Bank performance, especially because outcome data from results-based CASs will not be available for a few years.<sup>4</sup>

Overall, the slow progress in adopting measures of organizational effectiveness and Bank performance and defining more precisely how the Bank's contribution to development outcomes would be measured contributes to a lack of clarity about what staff are held accountable for.

### Changing the Organizational Culture, Incentives, and Structure

The lesson emerging from international experience is that managing for results can be achieved only with profound changes of organizational culture and incentives, and that changing mental models is the central challenge (see box 2.3). Aid organizations must shift toward creating an organizational culture that values lesson learning and an incentive framework that promotes managing for results.

*Better Measuring, Monitoring, and Managing for Development Results* (World Bank 2003a) observed that "considerable proactivity is needed to ensure a consistent and coherent approach and to correct disincentives that staff may perceive in current human resources and budget policies" and indicated that staff incentives would be reviewed and revised as

**Table 2.2**

### Selected Measures of Development and Organizational Effectiveness

Measures	Corporate Scorecard	Proposed IDA14 indicators
Development effectiveness	<ul style="list-style-type: none"> <li>• Percentage of people living on less than \$1 a day</li> <li>• Universal net primary enrollment rate (%)</li> <li>• Under-5 mortality rates (per 1,000 live births)</li> </ul>	<ul style="list-style-type: none"> <li>• Proportion of population below \$1/day poverty line</li> <li>• Ratio of girls to boys in primary and secondary education</li> <li>• Under-5 mortality rate</li> <li>• Sustainable access to an improved water source</li> <li>• Time required for business start-up</li> </ul>
Organizational effectiveness	<ul style="list-style-type: none"> <li>• CAS initial quality</li> <li>• CAS implementation/impact quality</li> <li>• Focus on Bank mission</li> <li>• Partnership/communication</li> <li>• Lending approvals (\$b committed and number)</li> <li>• Nonlending services volume (\$m budget and number)</li> <li>• Quality at entry; quality of supervision</li> <li>• Staff diversity index</li> </ul>	<ul style="list-style-type: none"> <li>• The number of countries that use a results-based CAS</li> <li>• CAS outcome ratings in CAS Completion Reports (CASCRs) as validated by OED through CASCR Reviews</li> <li>• Quality-at-entry indicators for IDA projects as assessed by QAG</li> </ul>

Sources: World Bank 2003c.

**Box 2.3****Changing Organizational Incentives:  
The Central Challenge**

International experience shows that organizations typically focus more attention on the methodological and technical aspects of results-oriented M&E systems than on changing mental models. A central challenge is to change incentives for staff and managers to focus on results, while simultaneously avoiding the risk of introducing perverse incentives. Perverse incentives appear when:

- A “checklist mentality” is created.
- Meeting indicators—as opposed to creating results—becomes the dominant goal.
- “Quick wins” dominate long-term sustainability.
- To ensure goal attainment, ambitions are set too low.
- The best people and the most resources are assigned to the best-performing activities rather than those that perform poorly and need resources the most to improve performance.
- Managers have so little real control over outputs that setting targets becomes a purely symbolic and ritual exercise, with later avoidance of reporting on results versus plan, or another ritual game of distributing the blame for failure.

To avoid such adverse behavior, the focus on measuring and monitoring results must be accompanied by clarity on the intended results and a wide understanding among staff of the rationale for measuring and monitoring. This will also help to avoid creating masses of irrelevant M&E data that are never used. The incentive system must encourage staff to implement the results focus in spirit, rather than encouraging them to see it as one more bureaucratic hassle to circumvent by cutting corners, or as an opportunity for spending resources indiscriminately to collect M&E data for its own sake.

*Sources:* Maxwell, forthcoming; Meier 2003; UNDP 2002; Binnedijk 2001; Schick 1998.

necessary in fiscal year 2003 to reflect enhanced results focus. However, such a review has yet to be undertaken. In reporting on the M&E Improvement Program, the 2004 OPCS Progress Report recognized that inappropriate incentives were a valid predictor of the challenges the program faced (World Bank 2004, p. 49). Yet it did not recommend urgent action, but rather treated it as “part of a long-term reform agenda” (World Bank 2004, p. 3). The terms of reference

for the recently established Organizational Effectiveness Task Force mention incentives but are silent on whether the Task Force will specifically address incentives related to achieving, monitoring, and managing for results. Even though Bank documents acknowledge the importance of getting the incentives right in order for the results agenda to succeed, management still needs to develop time-bound, actionable measures to address the issue.





# M&E in Bank Guidelines for Lending and Nonlending Products

**T**wo kinds of changes can help an organization move effectively toward a results-oriented M&E system: (i) changes in systems and procedures, notably staff guidelines and product templates; and (ii) changes in organizational culture, incentives, and structure. This chapter focuses on staff guidelines and product templates; organizational issues are discussed in Chapter 4.

Staff guidelines embody an organization's approach to specific issues and can help promote the systematic adoption of the intended approach across the organization. This chapter assesses how well the Bank's guidelines and templates guide staff in addressing the three key criteria for results-oriented M&E presented in Chapter 1. Box 3.1 discusses the M&E responsibilities of the Bank. For each product, available guidelines for product initiation, implementation, and completion were reviewed. Corporate assessments and QAG assessments, which embody Bank standards, were also reviewed. Annex A lists the products and items reviewed.

## **For Country Products, Guidelines Are Going in the Right Direction**

The Bank's main (and, until recently, only) tools to specify country program objectives and the M&E systems that track their achievement have been the Country Assistance Strategy (CAS) and the CAS Progress Reports.

As part of the Bank's move to a results focus, the traditional CAS<sup>1</sup> is being replaced by a results-based CAS, and the CAS Completion Report (CASCR) has been introduced on a pilot basis. This section assesses the extent to which the guidelines and templates for these country products have incorporated results-oriented M&E.

**Results-Based CASs.** Although no formal guidelines have yet been issued for the recently introduced pilot program of results-based CASs, a 2003 internal review of the CAS offered guidance and a template for a results-based CAS.<sup>2</sup> It offers instructions on linking results to country goals and provides an outline of a results framework with a sample matrix that defines the questions that should be answered. The matrix (see Annex B) is well designed, although it could be improved by requiring more explicit links between a Bank intervention, actions to achieve intermediate outcomes, and intermediate outcome indicators. The

**Box 3.1****M & E: Whose Responsibility?**

Clarity on roles and responsibilities is critical for the effective implementation of the M&E system and for the ownership and use of the resulting information. The borrower is responsible for monitoring of borrower objectives (for example, those expressed in the Poverty Reduction Strategy Paper, or PRSP). While it will be necessary for the Bank to rely on borrower M&E systems, monitoring the impact of Bank objectives (such as those expressed in the CAS) and measuring how Bank instruments support CAS objectives are Bank responsibilities.

For monitoring indicators fully under the Bank's control (for example, input and process indicators or measures of organizational effectiveness), the Bank will have to generate the necessary data. For monitoring specific process and output indicators and all outcome indicators relating to Bank objectives, the Bank's job will be easier if borrower M&E systems are good.

When the borrower M&E systems are poor and the borrower wants to improve them, the Bank, in cooperation with other development partners, can support borrower capacity development,

which will help the government monitor its own (PRSP) objectives, while at the same time facilitating the Bank's monitoring of the CAS objectives. Bank support should not encourage countries to set up enclave M&E systems that respond only to the Bank's own data needs, but should focus instead on achieving systemic capacity enhancements. The Bank should draw on borrower M&E data while also checking data quality during supervision, thus signaling its seriousness about M&E and, in the process, identifying the need for any additional support.

If the borrower systems are poor and there is little appetite for improving them, the Bank will find it difficult to monitor its CAS and other objectives. The Bank cannot substitute for the measurement of country results by the borrower. One option in such cases might be for the Bank to reconsider its engagement with the country until it is better able to measure the effectiveness of its own support. While limiting lending, the Bank could still continue efforts to enhance the country's interest in monitoring and evaluating results.

*Source:* Background work, 2004 AROE.

importance of well-designed indicators and the problems that weak government M&E capacities pose for implementing an M&E system have been recognized, but no specific advice has yet been offered. OPCS is reviewing the results-based CAS pilot and expects to deliver its report to the Board of Executive Directors in fiscal year 2005. Based on this review and the Board discussion, OPCS plans to issue guidelines in 2005.

**CAS Completion Reports (CASCRs).** The CASCR is the first systematic attempt at self-evaluation of country assistance programs. The CASCR template can be improved by requiring a clearer link to be made between the development objectives of lending/nonlending activities and their contribution to one or more of the CAS outcomes. A further improvement would be to require an assessment of the design and implementation of the CAS's M&E system and of the client's M&E capacity. In cases of weak client M&E capacity, the CASCR could point to a number of instruments that are now available to help build client M&E capacity (see box 3.2).<sup>3</sup>

The implicit guidelines for the new results-based CAS and the explicit guidelines for the CASCR are a major improvement over the earlier guidelines for (traditional) CASs. CASs using the new guidelines show that greater results orientation—additional guidance on the results framework and availability of a CASCR—can improve the CASs' results framework and M&E systems (see table 3.1). It is important for the proposed CAS guidelines governing them to be “living” guidelines, updated periodically to reflect additional learning and experience. Speedy extension of CAS completion reporting to all countries would allow the benefits from the new approach to be disseminated quickly.

### **For Sector Products, Guidelines Are Silent on Results**

The networks' main tools to set sector objectives and to track their achievement are the Sector Strategy Papers (SSPs) and Implementation Progress Reports (IPRs), respectively. Until recently, guidance to staff for SSPs and IPRs has been scanty. There are no

**Box 3.2****Building Client M&E Capacity****Statistical Capacity Development**

The Statistical Capacity Building Program (STATCAP) initiative uses adaptable program loans to help countries build statistical capacity. The cornerstone is the development of a Statistical Master Plan (financed through trust funds) that contains components to improve statistical policy and the regulatory and institutional framework, statistical infrastructure, statistical operations and procedures, human resource development, and physical infrastructure and equipment. So far, STATCAP programs have been launched in Burkina Faso and Ukraine. The programs include results frameworks that specify project development objectives, outcome indicators, and use of results information. If the STATCAP initiative is to succeed, it will be important to go beyond a technocratic approach to also address issues such as staff turnover in statistical agencies and the demand for M&E data among policymakers. The danger of generating masses of irrelevant or unused M&E data should be avoided by working closely with policymakers. The Bank needs to ensure that the STATCAP initiative is sufficiently different so that it is able to avoid the fate of past data efforts—for example, the

*Sources:* STATCAP 2004; OED 2002; background work, 2004 AROE.

*Social Dimensions of Adjustment Initiative for Africa, which did not generate the anticipated data strengthening and had to be abandoned after incurring much cost.*

**Evaluation Capacity Development**

The Bank's strategy for evaluation capacity development (ECD) was established in 1994 to help borrowers strengthen their M&E systems as an integral part of sound governance. ECD aims to improve M&E capacities in country institutions concerned with the monitoring, evaluation, and reform of public sector activities and programs—governments, civil society, universities, research institutes, think tanks, NGOs, private consulting companies, individuals, and so on. The Bank is currently supporting ECD in about 31 countries. A 2002 OED review of these efforts pointed out the limited progress in mainstreaming evaluation capacity building into Bank operations, despite the recommendation of two Bank Task Forces (in 1994 and 2000) that mainstreaming be pursued. Greater emphasis will be needed both on developing country capacities for measuring and monitoring results and for managing for results.

written guidelines for the SSP concept paper, and the written guidance on SSPs themselves is limited to a single-page suggested outline, which does not refer to a focus on outcomes or to a results orientation. The outline mentions performance monitoring and evaluation, but does not provide any guidance on this aspect. QAG recently reviewed SSPs and found that three out of four mainline sector strategies and 40 percent of the other strategies were rated less than satisfactory: they were vague, lacking both in selectivity and in practical operational guidance. An SSP Help Desk has been established in OPCS to provide guidance to SSP teams that are planning to engage in the preparation of new SSPs.

The Bank has not clearly established what sectors and networks are accountable for and how SSPs and CASs interact to produce results. In 2003, in an internal assessment, OPCS concluded that within a country-driven and results-focused business model, the main thrust of future sector strategy work should be

on implementation at the country level, with particular attention to bridging the gaps between sector priorities and country-program priorities. It has not, however, spelled out how this is to be achieved, how cross-regional sector strategies can deal adequately with varying country conditions and country priorities, and how adequate CAS-SSP linkage can be assured.

The Bank has not yet carried out systematic and regular self-evaluation of SSPs. Some IPRs aimed at reviewing SSP implementation have taken the form of Topical Briefings with little or no evaluative or retrospective content. Yet SSP self-evaluation could be an effective vehicle for assessing how well the Bank is succeeding in helping countries reach the MDGs.<sup>4</sup> Sector-by-sector IPRs are now to be replaced by a single Sector Strategy Implementation Update (SSIU) covering all sectors. An SSIU Help Desk has been set up and some guidance and a template produced. The first SSIU is slated for Board delivery in early fiscal year 2005. It remains to be seen how well a single report can capture

**Table 3.1 Results-Based CAS Adds Value**

Element	Cameroon and Honduras CASs	Cameroon CAS (part of the results-based CAS pilot)	Honduras CAS (not part of the results-based CAS pilot)
Results framework	Both CASs have a results framework based on the PRSPs and a "Country Program Matrix" setting out objectives, performance indicators, and Bank interventions.	The results framework is more explicit, distinguishing between strategic country objectives (for example, maintain strong macro performance) and CAS outcomes that the Bank expects to influence (such as annual budgets drawn up on the basis of the Medium-Term Expenditure Framework (MTEF)), as well as intermediate indicators to track implementation toward outcomes (for example, integration of MTEF into public finance law by 2005).	Results framework does not make explicit distinction between country and Bank (CAS) objectives.
Monitoring and evaluation	Both CASs discuss M&E relating to country PRSP and MDG objectives and Bank (CAS) objectives.	The results-based CAS shows the advantages, for M&E, of a more explicit results framework, linking the CAS outcomes and intermediate indicators to be monitored to the specific PRSP priorities (for instance, the PRSP objective of reducing the incidence of urban poverty leads to specific tracking of percentage of residents with access to water and sanitation and of city neighborhoods with access to functioning roads).	PRSP targets and MDGs are monitored but not directly linked to Bank (CAS) objectives. No intermediate indicators to track implementation toward expected CAS outcomes.
CAS Completion Report		The CASCR for Cameroon helped the CAS to define its results framework by systematically assessing achievements and shortcomings during the previous period and pointing to areas requiring improved M&E. For example, on public financing and service delivery, the CASCR notes: "budgetary reallocation in favor of priority sectors and regions became meaningless when it proved impossible to verify actual expenditures." The new CAS, therefore, emphasizes evaluation through budget tracking.	No CASCR available for Honduras. Brief discussion of what worked and what did not work.

Sources: Cameroon CAS (World Bank 2003d); and Honduras CAS (World Bank 2003e). The Honduras CAS was chosen as the comparator since it was a non-results-based CAS done at roughly the same time as the results-based Cameroon CAS, to illustrate how a results-based CAS differs from a non-results-based one.

implementation progress in the many sectors where the Bank works and how thoroughly it assesses the extent to which the Bank is managing for results in the various sectors.

### For Lending, Guidelines Are a Work in Progress

**Investment Lending.** Given the Bank's long involvement in investment projects, results-oriented M&E has received greater attention in guidelines at this level than at other levels. The current guidelines for the Project Concept Note provide detailed guidance for specifying

project development objectives.<sup>5</sup> They require staff to identify the principal outcome for the primary target group and analyze the exogenous factors that could prevent achievement of the objectives. They also require an explanation of how the proposed operation supports the borrower's objectives, policies, and strategies in the sector and how the proposed operation supports the relevant CAS objectives.

For Project Appraisal Documents (PADs), the current guidelines/template incorporate the main elements of results-oriented M&E, including a clear specification of project



development objectives, expected timing of results, exogenous factors that threaten their achievement, and arrangements for collecting and using M&E data for decisionmaking. A quality enhancement review, to be overseen by the sector director or manager after the project concept has matured into a project design, is required for all projects<sup>6</sup> and is supposed to cover the project’s results framework.<sup>7</sup>

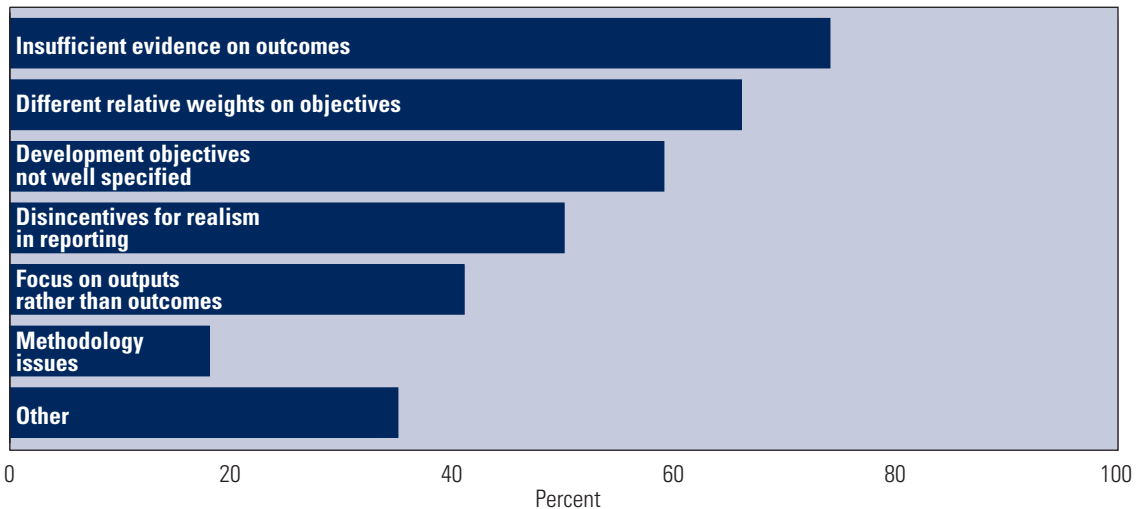
The template for Project Status Reports (PSRs) is being revised by OPCS for launch in fiscal year 2005. The May 2004 version is a significant improvement over the old PSR.

The guidelines and templates for Implementation Completion Reports (ICRs) should—but do not—adequately emphasize the need for ICRs to assess M&E and provide sound outcome data. The absence of outcome information in ICRs often produces a disconnect between OED and ICR ratings. In the Europe and Central Asia (ECA) Region, for example, over fiscal years 2001–03, the most common reason for a disconnect between the Regional and OED outcome ratings was insufficient evidence on outcomes, as shown in figure 3.1. Several OED ICR Reviews highlight the lack of adequate outcome data in ICRs (see box 3.3).

Adaptable Program Loans (APLs) can be key to the results agenda in that they promote adaptation and learning, but more guidance is needed, especially for preparation of subsequent phases of APLs. At completion, existing ICR guidelines are applied to APLs. These guidelines may need to be adapted for APLs to reflect their special characteristics.

QAG recently added questions about the results focus and M&E of Bank projects in its quality-at-entry and quality-of-supervision questionnaires—for example: “Is the results focus of the project design consistent with the outcomes specified by the task team?” (QAG 2004). But these questions need to be made more specific and guidance provided to panelists, because the term “results focus” may be subject to multiple interpretations. Also, QAG questionnaires do not require panelists to adequately examine the projects’ M&E systems. Recently, QAG further revised its quality-of-supervision questionnaires and plans to extend similar improvements to its other assessments. The full set of QAG questionnaires (and guidance to panelists) should include the three key criteria for results-oriented M&E: results framework, design of M&E, and implementation of M&E.

**Figure 3.1** Insufficient Evidence on Outcome Is Leading Reason for Disconnect in ECA



Source: ECA Region quality at exit data, 2003.

**Box 3.3****ICRs Lack Information on Results**

OED's review of the ICR for the Uganda Sexually Transmitted Infections Project found that the ICR had not provided sufficient output data (let alone outcome information) and had failed to demonstrate attribution or contribution. The ICR had not provided sufficient information on project outputs (aside from the number of condoms procured) to account for the use of funds, to substantiate claims that the project had a large influence on the national program, or to link the national program's (and project's) outputs to changes in behavior or HIV.

Although one of the main project objectives was to improve accessibility of water supply and sanitation services for the

Source: OED data.

urban poor, OED's examination of the Sierra Leone Urban Water Supply Project found that the project components were not specifically targeted toward the poor, and indicators for monitoring and evaluating impact were inadequate. The Staff Appraisal Report and ICR had assumed that the increased water supply to Freetown as a whole would improve accessibility for the poor, which may not be the case. Project benefits may be disproportionately enjoyed by those who are better off and able to afford water connections.

**Adjustment Lending.** There is currently no template comparable to the PAD for general adjustment operations. Guidelines for Poverty Reduction Support Credits (PRSCs) organize the results focus around the Poverty Reduction Strategy Paper (PRSP) and the Joint Staff Assessment, but provide no specific guidance on M&E. Guidelines for Programmatic Adjustment Loans (PALs) include little on results and M&E. The practice of segmenting support for complex and difficult institutional reforms into annual tranches, as is done in multiyear PRSCs and PALs, seems to encourage the counting of small, interim procedural steps as measures of progress. Even though results may not be immediately observable, greater effort is needed to identify proxies for results and to include these as triggers. OPCS plans to launch guidelines/templates for "traditional" and programmatic adjustment operations by the end of fiscal year 2004. It will be important that they ensure the needed shift away from a focus on processes and outputs to outcomes.

### **In the Bank's Analytical Products, M&E Is Still Rare**

The Bank's economic and sector work (ESW)/analytical and advisory activities (AAA)<sup>8</sup> seek to achieve results indirectly, by influencing policymaking, facilitating consensus building, or producing core diagnostics for the Bank.<sup>9</sup> This makes it difficult to apply results-based M&E and the Bank's accountability for

results; in consequence, the systematic measurement of results has not been attempted for the majority of the Bank's ESW/AAA products. Yet in order to ensure managing for results, it is important to be explicit about which results the ESW/AAA is seeking, to design and implement the ESW/AAA accordingly, and also to design and implement M&E systems to track the extent to which the intended results are materializing.

There are no Bankwide corporate guidelines for ESW, and most of the Regional guidelines are process-oriented. The networks have produced guidance for major task categories (for example, poverty assessments). Several corporate assessments of ESW have been undertaken recently as part of the "Fixing ESW" initiative. Papers prepared for this initiative addressed management tracking of ESW, but did not discuss how to improve (or establish) results frameworks or M&E for ESW.

The Bank's 1998 ESW governance framework introduced the Activity Completion Summary (ACS). The ACS was to be a self-evaluation tool for the task team, including a statement of task objectives and the extent to which these were achieved, the dissemination process and activities, and any lessons learned. The recent QAG review of fiscal year 2002 ESW found that only a handful of ACSs contained findings relevant to the achievement of success indicators, and noted that the ACS was "highly problematic" for task reporting and evaluation (QAG 2003). It

concluded that “the relatively weak results orientation in ESW task management is reflected in the scattered and incomplete use of the ACS to record task evaluation” and recommended that either meaningful incentives for task teams and managers to comply with ACS requirements be introduced, or that the ACS should be reconsidered. OPCS has recently revised the ACS template in SAP (Systems, Applications, and Products, a system used by the Bank) to link development objectives more closely with actual development results. But since the template itself was not the main problem, it is doubtful how much of a difference this revision will make.<sup>10</sup>

The Bank’s focus on results has found a promising echo in some new trends in ESW (and AAA) work. These activities—three are described in box 3.4—tend to combine a variety of tools, often cover multiple years, include more detailed setting of objectives (with different degrees of results frameworks), and incorporate more explicit M&E systems. A Poverty Reduction and Economic Management (PREM) Network note recommends the Bank’s traditional project cycle as a model for ESW and AAA, emphasizing the steps going from identi-

cation, preparation, and appraisal to implementation, supervision, and evaluation. In the AAA cycle, the latter steps represent a new addition, often missing in traditional analytical work, which would end with the preparation (and perhaps discussion and dissemination) of a Bank ESW report. As learning from these new initiatives increases, it should be codified in guidance to staff. So far, only the Middle East and North Africa Region (MNA) has prepared guidelines for this new approach (together with a Learning Exercise to disseminate lessons and modify the guidelines as learning takes place). The Bank’s research group, Development Economics (DEC), conducts analytical work and impact evaluations and could have a role in helping to increase the results focus of the Bank’s ESW/AAA.

In conclusion, the Bank is currently revising or introducing guidelines and templates for a number of its products. But these tools for each of the Bank’s product lines have gaps with respect to a results framework, design of M&E, or implementation of M&E, in initiation, implementation, or completion. Faster progress is especially needed for the majority of the Bank’s ESW/AAA products, for which the

### Box 3.4

### Results-Oriented ESW/AAA: Some Examples

The MNA Region’s Programmatic ESW (PESW) requires staff to define the objectives of the activity, agree on results and performance indicators, and secure the client’s commitment. The pilot program of PESW involves five activities in Iran, Morocco, and Tunisia. In Iran, for example, the Public Expenditure Review includes several intermediate outputs, such as joint analytical pieces on the budget process and a comprehensive set of training and learning activities for Iranian agencies—all expected to help achieve the objective of improving Iran’s budget process.

The Russia Poverty Program supports an improved methodology for poverty measurement, an improved Household Budget Survey, and a system of administrative data statistics for enhanced poverty monitoring and policy-oriented analytical work. Intermediate output indicators that will be monitored include two

poverty assessments. Close partnership with the government is expected to ensure open access to household-level data sets, a major objective of the program, which would help enhance the policy debate on poverty, development, and the reform agenda.

In Thailand, the CAS proposed the “Country Development Partnership,” rather than lending, as its main instrument. Country Development Partnerships set out medium-term integrated frameworks for reform programs to achieve specific results. Intermediate outputs include pension reform modeling, analyses of the financial sustainability of unemployment insurance, and a risk and vulnerability assessment. The Country Development Partnerships for social protection, supported by the Bank and other partners, include expected outcomes of reducing gaps in equity, coverage, and efficiency of the government’s social protection programs.

Sources: MNA Region internal PESW guidelines, 2003; World Bank 2003g.

systematic measurement of results has not yet been attempted, and at the sector level. The priority at the sector level is to ensure that the proposed Sector Strategy Implementation

Update (SSIU) emerges as a systematic self-evaluation tool for sector programs and assesses the extent of results-oriented M&E and managing for results at the sector level.



# Institutional and Organizational Factors Affecting the Bank's Focus on M&E and Managing for Results

Changes in guidelines and templates are important, but not sufficient, for ensuring a systematic adoption of a results-oriented approach in an organization. International experience with managing for results suggests that organizations are quick to make procedural and technical changes in response to the adoption of a results focus, but critical changes in organizational culture and incentives are the central challenge (Chapter 2).

This chapter examines the extent to which staff and managerial incentives, M&E systems for decisionmaking, and organizational structure are aligned to encourage results-oriented M&E and managing for results. The analysis is based primarily on semistructured interviews of 34 Bank staff and managers across the institution.<sup>1</sup> Annex C presents the Interviewee Matrix and the Interview Guide.

## **Traditional Approach Versus a Results-Oriented Approach**

The challenge of transforming the Bank into a more results-oriented organization is to move it from its current or “traditional” organizational culture, incentives, and structure, to a new, results-focused approach to decisionmaking and management. Table 4.1 shows the three key areas where change will be needed, and

subsequent sections discuss the current status of the Bank in dealing with each of those areas.

## **Staff and Managerial Incentives in the Bank Do Not Fully Support the Adoption of the Results Agenda**

Are incentives in the Bank sufficiently aligned with its results focus? While shareholder expectations may be creating positive incentives for the Bank to adopt results-oriented M&E, many staff and managerial incentives do not fully support the move toward results-oriented M&E and managing for results—in some cases, they inhibit it.

**Shareholder Expectations.** Within the Bank, a major incentive for intensifying the effort to manage for results has been strong and sustained pressure from the Bank's sharehold-

Table 4.1

## Changing to a Results-Oriented Bank

Decision type	Traditional approach	Results-oriented approach
Incentives, disincentives, and barriers	<ul style="list-style-type: none"> <li>• Staff are rewarded largely for delivering projects through the approval system, and expertise and task/program delivery are the paramount determinants of managerial selection.</li> <li>• Staff avoid taking risks and feel pressure to convey “good news.”</li> <li>• Regional, country, sector, and task budgets are derived from a combination of task numbers and unit costs, with overall allocations indirectly reflecting such variables as country performance (CPIA), which are now an explicit determinant for IDA country allocations.</li> </ul>	<ul style="list-style-type: none"> <li>• Staff and managers rewarded for achieving results on the ground, and managers for encouraging, recognizing, and rewarding collaborative cross-sectoral work.</li> <li>• Staff encouraged to experiment, and learn from mistakes, are not afraid to convey “bad news,” and are rewarded for systematically learning from past failures.</li> <li>• Budget allocation process is informed by data on progress toward intended results.</li> </ul>
Information systems for decisionmaking	<ul style="list-style-type: none"> <li>• Managerial decisions on goals and operational choices are based on information about inputs, including client requests, staff proposals and analyses, CPIAs, budget information, output status reports, and media reports. Lack of systematic tracking and reporting of outcome indicators.</li> <li>• Decisionmaking is based on what data are available—typically, input and output data.</li> </ul>	<ul style="list-style-type: none"> <li>• More emphasis on clearly specifying intended results and ensuring that managers have relevant and reliable M&amp;E information (especially on progress in achieving intended results) on which to base their decisionmaking.</li> <li>• Decisionmaking based on a variety of M&amp;E data (input, output, and outcome data, and also project-, sector-, and country-level data).</li> </ul>
Organizational structure	<ul style="list-style-type: none"> <li>• Too many groupings (for example, sectors/networks/sector boards), which makes cross-sectoral work difficult.</li> <li>• Proliferation of tasks born out of competition for budget resources and task leadership.</li> </ul>	<ul style="list-style-type: none"> <li>• Organizational structure facilitates cross-sectoral work.</li> <li>• Tasks and budgets matched to well-defined results.</li> </ul>

Source: Background work, 2004 AROE.

ers, particularly the IDA (International Development Association) donors. Beginning with IDA13, donors have expected IDA to report on selected aggregate development outcomes and on the delivery of the analytical work that underpins IDA's assistance to borrowers. A more comprehensive results-measurement system has been developed for use during IDA14. It includes an expanded set of measures of country outcomes, as well as indicators of IDA's contribution to them.

Within IDA recipient countries, the Bank now uses a country's rating on the Country Policy and Institutional Assessments (CPIAs) in its dialogue with country authorities. The authorities realize that their IDA allocation is linked to their policies as measured by the

CPIA. This has created an incentive, given the attractiveness of IDA funding, for countries to pay attention to their country policies and institutional environment. CPIAs currently measure policies and institutions, which some Bank analytic work shows are positively correlated with development outcomes. More explicitly introducing an element of results orientation into the resource allocation mechanism would give borrowers additional incentives to deliver results.

**Management Emphasis and Signals.** The renewed results initiative the Bank launched in September 2002 has raised staff awareness of results orientation. The Bank's top management has sent strong signals regarding the results

focus—notably through the “Results: Everybody’s Business” event for Bank staff held on January 7, 2004—which have been observed by virtually all interviewees. Interviewees were unsure, however, if management will persevere with the results initiative long enough for it to bear fruit, as well as about what the high-level corporate messages meant for their day-to-day work. They reported a lack of operationally relevant guidance from middle management about the implications of these messages. While increasing lending is not necessarily inconsistent with the results focus, interviewees observed that recent corporate messages on increasing IBRD lending<sup>2</sup> and the Infrastructure Initiative could compete with the focus on outcomes, and they reported receiving little guidance on how increased lending could be reconciled with the focus on results and outcomes. To reduce this risk, management should ensure that the right incentives are in place and clarify the accountability of staff at each level. The experience of the British Department for International Development (DFID), summarized in box 4.1, illustrates one way of doing this. Interviewees pointed out that they may face negative consequences for slow disbursements, but not for failing to achieve outcomes or for failing to distill and act on lessons of experience. Interviewees perceived that the Bank’s current incentive system in the project cycle primarily rewards

the early part of the project cycle. But an incentive system for results would reward the later stages more (see table 4.2). Changing this incentive system will be critical for encouraging results-oriented staff behavior.

Incentives for M&E have traditionally been weak in the Bank because of the lack of a learning culture. In a learning culture, mistakes are regarded as opportunities for learning. Staff are rewarded for learning from past mistakes and incorporating lessons of experience into future work. Interviewees said that they felt discouraged from monitoring results when things had not gone quite as anticipated. Interviewees also pointed out the need to create an organizational culture that is accepting and transparent about errors and shortcomings. OED’s *Knowledge Sharing* study found that there are no routine practices for validating lessons learned and promising practices—only a third of 15 surveyed network and Regional units reported having lessons learned or good practice validation processes in place (OED 2003c). The monitoring process should be designed to yield information that provides a sound and continuing basis for informed decisionmaking and learning (OED 2003b).

**Reward Systems.** Interviewees observed that management should reinforce its message on managing for results by recognizing staff who respond to the message. Such recognition could

#### Box 4.1

### Managing for Results in DFID: Linking Personal and Corporate Objectives

DFID sets measurable targets for itself using the U.K. Government system of Public Service Agreements (PSAs). The targets are a mix of outcome measures (related to the MDGs) and measures of DFID’s organizational effectiveness. The recent PSA was launched to staff with strong senior management support and an extensive, interactive communications strategy that drew clear links between each individual’s job and relevant PSA targets. Special efforts were made to clarify links for individuals working in support functions or on thematic or country teams that were excluded from the shortlist of “headline” targets. DFID has

mapped PSA targets directly onto its internal organizational structure. A director general/director was made directly accountable for each individual PSA target and for writing a delivery plan for them. A clear line of sight was created between every individual’s personal objectives, those of the team and department, and the PSA targets. Strategic moves to secure staff ownership of the PSA through clear communication, transparent and participatory design and review processes, and explicit linking of personal and corporate objectives, have been critical to its acceptance within the organization.

Source: Background work, 2004 AROE.

**Table 4.2****An Incentive System for Results Would Reward the Later Stages More**

<b>Current Incentive System</b>					
<b>Identification</b>	<b>Appraisal</b>	<b>Board approval</b>	<b>Signing and effectiveness</b>	<b>Supervision and monitoring</b>	<b>Evaluation</b>
<b>High</b>	<b>High</b>	<b>Very high</b>	<b>Medium</b>	<b>Medium</b>	<b>Low</b>
Incentive to identify new, self-standing projects	Matters for career advancement	Premium on loan approval for career advancement	Limited management attention, except when effectiveness is greatly delayed	Management attention to (and staff incentives for) supervision have increased but are still limited and focused on disbursement and implementation, not results	Evaluation often considered a junior staff or consultant task. Appraisal staff often re-assigned before evaluation
<b>Results-Oriented Incentive System</b>					
<b>Identification</b>	<b>Appraisal</b>	<b>Board approval</b>	<b>Signing and effectiveness</b>	<b>Supervision and monitoring</b>	<b>Evaluation</b>
<b>Medium</b>	<b>High</b>	<b>High</b>	<b>Very high</b>	<b>Very high</b>	<b>Very high</b>
Emphasis on contribution to CAS objectives through all means (not limited to new, self-standing projects)	Should be on par with achievement of results	Should be high as an important intermediate step for achieving results, not as a result itself	Management attention to speedy effectiveness should very be high	Project performance during implementation and quality of supervision should be major factors in career advancement	Monitoring of project performance during implementation and evaluation of results at closure should be important factors in career advancement

Source: Background work, 2004 AROE.

include private and public praise and feedback, professional advancement, career promotions, awards, pay increases, and bonuses. The limited effectiveness of the Bank's present performance evaluation and reward system is illustrated by the Staff Survey results: only 44 percent agree that staff are rewarded according to their job performance, and only 48 percent agree that when things go well in the job, their contribution is often recognized. Interviewees considered that getting results is not part of the current reward system and deserves greater emphasis. At the same time, there was some concern that the Bank may become too mechanistic and unrealistic in trying

to link country or project outcomes to reward or budgetary systems. Some new approaches are being tried in other organizations (see box 4.2) and the International Finance Corporation (IFC) has recently introduced a system of bonuses to reward staff. While these approaches continue to be developed and refined, experience with them should be reviewed to glean lessons for the Bank.

**Lack of Clarity about Additional Mandate.** The Staff Survey also illustrates the problem of overloaded staff: only 55 percent agree that work pressures are at acceptable levels, and a significant 29 percent have experienced



**Box 4.2****Pay Based on Results**

**DFID.** The U.K.'s DFID has introduced links between its senior management's pay and progress made toward Public Service Agreement (PSA) targets (for example, increase the proportion of DFID's bilateral program going to low-income countries from 78 to 90 percent, or progress toward the Millennium Development Goals in 16 key countries in Africa). The bonuses of the permanent secretary (chief civil servant), directors general, and directors now depend in part on progress made toward PSA targets.

**Government of Chile.** In 1998, Chile's government began providing pay incentives for employees and teams to improve processes that assist in achieving goals. When a public sector agency meets at least 90 percent of its organizational objectives (relating to human resources, quality of customer services, planning, control and territorial management, financial management, and gender focus), staff receive a 3 percent salary increase for the following year. Achieving 75–89 percent of objectives results in a 1.5 percent salary increase.

*Sources:* Background work, 2004 AROE; Chile 2003.

pressures that have harmed their health. Interviews for this AROE confirmed the perception that task managers and operational managers are overloaded by demands from clients, managers, and reviewers; in dealing with fiduciary issues; managing trust funds and team members; and in delivering their products. In the context of the results initiative, there was concern that although managing for results was an attractive way to work, it required a great deal of effort, thought, analysis, and discussion time to do it right. Interviewees reported a considerable learning curve in adopting the results-oriented approach, yet little time for it. These barriers at the staff level, which work against the effective adoption of the results agenda, touch many aspects of the Bank's operational and budgetary systems. If the results agenda is to move forward strongly, these barriers need to be examined and addressed by the institution, including the Organizational Effectiveness Task Force.

### **Decisionmaking Based on Outcome-Related M&E Information is Weak**

Managerial decisions in the Bank have been based on information inputs, including client requests, staff proposals and analyses, budget information, and output status reports. Systematic tracking of outcome indicators and reporting systems to produce them are relatively rare. This lack presents a real constraint to managing for results.

**Country-Level Decisionmaking.** If implemented effectively, results-based CASs are likely to yield useful data on country-level outcomes that could be used as the basis for decisionmaking by country directors and country teams. Results-based CASs require identification of CAS outcomes that the Bank expects to influence through its intervention. This requirement could lead to a clearer specification of the causal chains underpinning the choices in the results-based CAS, identification of monitorable indicators for outcomes, and definition of intermediate outcomes that the Bank can more directly influence and be responsible for. Results-based CASs could enhance country-level M&E systems and influence the Bank's country-level decisionmaking. Of course, this assumes the capacity to collect the necessary data, that the right data are actually collected, that they are collected in a timely manner, that they are available in an operationally relevant form, and that they are actually used in decisionmaking. Only a handful of results-based CASs have been prepared so far, and the quality of the early examples has been mixed. Their implementation has just begun, and it is too early to tell how effective they will ultimately prove to be.

**Sector-Level Decisionmaking.** Challenges remain at the sector level, and work at this level is less advanced than at the country level. The identification and tracking of the MDG and IDA

replenishment indicators are a start. The Implementation Follow-Up (IFU-7) initiative<sup>3</sup> and the recently introduced IFU-7 Annual Outcome Report are also likely to yield useful M&E information for sector-level decisionmaking. Efforts to identify good proxies or correlates of sectoral outcomes are still a work in progress, and appropriate indicators for many sectors (for example, transportation or rural development) have yet to be identified, although work has been initiated for some sectors. Systems to ensure the collection of the relevant sectoral data and their availability to appropriate Bank staff in a timely manner will also need to be developed.

**Project-Level Decisionmaking.** Given the Bank's long involvement at the project level, Bank M&E systems for project data are relatively well developed, but not without problems (see box 4.3). For the most part, project monitoring has been overly focused on input data. Even well-designed M&E plans in project appraisal documents are seldom implemented. The revised Project Status Report, to be launched in fiscal year 2005, emphasizes project development objectives indicators and intermediate outcome indicators—if well

implemented, it could significantly improve project monitoring. Ex post impact evaluations are being conducted for a number of Bank projects, often in collaboration with the Development Economics Vice Presidency (DEC). While this is a positive development, it does not diminish the need for strong ex ante monitoring that can help to guide project implementation and enhance results. Some interviewees noted that OPCS's focus on country-level outcomes was working against project-level M&E.<sup>4</sup> The Bank's recent Action Plan to Strengthen Portfolio Performance aims to provide training to operations staff and managers in project performance rating and reporting methodology, which could potentially strengthen project-level M&E and provide more relevant data on which to base decisionmaking, but it is too early to tell how effective this Action Plan will be.

**M&E Capacity in the Bank.** The Bank must assume responsibility for monitoring the effectiveness of its own support, and Bank management must ensure that Bank staff have the capacity to do so. Even though central groups in the Bank have held M&E workshops for staff and a number of training courses are planned, staff

### Box 4.3

### Documenting the Results of Bank Projects in the Europe and Central Asia Region (ECA)

The Environmentally and Socially Sustainable Development sector unit of ECA (ECSSD) decided to document the results of its environmental and social projects. The exercise showed many Bank achievements—for example, around 2 million households are members of water-users' associations identified with the Bank; 16 countries have reformed their extension services; and Bank-supported policy units in 6 countries advise ministers and lead major negotiations. But the exercise also showed that information about results is difficult to access. The information is scattered in multiple documents (such as PSRs, aide memoires, and PADs) and multiple information systems (for example, IRIS, SAP, PDS, and Imagebank), and is often available only from interviews with project team members. Furthermore, current Bank

documentation relating to project monitoring often contains input-type information that is unintelligible to outside readers, and there are few data on early or intermediate outcomes. Sometimes there is an attempt to quantify artificially. For example, some PSRs reported "The legal system is 15 percent improved (strategy has been drafted)" and "By year four, the agency will be 100 percent strengthened." ECSSD made a number of practical suggestions for improving the Bank's systems for capturing results. One such suggestion was that managers should not approve mission reports that do not report adequately on results and should reward those that do a good job. Mission aide memoires should include a section on results to get clients thinking along those lines.

*Note:* IRIS, SAP, PDS, and Imagebank are the Bank's internal information systems.  
*Source:* World Bank 2003h

demand more and better M&E support. Interviewees highlighted the importance of systematic review of the current state of M&E information and M&E systems, including development of staff capacity as an essential prerequisite to monitoring CAS outcomes. Interviewees working on Indonesia reported that they had found the use of an M&E expert working closely with Bank staff and clients particularly beneficial to enhancing M&E capacity (see box 4.4).

### **Need for Clarity about Roles and Responsibilities**

The Bank has made some organizational changes in launching the managing for results agenda. A Results Secretariat has been established in OPCS, which has effectively explained and mobilized support for the initiative. Results Focal Points have also been appointed in each anchor and Region. These staff have operated as champions and catalysts, but there are also instances where these staff are not clear about their mandate.

A key gap in the Bank's work on results has been the lack of articulation of what a results-oriented Bank is, what kinds of M&E data are

needed to inform decisionmaking at each level, and what kinds of incentives should prevail. Also lacking is a phased plan showing how, and by when, different parts of the Bank would implement results management. Such an articulation would help the Bank reach a common understanding of what is feasible and when, and thereby help manage expectations. It would also help to focus on the key data needed for decisionmaking and prevent the creation of a "measurement bureaucracy." And it would help the Bank to better monitor the results of its results agenda. Moving to a results-oriented Bank is likely to require significant and sustained effort. Activities have been initiated in the Regions (see box 4.5), but what they add up to or how well they fit within the larger institutional effort has not been specified. Nor have measurable targets been identified.

The Bank has also yet to specify detailed cost breakdowns for its results agenda. Management indicates that the incremental costs of implementing the results agenda have been estimated (and included in the budget) at around US\$5 million annually for the next three years, mostly to prepare CAS Completion Reports, strengthen country statistical capacity,

#### **Box 4.4**

#### **Indonesia CAS: Building M&E Capacity through Hiring an Experienced M&E Specialist to Work Closely with Bank Staff and Clients**

The Indonesia program provides an example of the use of an experienced M&E specialist working full time with the country team and the government to improve results orientation. The work began during CAS preparation to guide the development of the CAS results matrix, based on a realistic assessment of what could be monitored and accounted for. This work included a review of ongoing projects and their comparable objectives and indicators. The M&E specialist, together with the portfolio/operations manager, then worked closely with the country sector coordinators and staff to identify reasonable CAS outcomes (things they could measure and had the means to achieve) and indicators for the results matrix. Following completion of the CAS, the M&E specialist prepared an M&E strategy and work plan for the country team that would help the

*Source:* Interview with the Indonesia Country Team, World Bank, 2004.

team to follow through on actions needed to further refine CAS outcomes and indicators, fill in data gaps, and ensure the adequacy of monitoring systems and comparable treatment of M&E across projects. The Indonesia portfolio/operations manager and the M&E specialist continue to work closely with task teams, advising on appropriate objectives, outputs, and performance indicators and establishing solid M&E systems for new and ongoing projects. This work is continuing, and much more remains to be done to retrofit projects and strengthen the results focus for operations. In addition, the country team and the M&E specialist have begun several M&E diagnostic studies and/or technical assistance for five key Government of Indonesia line ministries in order to strengthen the country's overall M&E systems.

improve results reporting systems, maintain a small Results Secretariat, and enhance Regional support of results-based monitoring and evaluation. But pending a full specification of what a results-oriented Bank would look like, this figure may need revision. The OPCS Progress Report recognized that resource constraints were one of the three main factors that hindered the M&E Improvement Program.<sup>5</sup> The experience of the World Bank Private Sector Development (PSD) Department's "Doing Business" initiative, described in box 4.6, shows that it is expensive to assemble the elements of results focus.

**Matrix and Sectoral Structure.** Interviewees observed that the sectoral dimension of the matrix is fragmented, which has led to excessive management and transactions costs, task proliferation, and dilution of the results focus. QAG recently found that having too many sector boards has been costly and has led to overlap-

ping functions, which has weakened the integration of respective strategies into a coherent framework. The fragmentation of the sectors works against the cross-sectoral collaboration that is necessary for achieving results (World Bank 2003b). The manifestation of this issue is illustrated by the 2003 Staff Survey results, which show that only 62 percent of staff agree that "There is effective teamwork across the different groups with which my work group works." The Organizational Effectiveness Task Force, recently set up, will review the functioning of the matrix "with a view to fine-tuning it, as necessary, to ensure its effectiveness." It will be important how the Task Force defines organizational effectiveness—in particular, whether its definition adequately reflects the Bank's focus on managing for results—although the Task Force terms of reference do not specify this. Table 4.3 presents the different ways in which specific organizational approaches—decentralization of Bank operations to field offices, and matrix-

#### Box 4.5

#### Results Focus: Highlights of Regional Initiatives

The 2004 OPCS Progress Report summarizes Regional initiatives being supported around the Bank. A few highlights are:

**Africa (AFR):** The Africa Region accounts for one-third of the approximately 15 results-based CASs under development in the Bank: Cameroon, Ghana, Kenya, Mozambique, and Zambia. The Region emphasizes the need for continuing support to country teams in designing results-based CASs.

**East Asia and Pacific (EAP):** The Region is adopting a results orientation in Regional management through continued development of unit compacts, country portfolio performance reviews, and sectoral portfolio reviews.

**Europe and Central Asia (ECA):** The Region is playing a pioneering role in communicating the results of its operations through, for example, inclusion of "results on the ground" sections in all country briefs and week-long events highlighting achievements and innovations in country and sector programs.

Source: World Bank 2004, Annex B.

**Latin America and the Caribbean (LCR):** The Region is helping to build institutional and statistical capacity in countries to monitor and manage program implementation for results and is experimenting with specific investment loans that will link disbursements to performance, such as the proposed Brazil Family Health and Vigisus II and Argentina Maternal and Child Insurance projects.

**Middle East and North Africa (MNA):** The Region is focusing on country capacity building. In Morocco, the Bank will partner with a local institute to build a permanent in-country training of trainers program in results-based management and budgeting. In Egypt, the Region is working with the government to establish the required institutional capacity to implement integrated national results-based budgeting in the Ministry of Finance.

**South Asia (SAR):** The Region is focusing on staff training on results-oriented project design and supervision. India and Pakistan are commissioning results-focused portfolio evaluations to draw lessons and identify possible actions to enhance results at the project, sector, and country levels.

**Box 4.6****Producing and Maintaining Indicators Is Costly**

In October 2003, the PSD Department's "Doing Business" initiative put out five sets of indicators on the business environment (on starting a business, enforcing contracts, hiring and firing workers, getting credit, and closing a business). The direct cost was about \$1.4 million, and there was also pro bono assistance from around 3,000 local legal business and government professionals, conservatively valued at \$8 million. In 2004

*Source:* Interview with World Bank PSD Department staff, 2004.

the indicators were updated and three more indicators (on protecting investors, dealing with business licenses, and registering property) were added, at a direct cost of \$1.9 million. Next year, which is the final year of rapid expansion, these eight topics will be updated and another three added. The cost of maintaining this indicator set would eventually stabilize at about \$1.5 million per year.

management—can foster or inhibit managing for results.

**Anchor, Network, and Sector Roles.** There is lack of clarity about the roles of anchors, networks, and sectors. Managers and staff hold a wide range of views on their role in promoting managing for results. According to interviewees, some anchors have taken up the challenge of developing effective indicators of sector performance and dialoguing on these with Regional staff (for example, infrastructure); others have well-developed indicator reporting systems, such as PSD with the Doing Business initiative, ESSD with *The Little Green Data Book*, and PREM on governance indicators, while others have emphasized support to the Regions. Some Regional managers recommend that the anchors not waste time on developing indicators, because Regional

staff need to work through this with clients. Others advise that the networks do more analysis of what type of interventions yield various kinds of results. While there may be room for diversity across sectoral practices and thematic groups, it is evident that the roles of the anchors, networks, and sectors are not commonly understood or agreed.

**Country Directors and Country Team Roles.** There is greater clarity about the role of country directors and country teams, but the quality and extent of their focus on managing for results varies widely. Since the Bank's reorganization of 1996–97, the country director function has taken on a key managerial role. Country directors, with their considerable latitude and powers; country teams; and country program coordinators are uniquely placed and empowered to play a crucial role in the managing for results framework.

**Box 4.7****Getting Results in Pakistan's Petroleum Sector**

In 1999, Pakistan's government set out to reform Pakistan's important but corrupt oil and gas sector. The Bank decided to support a modest level of technical assistance to the sector, with no formal reports and no lending. An experienced Bank team was assigned to provide advice to the government. Certain outcomes regarding the reform of the sector were agreed, including market deregulation; market-based pricing; a regulatory framework for monopoly operations; upgrading the quality of products, including phasing out of leaded gasoline; and restructuring of state enterprises.

*Source:* Interview with World Bank Country Office staff, Pakistan, 2004.

Significant progress was achieved in the design and implementation of the reform program. This progress was made possible by the agreement on the desired outcomes, receptivity of the client to alternative options, and commitment of the government to implement difficult reform measures. The reliance on monitorable outcome targets allowed the task team to enjoy considerable autonomy to carry out a dialogue and flexibility as to the type of intervention, while at the same time being held accountable to achieve results.

Interviewees reported that concentrating authority in the country director helps the results agenda, and the (at times) competing sector and country objectives are effectively resolved through various mechanisms, particularly in the context of country programs. The question is how effectively country directors and country teams build and guide their teams toward focusing on results. Box 4.7 describes a case where reliance on monitorable outcome indicators allowed the country team to enjoy autonomy and flexibility in planning operations, without compromising accountability.

Interviewees observed that the most important way to get a results focus at the country level is by working through a results-based CAS. Management estimates it will take about three years to get virtually all country programs anchored in a results-based CAS. Annual Regional targets and plans for achiev-

ing them would seem necessary, in contrast to the current short-term target of 20 by the end of fiscal year 2005. In addition to the number of results-based CASs prepared, the Bank should also monitor more outcome-type measures, such as the number of countries that can provide data on key macro and sectoral/thematic indicators and the number of countries that have adopted medium-term expenditure frameworks and/or results-oriented planning and budgeting systems. The effectiveness of results-based CASs will depend on how they are implemented.

In conclusion, the Bank has begun to collect better information about the results of its country programs, notably through the results-based CASs, and implementation of the IFU-7 initiative may lead to improvements at the sector-program level. But efforts to address the organizational culture and incentives have so

<b>Table 4.3</b>		
<b>How Specific Organizational Approaches Can Foster or Inhibit Managing for Results</b>		
	<b>Bank decentralization</b>	<b>Matrix management</b>
<b>Can foster managing for results...</b>	<ul style="list-style-type: none"> <li>• by strengthening accountability for results at country level</li> <li>• by increasing attention to client needs and priorities</li> <li>• by facilitating M&amp;E for country-level results</li> <li>• by facilitating collaborative cross-sectoral work among Bank staff in the country office, and between Bank staff and clients</li> </ul>	<ul style="list-style-type: none"> <li>• by providing a mechanism for balancing demand-driven country results and the Bank's corporate objectives</li> <li>• by providing a mechanism for making trade-offs among competing Bank Regional and sectoral priorities</li> </ul>
<b>Can inhibit managing for results...</b>	<ul style="list-style-type: none"> <li>• by increasing transaction costs between headquarters and country office</li> <li>• by increasing costs of M&amp;E for sector-level results</li> <li>• if not accompanied by appropriate delegation of authority and devolution of functions</li> </ul>	<ul style="list-style-type: none"> <li>• by increasing transaction costs between Regional and sectoral staff</li> <li>• by creating "silos" that hinder multisectoral approaches</li> <li>• by shifting attention from focus on results to proliferation of tasks born out of competition for budget resources and task leadership</li> <li>• if there is lack of clarity about roles and/or uneven performance of sector boards, anchors, and other matrix structures</li> <li>• if country directors feel pressured to parcel out tasks to each sector/network</li> </ul>

Source: Background work, 2004 AROE.

far been lacking or, at best, weak. This confirms the emerging international experience on the actions critically needed versus those usually taken by organizations in moving to a results

focus. The challenge for the Bank is to go the full course in diagnosing and fixing an organizational culture and incentive system not designed for managing for results.







# Improving OED's Effectiveness

OED evaluates the relevance and efficacy of its own products and services. With a view to enhancing future effectiveness, this chapter discusses results-oriented M&E in OED guidelines, OED's response to recommendations for itself contained in last year's AROE, Bank staff awareness of OED studies and their perception of quality, OED's Country Assistance Evaluations, and OED's support to the International Program for Development Evaluation Training (IPDET).

## Results-Oriented M&E in OED Guidelines

Evaluation can help to reinforce results-oriented M&E by asking the right questions, assessing progress, and showing how performance can be improved based on lessons learned. This section discusses the extent to which OED guidelines at the country, sector, and project levels<sup>1</sup> require evaluators to assess the extent of results-oriented M&E in Bank programs. The criteria for results-oriented M&E presented in Chapter 1 are used to assess the OED guidelines.

## Country Evaluations

The two main OED evaluations at the country level are the Country Assistance Evaluations (CAEs), which have been conducted since 1995, and the recently initiated CAS Completion Report Reviews (CASCR Reviews), following the introduction of self-evaluation of CASs through CASCRs.

Even though OED's current guidelines do not require CAEs to evaluate each of the three key criteria for results-oriented M&E (results framework, design of M&E, and implementation of M&E), some of the actual CAEs do a good job, as illustrated by the Brazil CAE (see box 5.1). There is an emerging set of good practices that could be usefully incorporated in a set of guidelines that require evaluators to assess the soundness of the M&E system in the country assistance program, including the system of indicators and efforts to strengthen borrower M&E capacity.

OED has undertaken six CASCR Reviews so far. They have focused less on results-oriented M&E compared with CAEs; neither of the two CASCR Reviews (Mozambique and Ukraine) examined for the AROE addressed any aspect of the M&E system or capacity. Given the importance of CASCRs as the main self-evaluation tool for the Bank at the country level, it is

**Box 5.1****Brazil CAE: Good Assessment of the Results Framework of the Bank's Country Program**

OED's CAE for Brazil employed specific, appropriate, quantifiable indicators to assess progress against the CAS's qualitative objectives. The CAE traced and assessed the link between Bank assistance and outcomes (including relevant MDG out-

*Sources:* Background work, 2004 AROE; OED 2004a.

comes—for example, net enrollment and access to water). It also made a careful attempt to disentangle the role of the World Bank from that of other partners in contributing to intended outcomes.

important that OED guidelines for CASCRs require that evaluators explicitly assess the results framework, design of M&E, and implementation of M&E.

### **Sector, Thematic, and Corporate Evaluations**

Sector, thematic, and corporate evaluations are OED's main evaluations at the sector and corporate levels. Current OED guidelines for these evaluations are devoted mainly to the steps required to develop and implement the study, and include little reference to evaluative dimensions such as the design and implementation of M&E. As in the case of CAEs, a review of a sample of sector and thematic evaluations and corporate studies shows greater attention to results and M&E than suggested by the current OED guidelines. For example, the *Sharing Knowledge* review explained the need for the Bank to establish monitorable outcome objectives supported by performance indicators for knowledge sharing.<sup>2</sup> OED's *Power for Development* study, which reviewed 154 projects, found that "Bank project documents provide very little data to evaluate the impact of power sector reforms on the poor.... The data that are available tend to be anecdotal and not based on sound monitoring and evaluation (M&E) systems, or empirical evidence," and concluded that "monitoring and evaluation needs to be strengthened considerably."<sup>3</sup>

The review of the OED guidelines for sector and thematic evaluations and corporate studies suggests that the guidelines should be expanded to include specific criteria for assessing the results framework and the design and implementation of M&E in Bank activities.

### **Project Evaluations**

The Project Performance Assessment Reviews (PPARs) and the ICR Reviews are OED's oldest and best-established evaluation products. As a result, guidelines and good practices are also well established, but need to be updated to incorporate the new emphasis on results orientation, especially the requirement to systematically evaluate project M&E.

Although the PPAR guidelines do not ask the evaluator to assess the project's M&E systems and M&E capacity strengthening, some PPARs do cover them. This suggests that available good practice could be used to update the guidelines, and this, in turn, could extend and generalize the good practice. For example, the Cambodia Disease Control and Health Development PPAR assesses Bank and country objectives, results chains, and timelines. The PPAR also examines the adequacy of M&E systems, including indicators and data. It touches on issues of managing for results in being critical of the "lost opportunity" in the new project to "formally link the results of M&E activities to project decisions."

ICR Reviews (as CASCR Reviews, discussed above) are more limited in their coverage and scope. OED's ICR Review template currently has no explicit question on M&E. Greater attention is needed to examining the links between outcomes and Bank activities, and the design and implementation of M&E systems.

OED has recently revived its program of impact evaluations. Although the guidelines are still under development, the first in a series of impact evaluations—the Ghana Basic Education Impact study—was completed recently (OED 2004b). This study uses cutting-

edge quantitative and qualitative evaluation techniques and represents good practice.

### **OED's Response to Recommendations for Itself Contained in Last Year's AROE**

Last year's AROE made three recommendations for OED. Actions taken by OED with respect to each recommendation are presented below.

**Enhance borrower consultations on sector and thematic studies and PPARs.** OED has involved borrowers in its recent evaluations in a variety of ways. For example, the ongoing *Community-Driven Development* evaluation involved government officials in case study countries at the design stage of the evaluation by inviting them to an entry-level workshop. The evaluation also involved local researchers and organizations in qualitative and quantitative field-level data collection. In the Ghana Basic Education Impact study, in addition to the involvement of local researchers and organizations in data collection, country officials were also involved in data analysis (see box 5.2). The El Salvador Earthquake Reconstruction PPAR included key stakeholders in survey design, survey implementation, and data analysis. The ongoing Natural Disaster study engaged borrowers in discussions relating to the Bank's effectiveness. The HIV/AIDS study surveyed country participants attending the international Nairobi conference to assess the quality, knowledge, and usefulness of the Bank's analytic work.

The *Power for Development* study used regional workshops to engage in borrower

consultations. Meetings or workshops involving stakeholders such as government officials, nongovernmental organizations (NGOs), and the private sector were leveraged for feedback or consultations by the *Extractive Industries* and *Power for Development* studies. Extensive consultations engaged people at all levels in the assessment of a wide range of issues for a Bosnia-Herzegovina PPAR. Video conferencing facilities were used to enhance the borrower consultative process for a number of OED evaluations. The ongoing *Community-Driven Development* evaluation used teleconferencing facilities to work with local researchers to refine the design of its survey questionnaires.

**Improve internal dissemination of sector and thematic studies through better targeting and more active participation in sector and regional events.** To improve internal dissemination, OED presented its *Agricultural Policy Reform in the ECA Transition Economies* study to the Rural Development Unit of ECA, and the study was subsequently posted on the ECSSD<sup>9</sup> Web site. The *Social Development* evaluation team sent out summaries and reports to task managers, country directors, and relevant Bank groups. Special editions of the *Power for Development*, *Social Development*, and *Rural Poverty* studies were printed and distributed during key Bank events. The *Community-Driven Development*, *HIV/AIDS*, *Social Development*, and *Extractive Industries* evaluations set up dedicated study Web sites.

#### **Box 5.2**

#### **Involving Borrowers in an Impact Evaluation: Basic Education in Ghana**

The OED team conducting the Ghana Basic Education Impact study collaborated with two government agencies: Ghana Statistical Services (GSS) and the Ministry of Education (MOE). It consulted with MOE from the very earliest stages, including a presentation of the proposed study design to the directors of MOE and Ghana Education Service. The study involved a survey of 1,700 households and over 700 schools that was carried out by GSS. Working with OED, MOE was involved in the design of the school survey and the education module of the

household survey. Staff from both GSS and MOE visited the World Bank headquarters twice to work alongside the OED team in the analysis of the data. This collaboration allowed the real-time input of Ghanaian staff into the study findings as they emerged, providing policy context to guide the data analysis. The government agencies are producing their own report based on their own analysis of the data. This intensive relationship helped raise the profile of the report among government officials.

*Source:* Background work, 2004 AROE.

Dissemination was also carried out through consultations and meetings with the Sector Boards and relevant staff in the anchor and Regions. Consultations with Bank staff in urban and AIDS-related areas helped draw attention of Bank staff to OED's work. Some studies, such as *Social Development*, used a "heads-up" approach to brief the relevant Sector Board on potential study findings. The preliminary findings of OED's review of lines of credit and of Bank support to financial sector reforms were presented to the Financial Sector Board and to the financial sector groups in every Region. Related background papers were disseminated through a series of brown bag lunches around the Bank. Key partnerships were developed between OED and anchor units to launch joint products such as books, as in the case of the *Urban Review*. OED's secondary education portfolio review was conducted to feed into the Education Anchor's preparation of its *Secondary Education* strategy paper. Based on OED's recommendations in the *Power for Development* study, the PSD vice-presidency prepared a guidance note to staff on private participation in the power sector.

***Institute a transparent procedure to resolve differences in ratings between OED management and country- and project-level evaluations.*** This fiscal year, OED staff made presentations in six Regional units (Africa, Middle East and North Africa, Latin America and the Caribbean, East Asia and the Pacific, Eastern and Central Europe, and South Asia) and three Sector Boards (health, education, and rural) to clarify OED standards and rating criteria and to suggest how to prepare better ICRs. Whenever there was disagreement on ratings, OED held extensive consultations with Regions and related staff, including formal meetings upon a Region's request. OED project evaluations and CAEs can include footnotes proposed by operational staff and/or reference to their views whenever warranted.

### **Bank Staff Awareness about OED Studies and Their Perception of Quality**

In addition to supporting management's accountability to the Board, OED aims to help

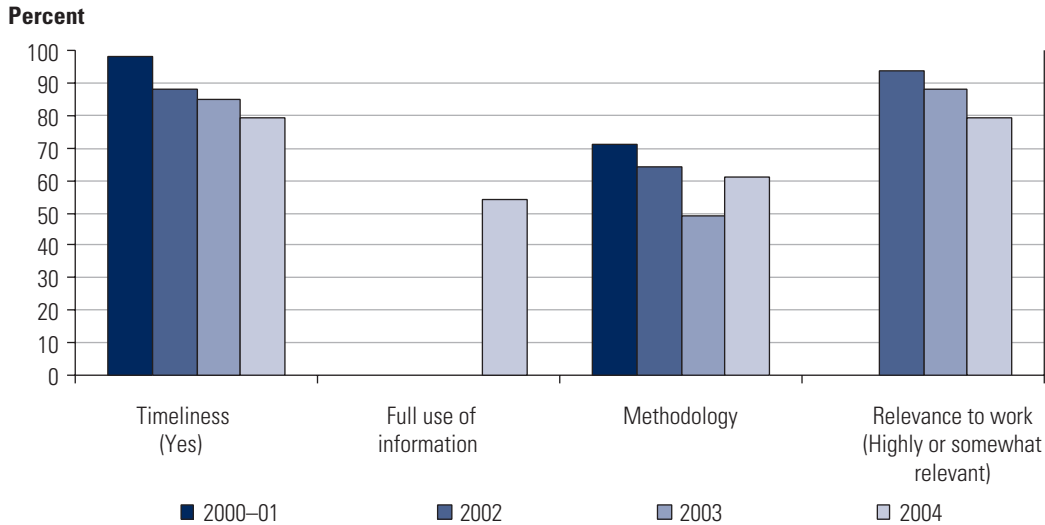
Bank staff learn lessons that can improve ongoing and future operations. OED, therefore, monitors staff readership and perceptions of its major studies through an annual survey. This year, OED surveyed 3,159 relevant Bank staff<sup>10</sup> from thematic groups, affiliated units, and networks about 12 OED sector, thematic, and corporate evaluations issued since last year.<sup>11</sup> For three corporate evaluations with broad relevance and implications across the Bank—the *Comprehensive Development Framework*, *Global Programs*, and the *2002 Annual Review of Development Effectiveness*—OED surveyed a random, rather than a targeted, sample of staff. The overall response rate was 21 percent (656 respondents). Detailed survey results are in Annex D.

***Readership.*** For all 12 studies taken together, just 27 percent of respondents reported having read the evaluation, participated in a discussion, or attended a related event such as a brown-bag seminar. Of these, 39 percent reported having read the Executive Summary and parts of the evaluation report. Of the remainder who had not read the evaluation, participated in a discussion, or attended a related event, half were not aware of the evaluation. Overall readership declined by 10 percentage points from last year's survey (which covered a different cohort of studies), mainly because fewer respondents reported having read the whole report (annex figure D.1). OED considers that readership levels will need to increase if it is to effectively disseminate lessons to operational staff.

***Aspects of Quality.*** Almost 80 percent of respondents who had read the study considered it timely and relevant to their work, but these figures have been declining for three years. In contrast, readers' views about OED methods are more favorable than they were last year (figure 5.1).<sup>12</sup> Over 60 percent of the responding readers reported that the evaluation used a methodology that was transparent and clear (12 percentage points more than the previous year). About two-thirds of respondents reported that the evaluation's recommenda-

Figure 5.1

### OED Studies Ranked on Timeliness and Relevance to Work



tions were based on its findings and conclusions (10 percentage point increase). But fewer than half the respondents found the recommendations to be cognizant of operational realities and sufficiently specific (annex figure D.2).

More than half the respondents rated as excellent or good the depth of analysis, presentation of evidence, presentation of conclusions, identification of key issues and vulnerabilities, and relevance of lessons to the Bank's work. But fewer than half of the respondents rated as excellent or good the reflection of the country context in the evaluation, persuasiveness of evidence, clarity of the rating scale, link between conclusions and evidence, reflection of borrower views, reflection of donor views, and provision of relevant options for the Bank to consider (see annex figure D.3).

**Respondents' suggestions to OED.** Over 40 percent of respondents suggested that OED improve its dissemination within the Bank and broaden external consultation. Other areas for improvement include the need to make OED recommendations more strategic and to increase dissemination outside the Bank (see annex figure D.5).

These results confirm that OED needs to significantly improve the communication and

dissemination of its evaluative findings both inside and outside the Bank. OED's new dissemination and communication strategy—currently under development—is intended to help address this issue. Areas that should receive greater attention in OED evaluations are: country context in the evaluation, persuasiveness of evidence, clarity of the rating scale, link between conclusions and evidence, reflection of borrower views, reflection of donor views, and provision of relevant options for the Bank to consider. Finally, OED needs to continue to pay attention to the timing and relevance of its evaluations. It should ensure that its evaluations address issues confronting Bank staff at specific points in time and provide staff with relevant evaluative evidence and lessons.

### OED's Country Assistance Evaluations

OED is currently conducting a retrospective of its Country Assistance Evaluations (CAEs). This retrospective draws on 25 CAEs completed in fiscal years 2001–03. It presents a discussion of the overall findings and lessons from these CAEs, reports on the usage of CAEs by the Committee on Development Effectiveness (CODE) and senior Bank staff, and suggests improvements to the current methodological approach. A number of lessons have been

distilled from these CAEs, including the usefulness of timely and high-quality ESW/AAA in informing Bank interventions, the positive role that the Bank's investment operations play in promoting institutional reforms, and the link between ownership and satisfactory outcomes in the Bank's adjustment lending—lessons that also emerge from other OED reviews, including the recent *Annual Review of Development Effectiveness* (ARDE).

Regarding the use of CAEs, a number of structured interviews were conducted with CODE members and with Bank managers and staff. CODE members indicated that they found CAEs generally to be of high quality and useful in their work, but of limited impact in country strategies. Bank managers and staff assessed CAEs positively only in cases where they could use the CAE directly in formulating a CAS or other product. The impact of a CAE depended on the degree of confidence in the competence/experience of the task manager, opportunities for interaction between the OED team and the country team, and sensitivity to the decisionmaking and operational processes. The responses from both groups also indicated that improvements to the presentation and methodological approach would enhance the effectiveness of CAEs. The retrospective will propose a number of improvements in these areas, including organizing the analysis in the CAEs around the objectives of the assistance strategies and the results achieved against each objective, clarifying and unifying the terminology, including more input from the beneficiaries of the Bank's assistance, and refining the counterfactual analysis.

### **OED's Support to the International Program for Development Evaluation Training**

OED conducts an annual course, the International Program for Development Evaluation Training (IPDET). During 2003, OED commissioned an external evaluation of IPDET that surveyed 236 participants in the first two IPDET

courses. The 90 participants that responded reported:

- A very high degree of satisfaction with IPDET: 89 percent rated the experience as good or excellent. This exceeds the benchmark target of 80 percent, which the World Bank Institute applies to its own courses if a course is to be judged of "high quality."
- A similar proportion judged IPDET to be a good investment in training and indicated that they had acquired new and improved evaluation skills. Respondents also expressed high levels of appreciation of IPDET training materials and of the IPDET listserv—a moderated electronic network involving graduates and instructors that provides an opportunity for information sharing and problem solving. Some 82 percent of respondents indicated that as a result of IPDET, they considered themselves to be part of a community of evaluation professionals.
- About half the participants indicated that their organization has strengthened its development evaluation policies and practices, and 84 percent of this group believe that their attendance at IPDET had at least some influence on this.
- Finally, about 31 percent of respondents noted that they had encountered barriers in applying the knowledge and skills acquired at IPDET in their jobs. These barriers included lack of organizational support, institutional or cultural resistance, and lack of understanding of evaluation in the workplace.

In addition to IPDET, since 1999 OED has provided country-level evaluation capacity development (ECD) support to 34 countries (and their corresponding Bank country teams): high-intensity support for 2 countries, medium-intensity support for 15, and indirect support for another 17 countries. OED has reviewed the influence of this work and will use the results to inform its future support to ECD (OED 2004c). An issue to consider in relation to the coverage of OED's support to ECD would be how the tradeoff should be made between the number of countries supported and the depth of support for each.



# Recommendations

The Bank has started the process of change toward results-oriented M&E and managing for results with an initial focus on improving its product guidelines and templates. The Bank has also begun to collect better information about the results of its country programs, notably through the results-based CASs. And although work at the sector level is less advanced than at the country level, improvements in information for decisionmaking at the sector-program level may result from the implementation of the IFU-7 initiative.

Several other initiatives are also under way in each Region to improve results orientation, such as results-oriented country and sector portfolio performance reviews and staff training on results-oriented project design and supervision. The Bank, however, still needs to close the remaining gaps in guidelines for the initiation, implementation, and completion of all its major products. Only a handful of results-based CASs have been prepared to date, and their effectiveness will depend on how they are implemented. The IFU-7 initiative, too, is still in its early stages. The Bank has not specified how the various Regional initiatives fit within the larger institutional effort, and it has yet to tackle systematically the changes in organizational culture and incentives that international experiences identify as the central challenge for managing for results. The recommendations in this AROE are intended to help the Bank move forward

faster and more effectively toward results-oriented M&E and managing for results.

## **For Bank Management**

- Develop and maintain a sustained vision, expressed in consistent messages to staff, of a results-oriented Bank, and how organizational culture, incentives, and structure need to change. In particular, management should address the sectoral fragmentation and unclear accountabilities that the matrix structure currently entails.
- Develop a phased and costed plan showing how, and by when, different parts of the Bank would implement results management. Revise the full set of product guidelines and QAG questionnaires, and provide other operational guidance for staff, to indicate how the content and process of operations need to reflect

managing for results, especially with respect to collecting and using results-oriented M&E data.

- Review and strengthen incentives such as managerial recognition, career advancement, and reward systems to promote results-oriented M&E, managing for results, and lesson learning.

### **For OED**

- Strengthen attention to results-oriented M&E in OED by requiring OED evaluations at the

country, sector, and project levels to address systematically the three key criteria for results-oriented M&E, and increase the weight that OED ratings give to the extent of results-oriented M&E and managing for results in Bank programs.

- Communicate and disseminate findings and lessons of OED evaluations more widely, both within and outside the Bank, consistent with the new communications strategy under preparation.







ANNEX A: PRODUCT INITIATION, IMPLEMENTATION, AND COMPLETION:  
GUIDELINES, TEMPLATES, AND ASSESSMENTS REVIEWED

Overall product	Intermediate products at each stage			Corporate assessments
	Initiation	Implementation	Completion	
CAS	CAS Concept Paper	CASPR (CAS Progress Report)	CASCR (CAS Completion Report)	OPCS CAS Retrospective; QAG provides demand-driven Quality Enhancement Reviews (QERs) of country portfolios.
Sector Strategies (Regional; Bankwide)	Concept Paper	Implementation Progress Report (IPR); Sector Strategy Implementation Update (SSIU)	N.A.	QAG reviews SSPs in its ESW/AAA assessment cycle; QAG reviews of sector boards; OPCS SSP reviews
Investment projects	Project Concept Note (PCN); Project Appraisal Document (PAD)	PSR; midterm review	Implementation Completion Report (ICR)	QAG assessments at entry (QAE) and during supervision (QSA); OPCS lending retrospectives
Adjustment projects	Program Document and Memorandum of the President	PSR; Board reviews of multi-tranche operations	ICR for traditional adjustment loans; ICR for programmatic lending	QAG assessments at entry (QAE) and during supervision (QSA); OPCS lending retrospectives
ESW/AAA	Concept Paper; Activity Initiation Report	Activity Progress Report	Completed paper; Activity Completion Summary (ACS)	QAG reviews of individual ESW/AAA, including strategy papers; QAG reviews of country AAA programs; OPCS ESW retrospectives/reviews



## ANNEX B: CAS RETROSPECTIVE'S THEMATIC RESULTS MATRIX

The 2003 internal CAS retrospective identified the questions that should be addressed in designing the CAS's results framework:

Thematic Results Matrix for CAS Retrospective						
Longer-term strategic goals	Shorter-term CAS outcomes				Bank priorities and performance	
	Sector-related issues	CAS outcome(s) that the Bank expects to influence through its interventions	Intermediate indicators to achieve expected CAS outcome(s)	Strategies/ actions to achieve expected intermediate indicators and outcome(s)	Bank and partner interventions in the sector	Measures of Bank performance in process and implementation
(MDG, MDG Plus)	Sector-related issues	CAS outcome(s) that the Bank expects to influence through its interventions	Intermediate indicators to achieve expected CAS outcome(s)	Strategies/ actions to achieve expected intermediate indicators and outcome(s)	Bank and partner interventions in the sector	Measures of Bank performance in process and implementation

**Longer-Term Strategic Goals.** The first column addresses the question of the strategic long-term goals, such as the MDGs or national goals, identified by the country as a priority. Not all MDGs need to be included here, only those to which the CAS outcome in the sector (or multisector outcome) will contribute. The current baseline, the national target, and the associated time frame would be captured. For example: (1) Increase in literacy rates for those over 15 years of age in the northeast of the country from 60 percent to 70 percent by 2010; increase national literacy level to 75 percent. (2) Improve the competitiveness of the manufacturing sector.

**Sector-Related Issues.** This answers the question: "What are the issues in achieving the selected MDG or other country development results?" This category describes critical issues that should be addressed to achieve country results, including specific risks to achieving longer-term de-

velopment results. For example: increasing literacy levels and contributing to increased competitiveness in the manufacturing sector require improvements in the quality of the educational system, especially in terms of learning outcomes of students. It also requires targeted programs for adult education. Deficiencies in the educational system for primary and secondary education include high dropout rates, high repetition rates, inadequate and inequitable resource allocation by school, insufficient provision of learning materials, and poor quality of teachers.

**CAS Outcome That the Bank Expects to Influence through Its Interventions.** This answers the question: "What outcome should the Bank's interventions influence directly?" This is the objective that will be the focus of the sector program (in the CAS period). It includes country-level outcome indicators against the outcome objective. It should relate to key sector-related issues, but does not

need to address all issues. For example: the Bank's interventions are designed to improve the learning outcomes of primary and lower secondary students in the northeast region of the country from 54 percent to 75 percent in grade 5 and from 33 percent to 55 percent in grade 10 (from 2003 to 2008).

**Intermediate Indicators to Achieve Expected CAS Outcome(s).** This answers the question: "What are intermediate indicators to track progress?" This column summarizes select country indicators that can be used yearly to track progress toward the CAS outcome, with baselines and targets specified. For example: (1) A reduction in dropout rates for primary and lower secondary students in the northeast from 20 percent to 15 percent by 2005, (2) 15,000 (nonqualified) teachers upgraded through certification program by 2004, (3) increase in the percentage of the national budget for education from 3.5 percent of GDP to 4.1 percent from 2003 to 2006, and (4) 50 percent of community-based school funds operational in the northeast region for purchase of learning and other classroom materials by 2005.

**Strategies and Actions in Achieving Expected Intermediate Indicators and Outcome(s).** This answers the question: "How will this be achieved?" This summarizes selective, critical government actions needed to achieve the specific strategy objective, which are supported by the Bank's program in conjunction with partners. For example: To improve the learning outcomes and reach the intermediate indicators, the Bank interventions will focus on a set of key strategies and actions, including better deployment of teacher training and management, technical assistance to the Department of Education, and increased involvement of parents in the system, as well as

strengthening ownership through school-based funds.

**Bank and Partner Interventions to Achieve CAS Outcomes.** This answers the question: "What are the Bank and partner interventions needed to assist government in achieving the CAS outcome and intermediate goals?" It shows the Bank's ongoing and planned lending and advisory knowledge services to achieve the desired outcome and partner support. For example: The Bank has a "Targeted Education Sector Project" and a "Public Sector Reform Adjustment Operation." An Out-of-School-Youth Learning and Innovation Loan is planned to enter the portfolio in 2004. An educational policy note will be undertaken in 2005. Partners are actively supporting education projects in basic education and mathematics.

**Measures of Bank Performance in Process and Implementation.** This answers the question "What is the expected performance by the Bank in its portfolio management and support to government?" It measures Bank performance across several dimensions unique to the sector or theme being discussed. This should be based on the Bank's past performance in the sector and or expected targets for major lending or AAA critical in supporting the achievement of the outcome. For example: past implementation of education projects has been slow. Thus, keeping the implementation progress rating at a satisfactory level is one of the Bank's measures. Timely implementation of components in the education sector projects will also be tracked. Planned interventions, such as the Learning and Innovation Loan, will be assessed at CAS mid-term. The Bank will increase its collaboration with donors and prepare an educational policy note to be delivered to government by CAS mid-term.

## ANNEX C: INTERVIEWEE MATRIX AND INTERVIEW GUIDE

This AROE conducted semi-structured interviews with 34 Bank staff to assess issues of organizational culture, incentives, and structure. The Interviewee Matrix and the Interview Guide are presented below.

Interviewee Matrix								
Category	AFR	ECA	EAP	LCR	MNA	SAR	Networks	Total
<b>Corporate level</b>								
Special adviser to MD								1
OPCS Results Secretariat								1
<b>Regional level</b>								
Regional vice presidents <sup>a</sup>								1
Chief economists <sup>a</sup>								1
Directors for operations <sup>a</sup>								3
Regional quality advisers				1	1	1		3
<b>Country level</b>								
Country directors	1	1				1		3
Country program coordinators		1	1		1			3
Country officers				1				1
Task managers	3	1		1	1	1		7
<b>Sector level</b>								
Sector directors		1	1			1		3
Sector managers			1	1				3
<b>Network level</b>								
Vice presidents							3	3
Results Secretariat Focal Points							2	2
<b>Total</b>	<b>4</b>	<b>4</b>	<b>4</b>	<b>4</b>	<b>3</b>	<b>5</b>	<b>5</b>	<b>34</b>

a. To maintain the confidentiality of the interviewee, Regional affiliation is not indicated.

Note: AFR = Africa; ECA = Europe and Central Asia; EAP = East Asia and Pacific; LCR = Latin America and Caribbean; MNA = Middle East and North Africa; SAR = South Asia.

## Interview Guide

1. What do you understand is the Bank's message about "managing for results?"
2. How would you define "results" in terms of Bank operations? [Probe: lending volume, number of products, outputs, outcomes, impact]. Give examples.
3. What relative emphasis do you think the Bank puts on different forms of "results?" Is that the "right" emphasis?
4. What organizational arrangements are there in your [Department, country team, Region, Sector/Network, Bankwide] for managing for results? Were changes made in structures, roles, responsibilities in moving to a results focus? Explain.
5. Do you find a change in what is expected of you or how you do your work subsequent to the Bank's adoption of the results agenda?
6. At what levels (of your organizational unit) are results data collected (who, how often) and what are the quality control mechanisms for that data?
7. How is the data used? [Probe: who gives what to whom, in what form, and how often?]
8. How does the data on results influence decisionmaking? [Probe: Decisions on country or sector strategies, administrative and program budgets, staffing, portfolio.]
9. To what extent is the results focus supported by:
  - Organizational structure (matrix, country teams, etc.)
  - Staff capacity/skills
  - Incentives [Probe: managerial recognition, pressure, fear, budget allocations, career benefits, performance evaluations, compensation, other forms of recognition/enforcement.]
  - What changes (including in the above) would enhance the results focus?
10. What are the main lessons of experience? What has worked/not worked and why?
11. What are the main challenges in managing for results?



## ANNEX D: OED SURVEY RESULTS

Since 1999, OED has undertaken an annual survey of its work to obtain feedback from clients on their perceptions of the quality and impact of OED evaluations. The underlying intention is to assess OED's influence and the development effectiveness of OED programs and activities within the Bank. The survey assesses OED evaluations in four broad aspects: readership, quality, evaluation influence on Bank processes and policies, and recommendations made to OED for improving the overall quality and relevance of its work.

This year, OED surveyed 3,159 relevant Bank

staff (level GE and above) from thematic groups, affiliated units, and networks<sup>1</sup> on 12 OED sector, thematic, and corporate evaluations that had been issued since the last survey.<sup>2</sup> For three corporate evaluations with broad relevance and implications across the Bank (the *Comprehensive Development Framework*, *Global Programs*, and the *2002 Annual Review of Development Effectiveness*), OED surveyed a random, rather than a targeted, sample of staff. Each staff member received a survey for only one study.

The graphs below detail the findings of this year's survey.

### Readership

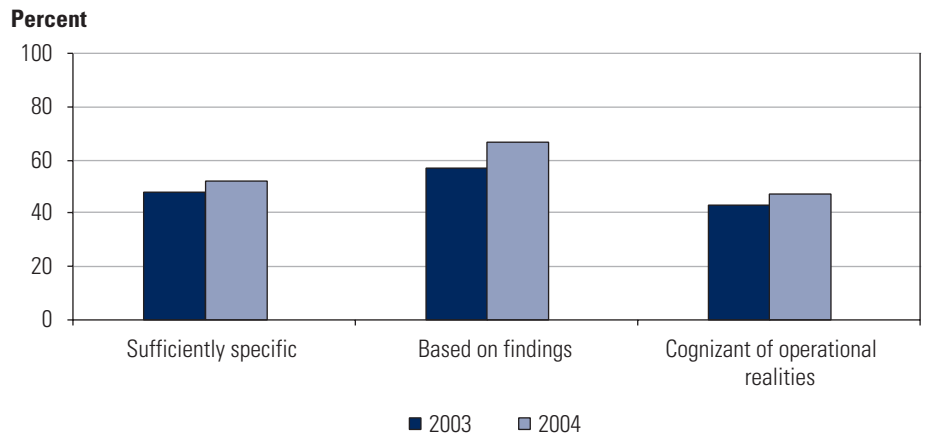
**Figure D.1**

**More Respondents Read the Executive Summary and/or Parts of the Report**



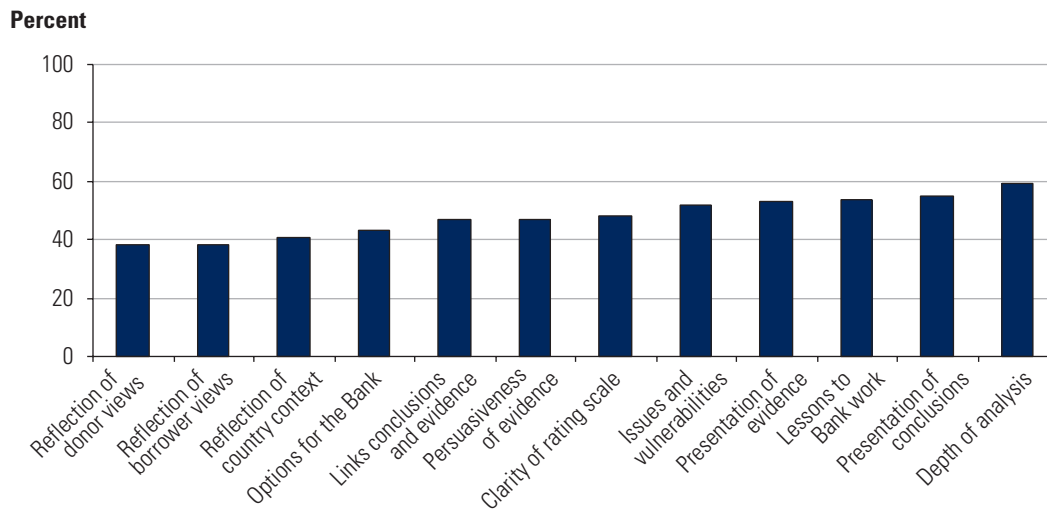
### Quality of OED Recommendations

**Figure D.2** Scope to Enhance Specificity and Operational Cognizance of OED Recommendations



### Details of Quality of Analysis

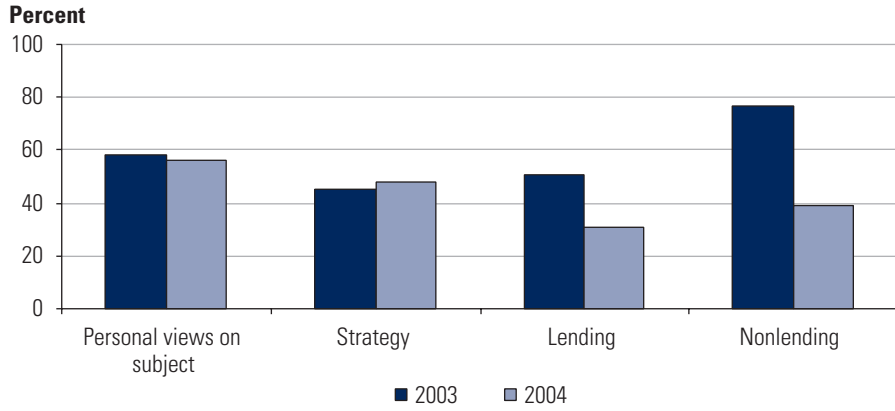
**Figure D.3** Details on Aspects of Quality of OED Evaluation



Note: Comparable question was not asked in prior years.

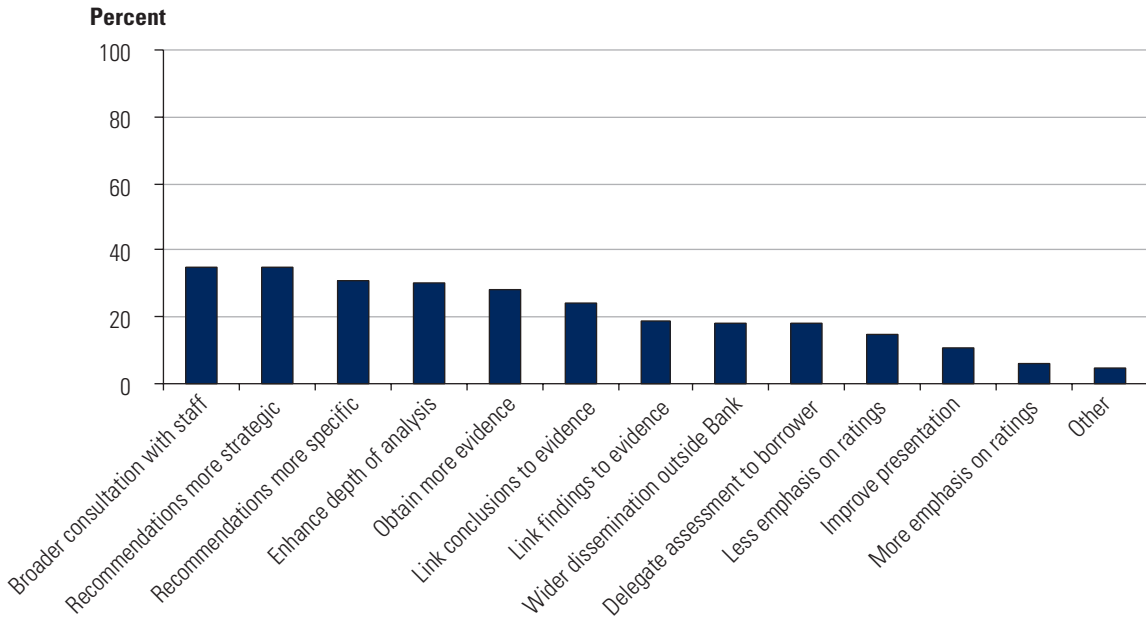
**Influence of OED Evaluations**

**Figure D.4** Influence of OED Evaluations on Staff Views and Bank Strategy (% Rated Strongly or Somewhat)



**Recommendations to OED**

**Figure D.5** Recommendations to OED: Enhance Borrower Consultation and Improve Dissemination within the Bank





## Introduction

Management welcomes the opportunity to discuss the *2004 Annual Report on Operations Evaluation* (AROE) by the Operations Evaluation Department (OED). In 2002 management made the decision to enhance its focus on results, building on the progress in improving the quality of World Bank products and strengthening control over delivery. In the context of this decision, the Bank adopted an ambitious agenda to better measure, monitor, and manage for results.<sup>1</sup> Implementation of the Bankwide agenda on better managing for results began in early 2003 with actions in three areas: (a) *in countries*, where development results are achieved, to assist countries to strengthen both their capacity to manage for results and their ownership of a results-based approach; (b) *in the World Bank*, to enhance the relevance and effectiveness of our contribution to results; and (c) *across development agencies*, to harmonize results-based approaches and better coordinate support to strengthen country capacity to manage for results. This year's AROE reviewed the Bank's monitoring and evaluation (M&E) systems slightly more than a year into the implementation of the results agenda.

**Recent Progress.** Although the Bank's recent results agenda is new, it builds on a long history of a focus on results. Most notably, the long-standing self-evaluation of the development outcomes of all Bank-supported projects, with independent validation by OED, is a model for development institutions. A careful review before the launch of this exercise showed that, in most respects, the Bank was at the cutting edge in its emphasis on results. While recent efforts to further enhance the Bank's results focus are still new, there has been notable progress in implemen-

tation across all three fronts in the last 18 months. Management reported on the implementation of the results agenda on March 31, 2004 (see World Bank 2004), and progress has continued since then. In the World Bank, which is the focus of this AROE, progress in the results agenda has been made in five areas.

- *Investment lending.* The Project Appraisal Document (PAD) was revised to focus on results-oriented design and monitoring, with project outcomes linked explicitly to the Country Assistance Strategy (CAS) and sector outcomes. It has been in use since the beginning of FY04.
- *Development policy lending.* Under the new Operational Policy 8.60, the Program Document sets out the country's program being supported and the specific results expected from the resource transfer, and the program design includes measurable indicators to monitor progress during implementation and to evaluate outcomes at completion.
- *CASs:* Seven pilot results-based CASs have now been completed and discussed by the Board. These pilots go further in specifying all key links in the causal chain, listing input, output, and outcome indicators that can be monitored regularly and used as a tool for corporate decisionmaking. Each CAS includes a CAS Completion Report (CASCRCR) and OED's review of the CASCRCR. In addition, another seven CASs have taken up the results framework methodology. A notable development is the use of transitional results matrices in Liberia and Haiti (both LICUS countries) to guide the planning and donor-coordination processes that support their recovery.
- *IDA14.* A new two-tiered reporting system was developed and proposed. The first tier moni-

tors aggregate progress in 14 country outcome indicators, and the second tier includes indicators on whether IDA programming supports countries in making progress.

- *Simplification, Modernization, Alignment, and Harmonization.* Implementation of the results agenda depends on efforts to scale up the impact and development effectiveness of the operations that the Bank supports. To scale up, the Bank is working to improve alignment of its support with country objectives, increase the degree of harmonization with other providers of development assistance around these country objectives, and to simplify and modernize Bank policies and procedures to concentrate only on those necessary for the achievement of development results. There has been progress in all aspects of this work. Bank CASs are increasingly aligned with country poverty reduction strategies. With regard to harmonization, joint assistance strategies, joint analytic work, and pooled funding arrangements have increased sharply.<sup>2</sup> Concerning simplification and modernization, Bank project documentation has been simplified and centered more clearly on development objectives, starting with the PAD, as noted above. Most important, the new policy on expenditure eligibility makes development objectives the primary determinant of Bank financing and reduces the disconnect between the types of expenditures that borrowers need to incur to implement their development programs and those that the Bank may finance. Going forward, in countries where assessed capacity is good, pilot programs will test increased reliance on country rules and mechanisms for environmental and social safeguards, true scaling up.

### AROE Analysis

The AROE provides a useful typology and overview of Bankwide M&E systems. The analysis is a timely input for the next steps in the results work program. Before addressing this analysis in detail, however, management would like to comment on OED's observation that there is a perception that aid has failed to produce significant results. That perception may indeed exist, but it is contradicted by the reality shown

in the data.<sup>3</sup> The last 40 years have seen well-documented increases in life expectancy, other health measures, and educational achievement; declines in illiteracy; and a turn-around in income poverty (from a steady increase in numbers to a decline since the early 1990s) that eclipse progress in any other comparable period in history. Development assistance has contributed to these outcomes. However, the challenge of the Millennium Development Goals (MDGs) is immense, and the Bank will need to scale up the impact of its assistance if it is to play its part in their achievement. The Bank's results agenda is designed to contribute to this scaling up.

**Concepts and Assessment Framework.** For the most part, the AROE uses the same concepts that management has used in setting out the results agenda. This is not surprising, as management has drawn on past OED work in designing the results agenda, and OED cites several management sources in its section on concepts. This alignment is important because the process of self-assessment validated by selective independent assessment, which is central to the results framework, works only when each side understands the tools used by the other. There is one caveat that management would like to highlight in this respect. Management has normally talked about managing for results, as opposed to results-based management. The latter term is sometimes linked to results-based budgeting. Management has pointed out that pure results-based budgeting is not well suited to the Bank, because cutting country program budgets when Bank performance falls short but country policy performance is good would unfairly penalize client countries. There are other, more appropriate ways for management to use the information coming from the results system to address shortfalls in performance by individual Bank units.

**The Bank's Corporate Approach to Results.** The AROE rightly points out that the results agenda is evolving. As OED notes, the agenda builds on earlier experiences with project M&E, which is a subset of the overall results framework; but more important, it builds on the success of work on qual-

ity and delivery. It is designed to provide the framework for scaling up the effectiveness of Bank assistance by using the Comprehensive Development Framework, moving more clearly to a country focus, aligning with the country's goals, and working in partnership with other donors and supporting fewer isolated projects that do not have a major impact (or, at least, a strong demonstration effect). As OED points out, earlier work on the M&E improvement program has provided valuable lessons for the work ahead. The most important lesson is country ownership—the Bank should not and cannot impose M&E systems on client countries. Client countries are rightly suspicious of the results agenda, as it can appear to be a fad that could absorb resources better used for new schools or health clinics, which would have more direct poverty-reducing benefits. It will take time and experience to get to the point where countries own the agenda and build results frameworks into all of their activities, to ensure that funds spent (for example, on new schools) have the intended outcome. (Chile, cited in the AROE, is one country where this has already happened.) The AROE suggests reconsidering Bank engagement and limiting lending to countries that do not take on the results agenda; management prefers to state its positive intent to help countries in their efforts to develop the internal demand for strong M&E systems.

**Results-based CASs.** Management welcomes OED's finding that its guidance to staff on the pilot results-based CAS program is, on the whole, on track. One of the elements that management looks for in a results-based CAS is a discussion of the country's results framework and the ways in which the Bank and other donors are supporting the country in this regard. Management will report to Executive Directors in the first half of this fiscal year on the experience with the pilot and, on the basis of that discussion, is likely to move to mainstream results-based CASs. At that time, it will update the guidelines to reflect experience drawing on work by OED.

**M&E in Lending.** Results-based CASs that are based on the client country's own objectives help deal

with one of the problems identified in the M&E Improvement Program—lack of country ownership—and they have already started to feed into the demand for better M&E in Bank-supported projects. The new PADs link results more clearly to CAS objectives and call for baseline data for monitoring results. The new Project Status Report (PSR), to be rolled out at the beginning of calendar year 2005, will pick up on these data and require updates as part of the regular monitoring process. Now that the new Operational Policy on development policy lending is in place, the draft guidelines on M&E for this product will become operational, filling an existing gap. With regard to development policy lending, management would like to note that it differs with OED's assessment that poverty reduction support credits and programmatic adjustment loans "seem to encourage the counting of small, interim procedural steps as measures of progress" and with the suggestion to use outcomes as triggers. Effective programmatic lending does indeed involve regular progress against realistic procedural actions that may appear small, but that are significant and often politically difficult; the fact that governments have demonstrated ownership by already taking actions should not be discounted. As noted in the good practice notes for development policy lending, these actions are those identified as leading to stated desired outcomes. Waiting for outcomes can mean that the financial support comes too late.

**Economic and Sector Work and Technical Assistance.**

Management takes a different view from OED with regard to the status of the architecture for evaluating economic and sector work (ESW) and technical assistance (TA). In March 2002, the Quality Assurance Group (QAG) commissioned an Advisory Panel to review experience with the ESW assessment programs, and it recommended changes for the future. The Panel's main recommendation was to shift the focus of ESW quality assessment from individual ESW tasks to the entire country portfolio of ESW and TA. The Panel saw this adjustment as both an important extension of accountability to the management level and a major opportunity for learning, notably as feedback in the preparation of a new CAS.

QAG adopted these recommendations and now assesses country ESW and TA programs. Over time, it is likely that CASCRs will become the main vehicle for self-assessment of overall country ESW and TA programs. At the level of individual ESW and TA activities, management has introduced appropriately simple and straightforward results self-assessment tracking templates that identify task objectives and success indicators at the upstream stage of each task. Management has put in place a monitoring function to ensure that these templates are used for all ESW and TA tasks.

**Sector Strategy Papers.** The AROE mentions that QAG reviewed sector strategy papers (SSPs) as part of its Sector Board Assessment and found serious shortcomings. Management notes that QAG Sector Board Assessments are not final products; rather, they were designed on an experimental basis to further clarify expectations around Sector Board products and services, given the fact that Sector Boards are relatively new, and thus they are more of a baseline assessment for management.

**Incentive Issues.** Management agrees with the AROE on the importance of incentives. The AROE discusses incentives for a results focus at three levels: shareholders, borrowing countries, and Bank staff. Management agrees that as part of the Monterrey Consensus for a higher level of funding for development, owners rightly expect a greater results focus at the World Bank and other multilateral development banks (MDBs); and management will report annually on MDB performance in the Global Monitoring Report. In connection with country incentives, the AROE mentions the IDA resource allocation system, notably the Country Policy and Institutional Assessment (CPIA) component. The AROE says that more explicitly introducing an element of results orientation into the resource allocation mechanism would give borrowers additional incentives to deliver results. However, as OED notes, higher CPIA ratings are correlated with better development outcomes. Outcomes, unlike policies and institutions, include elements that governments cannot control. For example, a

spike in primary commodity prices in a commodity-exporting country could positively affect outcomes but mask the impact of poor policies. In general, management would emphasize the importance of country ownership of the results agenda and, based on experience, would see any outside incentives through Bank conditionality as marginal. Internally, management agrees that incentives are an important element of the results agenda. Management started its results program with the results-based CAS because the country is the unit of account. This creates incentives that feed into a demand by client countries and country directors for a strengthening of the results focus of the lending and ESW that support CAS objectives. It also supports tracking of sectoral outcomes in the context of monitoring and learning from country programs. Given the immense needs of client countries, management does not see a conflict between increased lending commitments and disbursements and a focus on results. Management also differs with the AROE on where the incentive system should be in projects (2004 AROE, table 4.2). Management sees all of the stages of project preparation and execution as equally important. A satisfactory results framework at the project preparation stage is central to successful project execution. Project monitoring is equally important. For the new PSR, the system will not accept the PSR unless the module to report on project outcomes contains baseline data on project outcome indicators. Another measure management has taken to improve incentives is to make portfolio oversight an Overall Performance Evaluation item for managers. The AROE also cites the budget process, but its description of that process is incomplete. Regional budget determination is supported by three sets of analysis: country norms based on population, performance, per capita income, and inequality; costing of the proposed work program; and expenditure trends derived from external cost drivers.

**Results-Oriented in IBRD Lending and Infrastructure Initiative.** The AROE points out that increasing IBRD lending and the Infrastructure Initiative could compete with the focus on outcomes. Management notes that the results-oriented de-



sign of investment and development policy lending, ESW and technical assistance, and CASs applies to both IDA and IBRD countries.

**Structural Issues.** The AROE raises a set of structural issues. Many of these issues (for example, increasing the effectiveness of the sectoral dimension of the matrix) are under discussion in connection with the work of the Organizational Effectiveness Task Force. Management would like to note that the Task Force's terms of reference explicitly situate its work in the context of the Bank's focus on results and scaling up. The technical briefing of Executive Directors in August on the work of the Task Force emphasized this focus on results. The Task Force plans to submit its recommendations to management toward the end of 2004. Management is currently working with the Sector Boards to better refine their roles in the results agenda. One of the structural issues OED raises is the cost of the results agenda, citing the example of the indicators produced by Private Sector Development. This Bank production of country data is an exception and one that is likely over time to revert to countries. The burden of producing country results indicators is on our client countries, and the indicators need to be country-owned. However, it is true that donors can help: to help countries better manage for results, the Bank and other donors are increasing support for country statistical systems. The Bank's Trust Fund for Statistical Capacity Building, established in 1999, has approved over 50 projects valued at nearly \$12 million and covering more than 60 countries. It has also helped mobilize an additional \$11 million from other sources for these projects. Over 20 of these projects involved the preparation of national statistical development strategies. The effort has been informed by in-depth studies of statistical capacity in six IDA countries, carried out by the Bank in collaboration with other partners through PARIS21. Internally, the funding (new and reallocated funding) of the results agenda was incorporated in the Bank's budget starting in FY04 and continuing in FY05. Notably, Strategy, Finance, and Risk Management and Operations Policy and Country Services are already budgeted for their leadership work in fur-

thering the results agenda and monitoring progress.

### AROE Recommendations

The AROE has organized its recommendations for management consideration into three sets:

- Develop and maintain a sustained vision, expressed in consistent messages to staff, of a results-oriented Bank and the ways organizational culture, incentives, and structure would need to change to achieve that vision. In particular, management should address the sectoral fragmentation and unclear accountabilities that the matrix structure currently entails.
- Develop a phased and costed plan for showing how and by when different parts of the Bank would implement results management. Revise the full set of product guidelines and QAG questionnaires and provide other operational guidance for staff to indicate how the content and process of operations needs to reflect managing for results, especially with respect to the collection and use of results-oriented M&E data.
- Review and strengthen incentives such as managerial recognition, career advancement, and reward systems to promote results-oriented M&E, managing for results, and lesson-learning.

**Management Views.** Management understands the principles behind all the recommendations. Its differences with OED are in the details and timing, and about the degree to which we in the Bank comprehensively understand all that the results agenda entails and how much must be learned through trial and error. The Bank, like other aid agencies (AROE cited the Canadian International Development Agency and the United Kingdom's Department for International Development), recognizes that a results orientation is not easy because a focus on tangible results has to be balanced with a clear recognition that the Bank cannot be solely accountable for achieving results in client countries. The challenges for the Bank are also greater than those in Chile and Montgomery County (cited in the AROE) because both of those governments deliver services directly to users, while the Bank delivers support to country governments and not directly to service users.

Further strengthening incentives, in particular, is a priority, but it will take time, because management wants to avoid the risk of introducing perverse incentives for staff and managers. The AROE correctly points to risks of creating a “measurement bureaucracy” and a “checklist mentality,” in which meeting indicators—as opposed to achieving results—becomes dominant, quick wins dominate long-term sustainability, ambitions are set too low so as to ensure goal attainment, and the best people and most resources are assigned to the best-performing activities.

**Management Actions.** Despite these obstacles, the Bank is moving on budget and incentive issues. The recently announced results-focused corporate budget reform is a major element of the results agenda, and the Organizational Effectiveness Task Force is looking into structural and incentive issues. Other AROE recommendations are already part of the FY05 work program: continuing to improve the quality and focus of results-based CASs, strengthening results monitoring across sectors and for global programs, and beginning the development of a Bankwide results monitoring system that will combine data from country, sector, and global programs. This should also lead to greater cross-sectoral collaboration and closer involvement of sectors in managing for results. Management will also finalize the results measurement system for IDA14 during this fiscal year. The attached matrix provides the detailed response to OED’s recommendations. Like the quality agenda that began in the mid-1990s, the results agenda will take three to five years to implement fully; and as with the quality agenda, even after implementation there will be a constant need to adjust and fine-tune to reflect experience. This has also been the experience of other aid agencies, including those cited in the AROE. The lessons learned in assisting countries with their results monitoring systems apply equally to the Bank: monitoring systems should be simple and flexible and should be built up progressively. However, the key is to learn from what worked with regard to the quality agenda. Management is committed to an unwavering focus on results.

## Conclusion

The Monterrey Consensus underscored the shared responsibility for achieving development results, notably the MDGs and other goals aimed at sustainable growth, and the dual importance of more, and more effective, development assistance. At the same time, as was made clear in the Second International Roundtable on Managing for Development Results in Marrakech in February of this year, other agencies are looking for leadership from the Bank on how to link their own development assistance more effectively to development results. Management has responded to the leadership role by partnering with agencies in managing for results and disseminating good practices. Within the Bank, management’s commitment has been serious and sustained. Management has introduced strengthened results frameworks at the country and project levels, has linked project outcomes directly to CAS objectives and the country’s development objectives, and has started self-evaluating results in CASs. Thus the key elements for “the joined-up M&E architecture that extends logically from the country-level CAS down to products and services that are aligned with the CAS” (2004 AROE, p.10) are in place. However, it will not be easy to link incentives to results when results are achieved in countries and by the countries themselves, with the Bank in a supportive and partnership role, limited by the relatively small size of the Bank’s budget and lending program, relative to the overall investment levels of developing countries. The AROE provides a very useful overview of the state of the Bank’s M&E systems and the challenges management faces—which have also been spelled out in the Bank’s agenda for managing for results. The challenge in the coming years will be to build on the progress that has been achieved and to move steadily ahead with all three pillars of the results agenda—in countries, in the World Bank, and across development agencies. The key is relentless management focus on the results agenda, building on experience in implementing the quality and delivery agendas.

**Matrix: Management Response to the 2004 Annual Report on Operations Evaluation*****OED Recommendation***

1. Develop and maintain a sustained vision, expressed in consistent messages to staff, of a results-oriented Bank and how organizational culture, incentives, and structure need to change. In particular, management should address the sectoral fragmentation and unclear accountabilities that the matrix structure currently entails.

***Management Response***

Management has developed a clear vision of what a results-oriented Bank would look like, and it has communicated this vision in consistent messages to staff. The vision consists of actions in three areas: (a) *in countries*, where development results are achieved, to strengthen both capacity and demand to manage for results; (b) *in the Bank*, to enhance the relevance and effectiveness of our contribution to results; and (c) *across development agencies*, to harmonize results-based approaches and better coordinate support to strengthen country capacity to manage for results. Within the Bank, results-based CAS pilots and strengthened results-oriented project concept and appraisal documents adopted in FY04 are evidence of this commitment. Management will continue to consistently emphasize its commitment to the results agenda and to move the agenda forward as planned. Management commits to the following actions: the mainstreaming of results-based CASs after the upcoming review of experience (if that is the advice of Executive Directors), the roll-out of the results-based project status report (PSR) in FY05, the production and discussion of the Sector Strategy Implementation Update, progress by the Global Programs and Partnerships (GPP) Council in establishing a usable results framework for these activities (complicated by the fact that we are often just one of many partners), implementation of the results-focused budget reform, and effective follow-up to the findings of the Organizational Effectiveness Task Force.

2. Develop a phased and costed plan for showing how and by when different parts of the Bank would implement results management. Revise the full set of product guidelines and QAG questionnaires; and provide other operational guidance for staff to indicate how the content and process of operations needs to reflect managing for results, especially with respect to the collection and utilization of results-oriented M&E data.

The first recommendation here concerns our plan for managing for results. Management has adopted a phased plan (see World Bank 2004). The next steps are outlined above; their costs were included in both the FY04 and FY05 budgets, and this will continue to be the practice going forward. The second recommendation concerns product guidelines, where work is advanced. Management has accomplished or commits to the following actions: guidelines for results-based CASs will be finalized if their mainstreaming takes place as expected after the evaluation of the pilots; the full set of templates and guidelines for new investment projects will be complete with the roll-out of the PSR later this year, along with an extensive training program; the new OP on development policy lending includes specific provisions to strengthen the M&E architecture; new ESW and TA results systems are in place; and the secretariat to the GPP council will produce a results framework for global programs (but given the difficulties involved because of the partnership element, management does not commit to have the framework operational in FY05). With respect to QAG, the results focus of the questionnaire for the supervision assessment currently taking place was revised, with OED advice. Other QAG questionnaires will be similarly revised as part of the preparation for upcoming reviews.

3. Review and strengthen incentives such as managerial recognition, career advancement, and reward systems to promote results-oriented M&E, managing for results, and lesson-learning.

The Human Resources Vice Presidency has taken actions to address formal incentives: staff performance evaluation focuses on an assessment of results and the new promotion criteria call for a sustained record of operational results. The pilot results-based CASs (last year's experiment) lay out more clearly what country teams will be accountable for. The new Project Status Report to be piloted this year can't be completed unless it includes updated monitoring data and a manager's sign-off. As important, this year management will include

*(Matrix continues on the following page.)*

**Matrix: Management Response to the 2004 Annual Report on Operations Evaluation (continued)**

performance on portfolio implementation as an element in the performance evaluation of all operational managers. Incentives are also being reviewed in the context of the work of the Organizational Task Force and budget reform. Recognition is important and staff are recognized Bankwide in a variety of ways for their contributions to results, notably through the President's Awards for Excellence and OED's Good Practice Awards; and management will investigate a new award for results-focused work. Management will report on progress in the context of updates on the work of the Task Force, budget reform, investment lending modernization and simplification, and harmonization and alignment initiatives, as well as periodic reporting specifically on the results initiative, with the next report planned for mid-2005. Management is working on a draft Operational Policy on M&E and plans to complete it in this fiscal year.

---

---

ANNEX F: CHAIRMAN'S SUMMARY: COMMITTEE ON DEVELOPMENT  
EFFECTIVENESS (CODE)

---

The Committee on Development Effectiveness (CODE) met on September 13, 2004, to review the findings and recommendations of the *2004 OED Annual Report on Operations Evaluation* (2004 AROE).

**Overall Conclusions and Next Steps.** The Committee members welcomed the OED annual report, which addresses the key issues relating to consolidating the ethos of “managing for results.” Several members commented that the OED findings echoed issues that their chairs have drawn attention to on previous occasions. Accordingly, they were not surprised by such findings but remained concerned with their persistence and implications. The rich discussion revealed a diversity of views among OED, management, and various Committee members. The main issues that were raised clustered around the following topics: (i) organizational culture and incentives at different levels, (ii) diffusion or fragmentation of accountability, particularly in the context of the “matrix” structure; (iii) structure of the results architecture including the factors underlying the disconnect between the timeliness of OED findings and their incorporation in Bank operations; (iv) terms of reference of the Task Force on Organizational Effectiveness; (v) OED’s effectiveness; (vi) country ownership and how the results framework is best implemented; and (vii) harmonization. Several members requested a CODE and/or Board discussion on any proposed modification to tracking by OED and management of management’s implementation of recommendations arising from OED’s sector, thematic, and corporate evaluations, and wanted OED to report back on progress in next year’s AROE, and on the measures it takes to respond to its own recommendations.

Among the specific issues raised by the Committee were:

**Organizational culture and incentive framework.** Members unanimously agreed that the Bank must pay more attention to becoming a results-oriented institution. This requires changing the organizational culture and incentives; internalizing lessons learned from experience; and fostering actions that are not just new bureaucratic steps or mere additional reporting requirements. Most members noted the importance of monitoring and evaluation (M&E) to consolidate the culture for results management in the global development effort, poverty reduction, and the MDGs. While some members asserted that there is no conflict between increased lending and focus on results, another, quoting from OED interviews, questioned the compatibility of the increased focus on lending with the focus on outcomes (lending quality). Management agreed that there is more to do on organizational changes; however, it also noted that over the last year there has been good progress in initiatives such as country teams benefiting from cross-sectoral collaboration, the pilot results-based CASs, the introduction of the new Project Concept and Project Appraisal Documents, and incorporation of baseline data in projects against which to monitor progress. Management noted that there is no tradeoff between increasing lending and lending quality.

**HR management.** Members asked for additional information about the status and experience of performance-based HR management instruments—reward systems (i.e., career promotions, pay increases) and systems to address poor performance. They also sought further detail on

the improvements in clarifying staff responsibilities at approval, implementation, and ex-post evaluation of operations. One member agreed with the management response's position on strengthening incentives for staff, and cautioned against HR management practices that increase staff risk aversion to work in challenging environments and undertake difficult operations. Another member regretted that the target of mainstreaming M&E in Bank operations by FY04 has not been met and asked about the sufficiency of incentives, skills, and resources for internal capacity building. Management noted that implementing results-based performance and promotion is not easy, for example, when assigning staff to risky projects (or countries) and judging their performance relative to staff who work on easier projects. Management reported the adoption of new procedures and incentives for quality project supervision, concentrating more on progress toward development objectives after the approval of projects. The new project status report to be piloted in October and November, and mainstreamed in January, can't be completed without updating the monitoring data and without sign-off of managers. As importantly, this year management will include performance on portfolio supervision as an element in the performance evaluation of all operational managers.

**Assessing results and country outcomes.** One member suggested adding two elements to the criteria for assessing the results orientation of M&E (box 1.3): (i) client needs should be linked in and (ii) individual indicators ought to be of high quality (timely, reliable, valid, etc.). Another member noted the need to apply a robust methodology in assessing the Bank's development contribution, and asked why the Bank terminated the Corporate Scorecard exercise. Management answered that the Scorecard was just premature. A question was asked whether the Bank should take a comprehensive approach to M&E, or a phased approach, i.e. a learning-by-doing approach. Another member asked about the role of country systems and institutions and the resource implications to determine and monitor country outcomes. Members generally

agreed with management that one cannot say that aid has failed to produce significant results.

**Affordability.** Some members highlighted the need to watch the budgetary impact of implementing M&E recommendation. Others believed that producing and maintaining indicators is not costly if an operationally pragmatic approach is adopted, while the opposite is true for academic or purist approaches.

**Sector, anchor, or network accountability.** Members agreed with OED that management should eliminate diffused accountability and clearly establish the respective roles and responsibilities of the numerous sector anchors and networks in managing for results. Management acknowledges the need to do more in this regard.

**Disconnect between OED and management.** Some members regretted the striking differences of opinion regarding the implementation status of a few of OED's recommendations. They also noted the divergence between OED's recommendation emphasizing the need for a vision and a plan to implement results management and the management's position focusing on consolidation based on an existing vision and plan. Differing views were expressed about the adequacy of the management response, where one member found the response unsatisfactory. The diverging views related to linking results with resource allocation at the country, operational unit, or individual levels and associating inputs with country-level outcomes. One member observed that management and OED generally agreed on the M&E framework for the ESW and technical assistance portfolio. OED noted that the disconnection often arises from management's focus on plans and OED's focus on action. Management responded that it has taken actions, but in a fast-changing world there is a constant need to adjust and fine-tune actions, drawing on experience, that it is counterproductive to set too ambitious targets in the short run.

**Adoption of OED recommendations by Bank management.** Some members suggested that management and OED should review and possibly

reduce the number of recommendations that are tracked to a more manageable level, increase the precision of OED's recommendations, and follow up more systematically. They emphasized the importance of implementing the accepted recommendations. A few members disagreed with management's suggestion for unilaterally dropping several outdated recommendations. Management noted with some concern that the number of recommendations has increased over time, and indicated that it is willing to work together with OED on deciding which recommendations are outdated, completed, or will stand. Several members' sentiment was that CODE should be involved in this process. Going forward, management will try to be more disciplined in suggesting to CODE members which recommendations should be accepted. It will outline actions to complete the implementation of accepted recommendations.

**Gaps in guidelines.** Some members supported OED's recommendations to close the remaining gaps in guidelines for lending and nonlending products related to the results framework, design and implementation of M&E, and accountability of middle management. Regarding adjustment operations, some members shared OED's concern about the lack of specific M&E guidance, while others agreed with management that triggers should focus on measurable policy reform actions, and did not favor use of proxies for outcome. Management noted that actions in triggers are those identified as leading to stated desired outcomes. Outcomes are longer-term targets; waiting for outcomes can mean that the financial support comes too late. The EU experiment has faced some difficulties. Management also noted that this year the Bank has made progress in results frameworks in CASs, lending and nonlending activities, especially ESW; M&E is now required for development policy lending (DPL) which was not the case in the old policy for adjustment lending. A member welcomed the pilot program of results-based CASs as a vehicle to develop guidelines, but also cautioned that additional complexities may arise in the Country Partnership Strategies (CPS), which are increas-

ingly demand-driven, flexible, and less certain or specific in terms of expected outcomes and timeframes.

**Country ownership.** While members agreed on the importance of strengthening country M&E systems, they stressed that countries must own the results agenda. A few members felt that Bank support for country M&E capacity building should have been included in the scope of the study. Some agreed with management's perception that borrower countries are suspicious of the results agenda. However, the Bank should not impose M&E systems on clients or set up enclave systems that only support the Bank's own data needs. The challenge for OED is how it will conduct country evaluations while countries are still developing their results framework and how management would address capacity issues in LICUS (low-income countries under stress) cases. A member shared OED's concern relating to weaknesses in promoting results and M&E at the country level. OED stated that the Bank's engagement for improving M&E in countries that have weak M&E and little appetite to strengthen M&E should be step-by-step. OED also noted that it had just distributed to CODE a self-evaluation of Evaluation Capacity Development.

**Task Force on Organizational Effectiveness.** Several members suggested that the Task Force should not merely fine-tune the existing matrix structure but explicitly address the issue of sectoral fragmentation. Management noted that this issue is under discussion in connection with the work of the Organizational Effectiveness Task Force, which has taken a results focus as its starting point.

**OED's own effectiveness.** Some members suggested that OED's effectiveness may be increased by undertaking the following: (i) improving communication and dissemination of its work, and helping internalization of lessons from evaluation in Bank policies and operations; (ii) minimizing the additional burden on operational staff, given that the Bank's overall human and budgetary resources are becoming more severely limited; (iii) paying more atten-

tion to cost-effectiveness of its recommendations. One member expressed concerns about the decrease for three consecutive years in the ratios of respondents who considered OED studies timely and relevant. OED noted that it is working to increase the effectiveness of its work with a new communication strategy, and will work to improve the timeliness and relevance of its work during the next year. OED also replied that it weighs costs and benefits when it conducts evaluation. A member also asked OED to be clearer in future AROEs as to by when and what measures OED will take to respond to its own recommendations.

**Harmonization.** A few members questioned the exclusion of harmonization from the scope of the OED's study. One member inquired about results-based management with other development institutions. OED noted that there are examples of international agencies introducing some approaches to managing for results, and IFC is linking allocation of bonuses at departmental and individual level with results. Management noted that the Bank is well recognized as having taken the lead in managing for results in the international agenda, and that it is important to acknowledge the sustained progress made by the institution in this area.

*Pietro Veglio*  
Acting Chairperson



---

## ENDNOTES

---

### Chapter 1

1. OED reviewed the Bank's evaluation capacity development efforts in 2002 (OED 2002b) and recently completed a self-evaluation on evaluation capacity development (OED 2004b).

2. The Bank's frameworks for monitoring and evaluation can be found in tables 2.1 and 3.1, respectively, of the 2003 AROE.

3. The log frame approach was introduced to project-based development assistance by USAID, by most accounts in the early 1980s. It was quickly adopted by the United Nations Development Program (UNDP) (around 1985), followed by several European bilateral agencies. Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) spearheaded efforts to translate it into a highly structured, facilitative tool for participatory project planning under the title "Objective Oriented Project Planning." The World Bank adopted the log frame approach in 1997.

4. The causal chain is subject to risks. The log frame approach requires risk indicators to be identified and tracked in order to help account for any departures from the expected consequences of program activities.

5. This AROE uses the term "impact" to refer to change in outcomes attributable to the intervention.

6. See OECD/DAC 2002. The Bank defines "results" as "sustained improvements in country outcomes." Unlike the Bank's definition, the definition in this AROE does not restrict itself to country-level outcomes—local, regional, global, corporate, sector, and project outcomes are also important.

7. The links between outcomes and proxy variables for outcomes should derive from, or be verified by, research and other analytical work.

8. With minor adaptations, these criteria are applicable to each level: country, sector, and project.

9. See OECD/DAC 2002. This AROE uses the term "impact" to refer to change in "outcomes" attributa-

ble to the intervention. The Bank defines managing for results as "using information to improve decisionmaking and steer country-led development processes toward clearly defined goals" (see World Bank 2004).

### Chapter 2

1. In 1996, under the "Next Steps" program, the Bank developed sector-specific performance indicators for several sectors. In 1997, the Bank launched the Strategic Compact, which aimed to make the World Bank "more effective and efficient in achieving its main mission—reducing poverty." In 1998, OED recommended that the Bank adopt results-based management.

2. Management notes that at the project level, the use of M&E as part of design and implementation is now required for all operations. The new Project Supervision Report will require baseline data on the project outcome indicator or indicators.

3. OED's corporate, country, sector, and project evaluations provide Bank management with an assessment of the outcome of closed or mature operations.

4. A third exercise, the Global Monitoring Report (GMR), monitors developing country policies, developed country policies, and the role of International Financial Institutions (IFIs) in achieving the MDGs and related development outcomes. The role of IFIs was considered the least-developed part in the initial GMR. The indicators and methodologies to measure IFI role and contribution need further refinement.

### Chapter 3

1. The current guidelines for (traditional) CASs and the Bank's CAS Web site do not provide specific guidance to country teams on how to specify the causal chain that has to underlie a results-based CAS. Nor do current CAS guidelines provide basic design stan-

dards for the CAS's M&E system and CAS-specific performance indicators, beyond a reference to the *Performance Indicator Handbook* (World Bank 1996). The *Handbook* provides in-depth information on indicators, but operational guidelines should cull out of the literature the basic standards for the design and use of performance indicators that the Bank wants met in CASs. Having each CAS team work through fairly extensive documents is inefficient and dilutes the focus of the CAS team on basic standards. With respect to M&E implementation, the guidelines contain general, albeit sound, advice relating to long-term capacity building. They do not discuss the possibility of assessing borrower M&E capacity during CAS preparation.

2. The CAS review identified results orientation as the weakest of the four pillars of the Bank's CASs (vision, diagnosis, programming, results), with only 40 percent having satisfactory evaluation frameworks.

3. Management notes that the CASCR is not the key tool for assessing country capacity in M&E. It is the CAS that sets the M&E support strategy and the instruments to implement the strategy.

4. It could also help to assess how some of the challenges relating to MDGs identified in OED's *2002 Annual Review of Development Effectiveness* are being addressed—for example, focusing on outcomes among the poor, not just average outcomes, and identifying the results chain and monitoring the appropriate intermediate indicators.

5. Project development objectives should focus on the outcome for which the project reasonably can be held accountable, and not encompass higher-level objectives that depend on other efforts outside the scope of the project (World Bank 2003f).

6. Except for the so-called "simple" projects, for which it is optional.

7. Guidelines for the Project Appraisal Document, see World Bank 2003f.

8. ESW is a subdivision of AAA. It includes different types of analytic and diagnostic reports, conferences and workshops, and country dialogue/consultations. AAA encompasses ESW, nonlending technical assistance, and donor and aid coordination.

9. The core diagnostics include Poverty Assessments (PAs), Country Economic Memoranda/Development Policy Reviews/Social and Structural Reviews (CEMs/DPRs/SSRs), Public Expenditure Reviews (PERs), Country Procurement Assessment Reviews

(CPARs), and Country Financial Accountability Assessments (CFAAs).

10. CASCRs could serve as one tool for the ex post self-evaluation of ESW.

## Chapter 4

1. Special advisers to managing directors, OPCS Results Secretariat staff, Regional vice presidents, chief economists, directors for operations, Regional quality advisers, country directors, country program coordinators, country officers, task managers, sector directors, sector managers, network vice presidents, and network Results Secretariat Focal Points.

2. An increase in IBRD lending of \$1 billion in each of the next two fiscal years was proposed by Bank management at the 2004 Strategic Forum. Management has stressed that no tradeoff will be made on the quality of support.

3. The IFU-7 sectors are: (i) investment climate and finance; (ii) education for all; (iii) HIV/AIDS; (iv) child and maternal health; (v) water supply and sanitation; (vi) trade; and (vii) environmental sustainability. With varying degrees of specificity and completeness, the Bank has identified indicators for these sectors that address three issues: (i) what results on the ground are important?; (ii) how does the Bank contribute to these results?; and (iii) how do we know if implementation is on track?

4. That the country is the main unit of account should not mean that M&E relating to projects, sector strategies, or AAA/ESW is not needed. It simply means that it is not enough. Monitoring the effectiveness of its support at each level helps the Bank to manage its overall assistance program and demonstrate its accountability better.

5. OPCS 2004 Progress Report, page 45 (World Bank 2004).

## Chapter 5

1. At this time, the only existing guidelines for OED evaluators are contained in OED's internal procedures documentation.

2. It elaborated a results chain against which progress and performance in the knowledge-sharing area could be judged.

3. It pointed out the Bank's weak performance monitoring, which focused on inputs and outputs and provided little data on outcomes or impacts.

4. On level of adoption, the recommendations are rated as **High**: Fully adopted; **Substantial**: Largely

adopted but not yet fully incorporated into policy, strategy or operations; **Medium:** Adopted in some operational and policy work, but not to a significant degree in key areas; and **Negligible:** No evidence/plan for adoption, or plan and actions for adoption are in a very preliminary stage.

5. On status of recommendations, they are rated as **Active** and remain actionable by management, **Complete, Obsolete, or overtaken by events**, or **Difference of opinion** between management and OED.

6. Recommendations that were not rated by **both** management and OED have not been included in the calculation of percentages and disconnects. The Extractive Industries Study was not included in this year's figures, which were prepared before CODE discussed the Draft Management Response for this study. A recommendation from the Aid Coordination Review has been incorporated into a related recommendation in the more recent CDF evaluation.

7. Management notes that virtually all of the decline was due to a downgrade in the rating of recommendations from two older OED studies and a corporate issue. The rate of adoption of recommendations from recent OED reviews—where traditionally the level of adoption is low—was actually higher than that of last year. In these three cases of downgrading, subrecommendations that were completed or were overtaken by events were dropped, and OED's new rating reflects the rate of adoption of the remaining subrecommendations. As a result, in some cases, progress by management appears to have resulted in lower ratings.

8. This excludes corporate reviews.

9. Environmentally and Socially Sustainable Development sector unit of ECA.

10. In addition to Bank staff, OED's sector, thematic, and corporate evaluations have several other audiences both within and outside the Bank, including the members of the Board of Executive Directors, who were not surveyed this year.

11. The 12 OED evaluations surveyed comprised 7 sector and thematic evaluations (on *Extractive Industries, Social Development, Anti-Corruption, the Consultative Group on International Agricultural Research, Rural Poverty, Power for Development, and*

*Global Programs—Phase 1*) and 5 corporate evaluations (*Annual Review of Development Effectiveness 2002, Annual Review of Operations Evaluation 2003*, the multi-stakeholder *Review of the Comprehensive Development Framework*, and reviews on the *Heavily Indebted Poor Countries* and *Sharing Knowledge*).

12. Respondents in prior years were not asked about the “full use of information.”

## Annex D

1. Last year OED surveyed 856 members of the Bank staff from relevant thematic groups for the various OED studies produced in 2003. In addition to sector and thematic studies, OED also surveyed relevant task managers of PPARs assessed in 2002. PPARs were not a part of this year's survey.

2. Sector and thematic studies included *Extractive Industries, Social Development, Anti-Corruption, Consultative Group on International Agricultural Research, Rural Poverty, Global Programs 1, and Power for Development*. The corporate methods studies included *Annual Review on Operations Evaluation, Annual Review of Development Effectiveness, Review of the Comprehensive Development Framework, Heavily Indebted Poor Countries*, and *Sharing Knowledge*.

## Annex E

1. The Action Plan was discussed by the Bank's Committee on Development Effectiveness (CODE) in December 2002. After that discussion, the paper *Better Measuring, Monitoring, and Managing for Development Results: Implementation Action Plan* (World Bank 2002) was revised and submitted to the Board for information.

2. Documentation under preparation for the upcoming Second High-Level Forum on Harmonization and Alignment and Managing for Results (Paris, February 28–March 2, 2005) will summarize progress.

3. See *Development Effectiveness, Partnership, and Challenges for the Future* (DC2002-0006), April 9, 2002, and *The Role and Effectiveness of Development Assistance—Lessons from World Bank Experience*, available on the Bank's research website at <http://econ.worldbank.org>.



---

## BIBLIOGRAPHY

---

*Included are selected articles on the international experience with M&E and results.*

- Barrados, Maria. 2003. "Can Organizations in Government Learn?" Presentation at OECD Symposium on The Learning Government: Managing Knowledge in Central Government, February 3-4, 2003. Paris, OECD.
- Bhuiyan, Bayes, and R. Hay. 2003. "Health & Population Sector Program, Annual Program Review, Independent Technical Report." Bangladesh.
- Binnedijk, Annette. 2001. "Results-based Management in the Development Co-operation Agencies: A Review of Experience." Paris, OECD.
- CDF Secretariat. 2003. *Toward Country-led Development: A Multi-Partner Evaluation of the Comprehensive Development Framework*. Washington, D.C.: The World Bank.
- Chile, Ministry of Finance, National Budget Office. 2003. *Systems of Management Control and Results-Based Budgeting: The Chilean Experience*. Santiago.
- Danida. 1996. *The Logical Framework Approach. A Flexible Tool for Participatory Development*. Copenhagen.
- Flint, Michael, C. Cameron, S. Henderson, S. Jones, and D. Ticehurst. 2002. "How Effective Is DFID?" DFID Evaluation Report EV 639, London.
- Gates, Stephen. 1999. *Aligning Strategic Performance Measures and Results*. New York: The Conference Board.
- Hatry, Harry, and K. Yansane. 2002. *Comprehensive Development Framework Evaluation. Results Orientation: An Early Look*. Washington D.C.: The World Bank.
- Kettl, Donald F. 1997. "The Global Revolution in Public Management: Driving Themes, Missing Links." *Journal of Policy Analysis and Management* 16 (3).
- Kusek, Jody Zall, and Ray C. Rist. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners*. Washington, D.C.: The World Bank.
- . 2002. "Building Results-based Monitoring and Evaluation Systems: Assessing Developing Country Readiness." *Zeitschrift für Evaluation* 1. Bern; Leverkusen. Germany.
- Kusek, Jody Zall, Ray C. Rist, and E. M. White. 2005. "How Will We Know the Millennium Development Goal Results When We See Them?" World Bank Africa Region Working Paper Series No. 66. Washington, D.C. (Also, forthcoming, in *Evaluation*.)
- Maxwell, Simon. Forthcoming. "Lost in Translation? Implementation Constraints to Results-based Management." In UNDP/DFID, *Enhancing Development Effectiveness: A New Focus on Managing for Results*. New York: Earthscan.
- Meier, Werner. 2003. "Results-based Management: Towards a Common Understanding among Development Cooperation Agencies." Discussion Paper, Results-Based Management Group, Ottawa.
- Morgan, Peter, and Anne Qualman. 1996. *Institutional and Capacity Development, Results-based Management and Organizational Performance*. Ottawa: CIDA.
- OECD (Organization for Economic Co-operation and Development). 1994. *Performance Management in Government: Performance Measurement and Results-oriented Man-*

- agement. OECD Occasional Papers on Public Management, no. 3. Paris.
- OECD/DAC (Development Assistance Committee). 2002. *Glossary of Key Terms in Evaluation and Results Based Management*. Paris: OECD.
- OED (Operations Evaluation Department, The World Bank). 2004a. *The CGIAR at 31: An Independent Meta-Evaluation of the Consultative Group on International Agricultural Research*. OED Study Series. Uma Lele, task manager. Washington, D.C.: The World Bank.
- . 2004b. *Books, Buildings, and Learning Outcomes: An Impact Evaluation of World Bank Support to Basic Education in Ghana*. OED Study Series. Howard White, task manager. Washington, D.C.: The World Bank.
- . 2004c. "Evaluation Capacity Development: OED Self-Evaluation." Keith Mackay, task manager. [www.worldbank.org/oed/ecd](http://www.worldbank.org/oed/ecd)
- . 2003a. *2003 Annual Report on Operations Evaluation*. Deepa Chakrapani, task manager. Washington, D.C.
- . 2003b. *2002 Annual Review of Development Effectiveness: Achieving Development Outcomes: The Millennium Challenge*. OED Study Series. Soniya Carvalho, task manager. Washington, D.C.: The World Bank.
- . 2003c. *Sharing Knowledge: Innovations and Remaining Challenges: An OED Evaluation*. OED Study Series. Catherine Gwin, task manager. Washington, D.C.: The World Bank.
- . 2003d. *Debt Relief for the Poorest: An OED Review of the HIPC Initiative*. OED Study Series. Madhur Gautam, task manager. Washington, D.C.: The World Bank.
- . 2002a. *2002 Annual Report on Operations Evaluation*. Laurie Effron, task manager. Washington, D.C.
- . 2002b. "Annual Report on Evaluation Capacity Development." Keith Mackay, task manager. Washington, D.C.
- . 2002c. *The World Bank's Approach to Global Programs: An Independent Evaluation—Phase 1*. Washington, D.C.
- . 2002d. *IDA's Partnership for Poverty Reduction: An Independent Evaluation of Fiscal Years 1994-2000*. OED Study Series. Catherine Gwin, task manager. Washington, D.C.: The World Bank.
- . 2001. *Integrating Gender in World Bank Assistance*. Gita Gopal, task manager. Washington, D.C.
- . 2000. *The World Bank Forestry Policy: Striking the Right Balance*. OED Study Series. Uma Lele, task manager. Washington, D.C.: The World Bank.
- . 1999. *Investing in Health: Development Effectiveness in the Health, Nutrition and Population Sector*. OED Study Series. Susan Stout and Timothy Johnston, task managers. Washington, D.C.: The World Bank.
- . 1985. "Built-in Project Monitoring and Evaluation: An Overview." Washington, D.C. OED, OEG, and OEU.
- OED, OEG, and OEU. 2003. *Power for Development: A Review of the World Bank Group's Experience with Private Participation in the Electricity Sector*. OED Study Series. Washington, D.C.: The World Bank.
- OED, OEG, OEU, and MIGA. 2003. *Extractive Industries and Sustainable Development: An Evaluation of World Bank Group Experience*. Vols. 1-4. Washington, D.C.: The World Bank.
- Perrin, Burt. 2002. *Implementing the Vision: Addressing Challenges to Results-focused Management and Budgeting*. Paris: OECD.
- QAG (Quality Assurance Group, The World Bank). 2004. "Quality at Entry Investment Lending Questionnaire." [http://siteresources.worldbank.org/QAG/Resources/QEA6\\_Investment\\_Questionnaire.pdf](http://siteresources.worldbank.org/QAG/Resources/QEA6_Investment_Questionnaire.pdf)
- . 2003a. *Annual Review of Portfolio Performance*. Washington, D.C.
- . 2003b. *Quality of ESW in FY02: A QAG Assessment*. Washington, D.C.
- Schick, Allen. 1998. "Why Most Developing Countries Should Not Try New Zealand Reforms." *The World Bank Research Observer* 13 (1).
- Scott, Alison. 2004. "Assessing Multilateral Effectiveness." DFID, London. Photocopy.
- Senge, Peter. 1990. *The Fifth Discipline. The Art and Practice of the Learning Organization*. London: Century Business.
- UNDP (United Nations Development Program). 2002. "RBM in UNDP: Overview and General Principles." New York.
- White, Howard. 2004a. "M&E for Results." Background paper, 2004 AROE. OED, Washington, D.C.

- . 2004b. “The Triple A Standard for Agency Performance Measurement: Aggregation, Alignment, and Attribution.” Background paper, 2004 AROE. OED, Washington, D.C.
- . 2003. “Using Development Goals and Targets for Donor Agency Performance Measurement.” In Richard Black and Howard White, eds., *Targeting Development*. New York: Routledge.
- World Bank. 2004. “Implementation of the Agenda on Managing for Results: Progress Report.” Washington, D.C. <http://www.managingfordevelopmentresults.org/>
- . 2003a. “Better Measuring, Monitoring, and Managing for Development Results: Implementation Action Plan.” Washington, D.C.
- . 2003b. “Managing for Results: Implementing the Agenda.” Presentation at meeting of the Operational Vice-presidents, 8 September 2003. Washington, D.C.
- . 2003c. “IDA Results Measurement System: Update Note.” Washington, D.C.
- . 2003d. “Cameroon Country Assistance Strategy.” Washington, D.C.
- . 2003e. “Honduras Country Assistance Strategy.” Washington, D.C.
- . 2003f. “Template and Guidance for the Project Concept Note.” <http://info.worldbank.org/etools/docs/library/37492/GuidelinesforCompletingProject.pdf>
- . 2003g. “Enhancing Measurement, Monitoring, and Analysis of Poverty in Russia.” ECA, ECSPE. <http://www.worldbank.org.ru/ECA/Russia.nsf/ECADocByUnid/48D7A70C140A197CC3256CCC0046E34A>
- . 2003h. “Measuring the Impact of ECSSD Interventions: Final Report for Financial Year 2003.” World Bank Europe and Central Asia Region, Environmentally and Socially Sustainable Development. Washington, D.C.
- . 2002. “Better Measuring, Monitoring, and Managing for Development Results.” Paper presented to the Development Committee meeting, September 28, 2002. Washington, D.C.
- . 2000. “Report on the Working Group on Improving Quality of Monitoring and Evaluation in Bank-financed Operations.” Washington, D.C.
- . 1996. *Performance Indicator Handbook*. Washington, D.C.
- . 1993. *Getting Results: The World Bank’s Agenda for Development Effectiveness*. Washington, D.C.
- . 1992. *Effective Implementation: Key to Development Impact. Report of the World Bank’s Portfolio Management Task Force (The Wapenhans Report)*. Washington, D.C. World Bank and International Monetary Fund.
2002. *Better Measuring, Monitoring, and Managing for Development Results*. Washington, D.C. : The World Bank. [http://site resources.worldbank.org/DEVCOMMINT/Documentation/90015418/DC2002-0019\(E\)-Results.pdf](http://site.resources.worldbank.org/DEVCOMMINT/Documentation/90015418/DC2002-0019(E)-Results.pdf)

