

IEG

INDEPENDENT EVALUATION GROUP



THE WORLD BANK



2006 Annual Report on Operations Evaluation



THE WORLD BANK GROUP

WORKING FOR A WORLD FREE OF POVERTY

The World Bank Group consists of five institutions—the International Bank for Reconstruction and Development (IBRD), the International Finance Corporation (IFC), the International Development Association (IDA), the Multilateral Investment Guarantee Agency (MIGA), and the International Centre for the Settlement of Investment Disputes (ICSID). Its mission is to fight poverty for lasting results and to help people help themselves and their environment by providing resources, sharing knowledge, building capacity, and forging partnerships in the public and private sectors.

THE INDEPENDENT EVALUATION GROUP

ENHANCING DEVELOPMENT EFFECTIVENESS THROUGH EXCELLENCE AND INDEPENDENCE IN EVALUATION

The Independent Evaluation Group (IEG) is an independent, three-part unit within the World Bank Group. IEG-World Bank is charged with evaluating the activities of the IBRD (The World Bank) and IDA, IEG-IFC focuses on assessment of IFC's work toward private sector development, and IEG-MIGA evaluates the contributions of MIGA guarantee projects and services. IEG reports directly to the Bank's Board of Directors through the Director-General, Evaluation.

The goals of evaluation are to learn from experience, to provide an objective basis for assessing the results of the Bank Group's work, and to provide accountability in the achievement of its objectives. It also improves Bank Group work by identifying and disseminating the lessons learned from experience and by framing recommendations drawn from evaluation findings.



2006 Annual Report on Operations Evaluation



©2006 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
E-mail: feedback@worldbank.org

All rights reserved

1 2 3 4 5 09 08 07 06

This volume is a product of the staff of the International Bank for Reconstruction and Development / The World Bank. The findings, interpretations, and conclusions expressed in this volume do not necessarily reflect the views of the Executive Directors of The World Bank or the governments they represent.

The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgement on the part of The World Bank concerning 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 International Bank for Reconstruction and Development / The World Bank encourages dissemination of its work and will normally grant permission to reproduce portions of the work promptly.

For permission to photocopy or reprint any part of this work, please send a request with complete information to the Copyright Clearance Center Inc., 222 Rosewood Drive, Danvers, MA 01923, USA; telephone: 978-750-8400; fax: 978-750-4470; Internet: www.copyright.com.

All other queries on rights and licenses, including subsidiary rights, should be addressed to the Office of the Publisher, The World Bank, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2422; e-mail: pubrights@worldbank.org.

Cover illustration credit: ©Images.com/Corbis.

ISBN-10: 0-8213-6845-1
ISBN-13: 978-0-8213-6845-9
e-ISBN:0-8213-6846-X
DOI: 10.1596/978-0-8213-6845-9

Library of Congress Cataloging-in-Publication Data have been applied for.

World Bank InfoShop
E-mail: pic@worldbank.org
Telephone: 202-458-5454
Facsimile: 202-522-1500

Independent Evaluation Group
Knowledge Programs and Evaluation Capacity
Development (IEGKE)
E-mail: eline@worldbank.org
Telephone: 202-458-4497
Facsimile: 202-522-3125



Printed on Recycled Paper



Contents

v	Acronyms and Abbreviations
vii	Acknowledgments
ix	Executive Summary
xv	Summary of Management Response to IEG Recommendations
xix	Chairman's Summary: Committee on Development Effectiveness (CODE)
3	1 Introduction
	3 Using M&E Information for Results
	4 Overview of the 2006 AROE
7	2 Monitoring and Evaluating Results in Bank Operations
	7 Recent Developments in Monitoring
	8 How a Results Chain Works
	9 Possible Uses of Results Frameworks
	10 Current Status of Monitoring Results
	14 Challenges to Developing Results-Based Monitoring
	17 Self-Evaluation and the Bank's Results Agenda
	19 Implications for Future Work
25	3 Improving IEG's Effectiveness
	25 IEG's Results Framework
	26 IEG's Outputs: Increasing the Focus on Results and Impact in Independent Evaluation
	27 Strengthening the Focus on M&E in IEG Evaluations
	29 IEG Initiatives for Applied Learning and Real-Time Use of Evaluation
	29 Deepening Strategic Partnerships
	30 Strengthening IEG's Results Framework
	31 Quality and Relevance of IEG Outputs
	32 Communicating Knowledge from IEG Evaluations
	33 IEG Intermediate Outcomes: Evaluation Recommendations Incorporated into Bank Operations and Policy
	34 IEG Intermediate Outcomes: Use of Evaluation Findings by Bank Staff at the Operational Level

4	37	Conclusions and Recommendations
	39	Recommendations for Management
	39	Recommendations for IEG-WB
41		Appendixes
	43	A: Overview of Monitoring and Evaluation in the World Bank
	51	B: Implications of the World Bank's Results Agenda for Monitoring and Evaluation
	59	C: Review of CAS Results Frameworks
	69	D: Focus Group and Individual Interviews with CODE Members, Country Directors, Sector Managers, and Task Team Leaders
	79	E: Client Survey Results
	93	F: Overview: Changes in IEG's Dissemination Methods
	97	G: Accelerating the Results Agenda
	101	H: World Bank Management Action Record
	105	I: Management Response
109		Endnotes
113		Bibliography
		Boxes
	12	2.1 What Constitutes a Strong Results Framework?
	15	2.2 What Are We Hearing from Managers and Staff?
	18	2.3 GTZ's e-VAL System: Using Computer-Supported Evaluation Tools for Rapid Progress Reviews
	20	2.4 Integration of Objectives in Results Frameworks
	21	2.5 Reducing the Gap Between Country- and Project-Level Outcomes
	29	3.1 Real-Time Use of IEG Knowledge—Accountability through Education Indicators in Peru
		Figures
	8	2.1 Sample Results Chain for Agricultural Development
	13	2.2 Results Chains Better Developed in IDA Countries than in IBRD or Blend Borrowers
	13	2.3 Country Assistance Strategies Tend to Lack Adequate Baselines and Targets
	13	2.4 Precision of Milestones Could Be Improved
	13	2.5 Larger Programs Tend to Have More Indicators, Outliers May Need Review
	26	3.1 IEG Results Chain
	27	3.2 Uses of Evaluation Knowledge
	32	3.3 Bank Staff and Executive Directors and Advisors Generally Satisfied with Quality of IEG Evaluations
	34	3.4 Level of Adoption of IEG Recommendations Similar to That in Previous Years
		Table
	16	2.1 Regional Client Services

ACRONYMS AND ABBREVIATIONS

AAA	Analytical and advisory activities
ACS	Activity Completion Summary
ADB	Asian Development Bank
AIS	Activity Initiation Summary
AROE	Annual Report on Operations Evaluation
ARPP	Annual Report on Portfolio Performance
BP	Bank procedure
CAE	Country Assistance Evaluation
CAS	Country Assistance Strategy
CASCR	CAS Completion Report
CBD-CDD	Community-based and -driven development
CC	Country coordinator
CD	Country director
CGAP	Consultative Group to Assist the Poor
CODE	Committee on Development Effectiveness
COMPAS	Common Performance Assessment System
CPFR	Country Portfolio Performance Review
CPR	Country program review
DAC	Development Assistance Committee (Organisation for Economic Co-operation and Development)
DEC	Development Economics Vice Presidency
DGF	Development Grant Facility
DIME	Development Impact Evaluation Initiative
DPL	Development Policy Loan
ECD	Evaluation capacity development
ECG	Evaluation Cooperation Group
ED	Executive director
EI	Extractive industries
ESW	Economic and sector work
GMR	Global Monitoring Report
GPP	Global program and partnership
GPR	Global Program Review
GRM	Grant reporting and monitoring
GTZ	Deutsche Gesellschaft für Technische Zusammenarbeit
IBRD	International Bank for Reconstruction and Development
ICR	Implementation Completion Report
IDA	International Development Association
IEG	Independent Evaluation Group (World Bank)
IFC	International Finance Corporation
ISR	Implementation Status and Results Report
LICUS	Low-Income Country Under Stress

LOC	Line of credit
M&E	Monitoring and evaluation
MAR	Management Action Record
MDB	Multilateral development bank
MDG	Millennium Development Goal
MIGA	Multilateral Investment Guarantee Agency
NGO	Nongovernmental organization
OED	Operations Evaluation Department (now IEG)
OD	Operational directive
OP	Operational policy
OPCS	Operations Policy and Country Services (World Bank)
PAD	Project Appraisal Document
PDO	Project development objective
PPAR	Project Performance Assessment Report
PSR	Project Supervision Report
PRS	Poverty Reduction Strategy
QAG	Quality Assurance Group
RBCAS	Results-Based Country Assistance Strategy
RMLS	Results Monitoring and Learning System
RMS	Results Measurement System
RSG	Results Steering Group
SAP	Systems Applications and Products
SM	Sector manager
SSIU	Sector Strategy Implementation Update
SPC	Strategy and Performance Contract
STATCAP	Statistical Capacity Building Multicountry Adaptable Program Loan
TFSCB	Trust Fund for Statistical Capacity Building
TTL	Task team leader
VPU	Vice presidential unit

OED changed its official name to the Independent Evaluation Group (IEG) in December 2005. The new designation “IEG” will be inserted in all IEG’s publications, review forms, databases, and Web sites.



Acknowledgments

This report was prepared by a team comprised of Janardan Prasad (J.P.) Singh (task manager), Jan Rielaender, and Yoshine Uchimura. The peer reviewers were Sid Edelman (Independent Evaluation Group [IEG]-International Finance Corporation), Gregg Jackson (George Washington University), and Alfred Nickesen (World Bank Central Operational Services Unit), who provided invaluable advice and guidance in preparing the report.

Janet Mancini Billson and Martin Steinmeyer from Group Dimensions International conducted focus group interviews for this report. Susan Collins and Howard Smith from the Advanced Performance Consulting Group evaluated Country Assistance Strategy results matrices.

Thanks also go to Forum One Communications for support in implementing and analyzing the client survey, as well as to all those who responded to the survey and participated in focus groups.

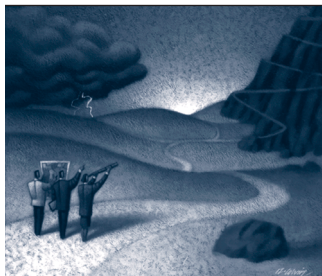
Many IEG staff, including Nils Fostvedt, Patrick Grasso, Keith Mackay, and Ray Rist, provided extensive support and contributions to all phases of this report. Aida Tapalova provided research assistance.

Others who have provided valuable input and assistance include Vivian Jackson, Pierre-Joseph Kingbo, Princess Moore-Lewis, Maria Gabriela Padrino, Juicy Qureishi-Huq, and Barbara Yale. Julius Gwyer, Maria J. Mar, Alex H. McKenzie, and Yang Soo Yoon offered computer and technical support.

Document production support was provided by William Hurlbut. Caroline McEuen edited the report.

The report was prepared under the direction of Klaus Tilmes, Manager, Knowledge Programs and Evaluation Capacity Development.

Director-General, Evaluation: *Vinod Thomas*
Director, Independent Evaluation Group–World Bank: *Ajay Chhibber*
Manager, Knowledge Programs and Evaluation Capacity Development: *Klaus Tilmes*
Task Manager: *Janardan Prasad (J.P.) Singh*



Executive Summary

The 2002 International Conference on Financing for Development in Monterrey emphasized the need to measure, monitor, and manage for results and concluded with joint statements issued by the leaders of the multilateral development banks. The World Bank began implementing its Managing for Development Results action plan in 2003. The plan calls for steps to (a) strengthen countries' capacity and demand to manage for results; (b) enhance the Bank's relevance and effectiveness in contributing to results; and (c) harmonize results-based approaches across development agencies.

The Bank's results agenda shifts attention from outputs of development activities, such as dollars lent or reports written, toward country and sectoral as well as global outcomes, such as the Millennium Development Goals, economic growth, and environmental and social sustainability. This shift has implications for monitoring and evaluation (M&E). In a results framework, monitoring includes tracking outcomes as well as outputs, while evaluation assesses the contribution of Bank operations to achieving results.

Managing for results requires better information on program performance at the project, sector, country, and global levels to guide decision making and reporting. The focus on results is a joint responsibility of countries and development agencies and requires parallel efforts to strengthen demand and the capacity to make use of information. The results agenda combines heightened accountability at the country and development

agency levels with learning to understand better what works, what does not, and why.

The *2006 Annual Report on Operations Evaluation* (AROE) updates the actions taken since the 2004 and 2005 AROEs to strengthen the results focus in M&E. First, it analyzes the implications of managing for results on Bank operations. Second, it assesses the extent to which the Bank's M&E systems provide staff with the information they need to better manage for results. Third, the report evaluates products and services of the Independent Evaluation Group-World Bank (IEG-WB) as part of a continuous process of self-evaluation and assesses their quality, influence, and use among both internal and external audiences. The report also raises some unresolved issues for further consideration. Finally, it makes recommendations on how M&E can be strengthened to increase the World Bank's effectiveness.

World Bank Progress in Implementing the Results Agenda

The Bank has made progress in putting in place an architecture for managing for development results, but more work is needed to extend this framework beyond country programs and projects.

Policies and procedures have been revised to manage better for results. Results-Based Country Assistance Strategies (RBCASs), Country Assistance Strategy (CAS) Completion Reports, and IEG validation of the completion reports provide the basis for managing for country-level results. The results frameworks are intended to clarify the Bank's contribution to the achievement of country development goals. This addresses a weakness of previous CASs, which discussed country goals and Bank operations, but did not link them systematically.

Investment Loans and Development Policy Loans now have formal results frameworks that link individual Bank projects to country development goals. Even if more difficult to do in many respects, similar efforts are needed in sectoral program areas such as infrastructure and health initiatives and in global initiatives, whether in education or the environment. The Bank needs to assess more systematically and strategically what has worked well in different sectors, as is being currently attempted for health and for infrastructure.

Some Bank operational teams are using self-evaluation to improve performance, but these efforts are not yet systematic.

While still at a pilot stage, evaluative methods such as rapid results approaches, deployed during implementation, can improve the effectiveness of Bank projects and programs. Impact evaluations have been recognized as a formal analytical and advisory activity product, and the Development Impact Evaluation Initiative has identified some 60 ongoing impact evaluations, mostly in social sector and poverty programs. A variety of approaches can be used to assess program impact, including well-focused objectives-based assessments, rates of

return analysis, retrospective impact evaluation, and randomized experiments.

The Bank's country partners are still at an early stage in adapting to a results focus.

Lack of capacity and the additional cost of data collection are seen as constraints to M&E in the Bank's client countries. A few countries, particularly those preparing Poverty Reduction Strategies, are receiving Bank support to strengthen their national statistical systems that support improved poverty monitoring. Some countries, most visibly in Latin America, are working with Bank support to strengthen systems for M&E, both governmentwide and at the sectoral level. Engagement with both the country and multilateral and bilateral partners is essential in these efforts.

The Bank's approach to the results agenda needs to be supported by complimentary action on several fronts where early resolution could enhance the effectiveness of the results agenda.

The AROE analysis confirmed the importance of focusing on country outcomes and aligning Bank operations around them, often within multisectoral settings. Yet incentives for cross-sectoral teamwork remain weak. Also, with growing globalization, the potential role of cross-country interventions—and the Bank's contribution to them—is likely to grow.

Yet inadequate measurement of the effects of such interventions in areas such as health, water and sanitation, and the environment holds back current or future support for them. The Bank is encouraging staff to take on difficult but important challenges, such as issues of governance or the environment or working in tough country circumstances, such as those of fragile states. In these situations, the focus on measurable and achievable results needs to be factored into the complexity of the process and the difficulty of outcomes so that attention and incentives do not shift exclusively toward the outcomes that are easiest to obtain or measure.

Are M&E Systems Providing What Is Needed to Manage for Results?

The use of the results framework as an effective monitoring and management tool remains in question.

An AROE review of 25 CAS results frameworks prepared during fiscal 2005–06 found that about one-third of the CASs had weak results frameworks that made it impossible to follow the links between country goals, CAS outcomes, and Bank activities. In addition, more than half did not include baselines and did not set specific, time-bound targets. The absence of performance indicators with baselines and targets reduces the usefulness of the results frameworks.

In cases where indicators are present, there are commonly too many, which makes the framework an unwieldy tool. At the project level, the Quality Assurance Group found that the clarity and realism of development objectives could have been improved in about 40 percent of projects reviewed in fiscal 2004–05. In addition, 40–50 percent of the projects reviewed lacked well-articulated results frameworks with baselines and systems for tracking and reporting progress. An Operations Policy and Country Services (OPCS) review of the quality of results frameworks in Implementation Status and Results Reports (ISRs) that was prepared in fiscal 2005 found that 40 percent did not contain any outcome indicators with baseline data for either the project development objective or intermediate outcome indicators, and were not considered acceptable.

Disconnects in outcome ratings between country and sector policy outcomes and project outcomes point to a gap in understanding the factors that contribute to results.

An IEG retrospective of Country Assistance Evaluations (IEG 2005d) found that in a third of the country programs reviewed, aggregate project outcomes were satisfactory, but the overall country assistance programs were unsuccessful. Similar disconnects have emerged at the sector level. While the Bank's primary education portfo-

lio performed above Bank average and met stated expansion targets, most such programs failed to look beyond increased enrollment and completion rates to how well students were learning. An IEG evaluation of the Bank's trade assistance (IEG 2006a) found that Bank advice and lending contributed to systematic reductions in protection and in opening up economies. But the impact on employment and poverty reduction was less than expected, and the Bank did not conduct sufficient analysis to inform its policy and lending in this area. Systematic evaluation during and on completion of Bank activities could help improve understanding of factors critical to success that can be used to refine the results framework at the country, sector, and project levels.

Bank managers and operational staff encounter a variety of challenges in managing for results.

Focus groups and interviews conducted for this AROE indicated that managers and operational staff struggle to link higher-level goals to Bank operations and to identify performance indicators. The possible uses of performance information remain unclear to many staff, and this acts as a disincentive to using monitoring to achieve outcomes. Competing pressures in the Bank, combined with the lack of clarity about the utility of M&E, relegate performance information to a lower priority.

The Bank has made progress in putting formal policies and procedures in place, but those changes have not yet translated into improved incentives and consistently applied operational practices.

The challenge remains to change the Bank's culture and the staff perceptions that serve as informal incentives and disincentives to managing for results. A particular issue worthy of further work is to incorporate the relative difficulty or ease of success into the results frameworks, and to find the appropriate balance between measures of success and measures of risk.

There is a particular need for such balance in considering and encouraging Bank engagement in low-income countries under stress, for

example, or on difficult thematic issues such as corruption or environmental degradation. The recommendations of the 2004 AROE remain valid: the Bank needs to maintain and communicate its vision for a results-oriented institution, strengthen incentives, and develop a cohesive and fully funded action plan.

The Bank has outlined what needs to be done to manage for results during the next stage, but lack of prioritization and funding raise concerns about the likelihood of success.

The 2006 OPCS results progress report (World Bank 2006b) identified possible actions to accelerate the implementation of the agenda but did not identify priority areas and funding sources. M&E improvement initiatives have failed in the past because of fragmentation of efforts, resource constraints, and lack of staff capacity. These lessons need to be considered.

How Is IEG Contributing to Meeting the Results Agenda?

IEG's mandate is to assess whether the World Bank Group's programs and activities are producing the expected results. While IEG is fully independent, it is placed within the World Bank Group to provide strategic decision makers and operational staff with knowledge that helps them work more effectively.

IEG is reinforcing its focus on M&E and is deepening its work to develop evaluation capacity among Bank clients.

IEG is strengthening its assessment of M&E in its evaluations. Country, sector, and project evaluations now give more attention to M&E and the effectiveness of results frameworks, but more remains to be done. IEG is increasing its support of results-oriented M&E capacities in client countries and refocusing its high-intensity support in a few targeted countries, while maintaining less-intensive support to a broader range of countries.

Greater emphasis is needed on applied learning and real-time use of evaluation findings to improve Bank performance.

The AROE survey queried clients about the quality, use, and influence of IEG evaluations and found that IEG provides good services to the Board. However, task team leaders are the least likely of the groups queried to report that they incorporate evaluation findings into planning, designing, and implementing Bank projects and programs. IEG needs to find more effective and innovative channels to bring its findings to task teams in a way that would help improve design of future operations, and work with Bank management to ensure relevant evaluative lessons are reflected in future programs.

Continued emphasis is also needed on the development of the newly established initiative of providing quick-turnaround products for just-in-time knowledge to decision makers at the strategic and operational levels.

IEG is increasing the follow-up, monitoring, and accessibility of the Management Action Record, its primary tool for tracking management progress in implementing IEG recommendations.

One of the ways IEG can influence the Bank's effectiveness is through the recommendations it makes to management in its evaluations. The Management Action Record (MAR) tracks progress in incorporating findings and recommendations in Bank operations and policy. The 2006 MAR contains 79 recommendations. IEG rated the level of adoption of its recommendations by management "high" or "substantial" for 66 percent of the recommendations. IEG and management agreed on the rating of adoption for 65 percent of the recommendations. A summary of the 2006 MAR can be found in appendix H.

IEG has a critical role in the workings of the results agenda in the World Bank Group.

The AROE includes a results framework for IEG that identifies its contribution to the Bank's results agenda. Within that framework, IEG's outputs are the findings, lessons, and recommendations from its evaluations and evaluation

capacity support in client countries. Outcomes are twofold. First, for accountability, the evaluations are used by the Board to fulfill its oversight functions. Second, Bank management—as well as countries—are to use the recommendations from evaluation in policies and procedures.

Even if they are difficult to measure and target, greater attention needs to be given to these outcomes of evaluative work. There has been progress on this count. For example, early findings from the recent evaluation of the Bank's primary education assistance (IEG 2006c) were incorporated in the 2005 Education Sector Strategy Update. The trade assistance evaluation (IEG 2006a) is being used to recalibrate the Bank's approach to trade, with pilot initiatives getting under way in several Regions. The recent evaluation of natural disaster assistance (IEG 2006d) made special efforts to provide useful findings to operational staff dealing with current events as well as to inform emerging international responses.

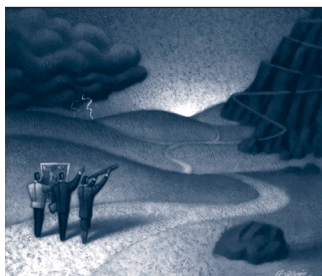
Reviews of Country Assistance Strategy Completion Reports (CASCRs) could provide inputs into the preparation of new CAS. For learning, IEG products are aimed at improving the Bank's policy advice and program and project design: all parties can do much more to strengthen this learning agenda. The final outcome for IEG outputs is the use of its evaluative knowledge to improve the effectiveness of Bank operations.

Recommendations for Management

- Building on progress achieved so far in advancing the results agenda, agree on a three-year action program supported by appropriate funding for the implementation of the next stage, differentiating new allocations and re-deployed resources. Critical action items would be the following:
 - Support country directors and country teams in their efforts to refine and use results frameworks at country and sectoral levels to manage country programs. Similar efforts need to be made to strengthen the results frameworks of thematic and global programs.
- Assess the effectiveness of (self-) evaluation approaches during project and country program implementation and provide guidelines to staff on their use.
- Strengthen incentives and accelerate a results-oriented training and communications program for management and staff to encourage use of M&E information.
- Identify and support in-depth learning opportunities to develop and use results-based approaches with task teams, particularly in challenging country environments and in complex, multisectoral settings.
- Provide support to task teams and technical advice to countries that intend to institutionalize M&E systems to strengthen the collection and use of performance information in order to enhance the countries' capacity and demand to manage for results.

Recommendations for IEG-WB

- To further strengthen IEG's contribution to the workings of the results agenda in the World Bank Group, IEG should continue to follow its own results framework and monitor it through the AROE. Its focus on the usefulness of evaluation findings for its core audiences should be enhanced: for the **Board** for oversight, for **management** through the incorporation of recommendations into Bank policies and strategies, for **Bank staff** through the use of evaluation findings for policy advice to country partners and in project design, for **external partners** through the use of evaluation findings to improve their programs and policies, and for the **countries** more broadly. In playing this role, IEG should:
 - Improve the timeliness of its evaluations.
 - Strengthen the operational relevance of the findings.
 - Increase access to and exchange of the lessons.



Summary of Management Response to IEG Recommendations

IEG recommendation	Management response
<p>1. Building on progress achieved so far in advancing the results agenda, agree on a three-year action program with a corresponding budget to fund implementation of the next stage, differentiating new allocations and redeployed resources. Critical action items would be:</p> <ul style="list-style-type: none"> • Support country directors and country teams in their efforts to refine and use results frameworks at country and sectoral levels to manage country programs. Similar efforts need to be undertaken to strengthen the results frameworks of thematic and global programs. • Assess the effectiveness of (self-) evaluation approaches during project and country program implementation and provide guidelines to staff on their use. • Strengthen incentives and accelerate a results-oriented training and communications program for management and staff to encourage use of M&E information. • Identify and support in-depth learning opportunities to develop and use results-based approaches with operational teams, particularly in challenging country cases and in complex multisectoral settings. 	<p>At present, management does not agree that new budget allocations are required to implement the results agenda. A results focus is the responsibility of country teams and task teams and is to be included in the formulation of RBCASs and task preparation and task supervision budgets. Management regularly monitors the progress of the work on the results agenda and will continue to formally report on progress to CODE every two years. The most recent progress report includes a specific action plan that is now being implemented. A key step forward has been the formation of a Bank-wide Results Steering Group (RSG), which is now functioning. The RSG is responsible for tracking progress on the action plan articulated in the March 2006 progress report and for promoting communication and coordination across network and Regional level efforts. An early task for the RSG is to review and assign priorities for implementing the action plan and to work with operational management teams to ensure that these priorities are included in Regional and network budgets and work programs. If this process identifies major budget pressures that go beyond what can be reasonably funded through reallocation, management would discuss this issue with executive directors using established budgetary procedures as outlined in the recent budget document approved by executive directors.</p> <p>Management is pleased to report progress on other critical actions identified in this recommendation:</p> <ul style="list-style-type: none"> • Regional Quality Teams are supporting country directors and country teams in their efforts to refine and use results frameworks at country and sectoral levels to manage country programs. As an example, the Latin America and Caribbean Region together with OPCS carried

(Continues on the following page.)

IEG recommendation	Management response
	<p>out during fiscal 2006 a regional portfolio review of results orientation in operations that was backed by regional management and presented to country directors and country and sectoral teams. Results frameworks at thematic and global program levels are also being strengthened. Six new results frameworks are being developed as part of either sector strategy updates or the 2006 Sector Strategy Implementation Update: health, nutrition and population, environment, forestry, social protection, finance, and transport. Results frameworks in global programs are also being strengthened. Each vice presidential unit is now explicitly accountable for periodically reviewing and adjusting global programs to ensure that they focus on results. New procedures will allow the Bank to track its global programs from their start through to evaluation and impact assessment.</p> <ul style="list-style-type: none"> • The Country Director/Country Management Peer Learning Event in September 2006 will include a specific discussion of the results agenda to identify methods for accelerating progress in enhancing staff understanding of how to construct appropriate results chains at the country level. Management will work to establish an analogous peer learning structure for sector directors and sector managers during fiscal 2007. • Management does not see the need to issue additional guidance to assess the effectiveness of (self-) evaluation approaches during project and country program implementation, as there are mechanisms and processes already in place for this purpose. The Quality Assurance Group reports on effectiveness of quality of supervision during project implementation, and CAS Progress Reports report on evaluation of country program implementation. The self-evaluation in a CASCR is now validated independently by IEG. In addition, country program reviews (CPRs) have become another key instrument to assess the effectiveness of (self-) evaluation approaches during project and country program implementation. They also strengthen incentives and have helped to accelerate a results-oriented training and communications program for management and staff to encourage use of M&E information. During fiscal 2006 OPCS issued a review note of some 50 country CPRs providing guidance to teams on modalities and current practice of results-oriented country program reviews. The note examines the evolution of Country Portfolio Performance Reviews (CPPRs) and their growing importance as a management tool and a tool for policy dialog. Going forward OPCS will help country teams strengthen CPRs through guidance to individual teams and facilitating cross-country learning. • Management is assessing the experience to date with several innovations in the application of the rapid results approach to accelerating implementation of the results agenda. That assessment, to be completed by December, will then feed into updating the results learning program.

IEG recommendation	Management response
<p>2. Provide budgetary support to task teams and technical advice to countries that intend to institutionalize M&E systems to strengthen the collection and use of performance information in order to enhance the countries' capacity and demand to manage for results.</p>	<p>Management will consider its agreed actions complete with the functioning of the RSG (accomplished), the implementation of the changes noted above for global programs (in progress), and the successful implementation of the planned learning activities in fiscal 2007. Management will report on the completion of these activities in the context of the regular progress report to CODE in early calendar year 2008.</p> <p>Management does not agree to provide special budgetary support to task teams (see above) but does agree that greater country support will be useful. Management is working to identify ways, likely through partnerships with other donor agencies that are facing analogous challenges, to provide greater funding support to countries that are working to improve their capacity to manage for results. Management will report on progress in the context of informing Executive Directors on the upcoming Roundtable on Results to be held in Hanoi in February 2007.</p>



Chairman's Summary: Committee on Development Effectiveness (CODE)

On July 19, 2006, the Committee on Development Effectiveness (CODE) considered the *2006 Annual Report on Operations Evaluation* prepared by IEG and the draft management response.

Background

The 2005 AROE assessed the use and usefulness of monitoring and evaluation information to improve results and enhance Bank effectiveness at the country level. In the Management Action Record (MAR), one of the recommendations for management was to establish clear lines of accountability for the follow-up and oversight of adopted recommendations and to make management responses more action specific to include timelines for implementation. For IEG, the recommendation was to institute a formal system for tracking and managing its recommendations. The Committee supported the proposed changes to the MAR, including a pilot three-year expiration criterion for MAR recommendations from fiscal 2006 onward. On March 8, 2006, CODE discussed two draft documents prepared by management that are relevant to the AROE: *Accelerating the Results Agenda: Progress and Next Steps* (World Bank 2006b) and "Updating the Bank's Operational Policy on Monitoring and Evaluation" (World Bank 2005d).

Report Findings

The 2006 AROE found that progress had been made on the 2003 action plan. Procedures were in place and results frameworks were being used in projects and RBCASs, but their effectiveness as a monitoring tool remained in doubt, and more work was needed at the sector and global program levels. Bank managers and staff faced challenges in managing for results such as conceptual difficulties in developing results chains, lack of client capacity and interest, competing priorities in the Bank, and lack of appreciation for the use of performance information and evaluations for management decision making. Moreover, disconnects in outcome ratings between country and sector policy outcomes and project outcomes pointed to a gap in understanding contributing factors to results. The AROE recommended establishing fully funded priority actions. Assessing IEG's effectiveness, the report recommended that IEG's results framework should continue, and even enhance, its focus on the use of evaluation findings.

Management Response

Management broadly agreed with the main thrust of IEG recommendations, which are in line with its own findings and the priorities for action set out in its February 2006 update to

CODE, with one major exception. It noted that a results focus was the responsibility of country teams and task teams and needed to be included in the formulation of RBCASs and in task preparation and task supervision budgets.

Overall Conclusions

Speakers welcomed the opportunity to review management's and IEG's actions to strengthen Managing for Development Results. They noted the findings that the Bank has strengthened the results focus of its operations, while major challenges lie ahead. The Committee was gratified to observe a broad agreement between IEG and management on the main findings and recommendations, which it also generally endorsed. It also accepted IEG's clarification that it was not suggesting additional budgetary allocations but rather stronger prioritization in deployment of resources, a position closer to what was enunciated by management in its response. There were also a few questions regarding specific statistics reported, including some from the client survey and the MAR.

Several members concurred with IEG that accountability was as important as incentives and resources are to promoting the results agenda but also to achieve better results. Members also endorsed greater IEG efforts to enhance the usefulness and application of IEG products, particularly by operational staff. In this regard, the timing of Country Assistance Evaluations (CAEs) vis-à-vis the CAS was cited as an area where some improvement might be possible. A number of speakers raised questions about the effectiveness of the results architecture being put in place, including prioritization of areas and statistical capacity. The role of the Board in raising results issues when discussing Bank's operations was highlighted.

Next Steps

There was agreement about a two-year cycle for AROE; the report will be considered by the Committee, followed by a Board discussion. Given the likely Board agenda in August, the Committee recommended that the Board consider this year the AROE and the draft management response on a streamlined

procedure (that is, without discussion). (Initially a Board meeting was scheduled for August 22).

The following main issues were raised during the meeting:

Results architecture and results framework. The Committee welcomed the progress of the Bank in implementing the results agenda. Yet several speakers had questions about the effectiveness of the results architecture being established. A member expressed concerns about possible distortions favoring more measurable activities and against more complex priority areas such as governance or rural development. The robustness of the system of indicators, given weak statistical capacity in many countries, was also questioned. A member stressed that countries were at the center of the effort to strengthen the focus on results and that a substantive demand from governments in this respect was instrumental. *Management noted that there was limited demand from governments due in part to data gaps at the country level (yet these gaps would not be so expensive to fix) and lack of relevance in the decision-making process. Management found that there was a risk of distortions to only wanting to do things that are measurable.*

Some speakers commented on "disconnects" between country and sector policy outcomes and project outcomes. In this regard, efforts to focus on results were perceived as uneven and inconsistent among and within countries, Regions and networks, rendering it difficult to assess the Bank's overall contribution to development results. Two members noted that frameworks appeared to be better developed in International Development Association countries than in International Bank for Reconstruction and Development countries, although issues of quantity and quality of data were raised. *IEG noted that some reasons for "disconnects" and perception of inconsistency were the difficulty of indicators and measurement due to the political economy of reform or the problem of attribution of a single straightforward intervention on a range of outcomes. In*

response to comments about the importance of harmonization, *management noted that the multilateral development banks were entering the second year of doing a comparative assessment of their respective works on results (COMPAS)*. One member urged management to redouble their efforts to monitor and evaluate results of trust fund activities. *Management stated that they are working with trust fund and global programs to increase their results focus.*

Members stressed the need for staff to better understand the results framework as a tool to link Bank operations with outcomes in country programs and projects, and to strengthen client capacity to manage for results, particularly low-income countries. Further analyses on project outcomes, baselines, and targets, especially in RBCASs and Poverty Reduction Strategies, were suggested. In this regard, consideration should also be given to time (time needed to implement the results framework and the time lag between project completion and outcome). *Staff from the Africa Region provided information about efforts in promoting the results agenda and building capacities on results.*

Organizational issues and incentives. Members commented on the need to change the organizational culture, align incentives and reward structure, and disseminate successful experiences for learning purposes. *Management replied that the current culture, in which staff perceive greater incentive to report up than across, will be addressed by the development of the Results Monitoring and Learning System.* The need for a better communication strategy was also mentioned. More analysis was sought about overload, time pressures, availability of resources, and difficulties in developing the results chain. Speakers and management shared the Director-General, Evaluation's views regarding the key role of accountability in implementing the results agenda. One member sought information on recent activities of the RSG. *Management responded that the group has started to function and is expected to play a key role in increasing attention Bank-wide to results and in communicating and consolidating gains across the Bank.*

Funding priorities. IEG proposed designated funding for implementing the results agenda differentiating new allocations and redeployed resources. Some members commented on special budgetary support to task teams. The Committee took note of IEG's clarification that it was not asking for additional budgetary allocation but rather stronger prioritization in deployment of resources.

Recommendations for IEG. The Committee welcomed the AROE's recommendation on the role of IEG in promoting the results agenda. There were other comments on the need to improve timeliness of IEG evaluations, including timing of CAE and CASCRs; to analyze the relative strength of results frameworks by region and sector, particularly to help operationalize the findings; and to track and publish the results of sectoral, Regional, or project analysis as well as their trends to facilitate learning. One member felt IEG needed to focus on new and ongoing operations. Another member commended IEG for the efforts to develop its own results framework and noting the difficulties of constructing measures that link final outcomes to IEG interventions. Regarding the client survey, one member expressed concern about the perceived decline in the quality and depth of IEG's analysis and recommended that IEG analyze the trend. *IEG clarified that surveys did not show such trends in regard to external clients but that this was an internal issue that had to be sorted out.*

Regarding the content of the AROE itself, a member made some suggestions to add some flesh to the report's substance and stressed the need to use the AROE to help promote "out-of-the-box" thinking on the results focus by exploring such topics as how to better link the evaluation findings to administrative budget and human resources management and how the increased use of innovative work and experimentation—for example, conditional cash transfer—can foster the results agenda. *IEG promised a more specific AROE in two years and took note of the suggestions.*

Management Action Record. A member expressed concern about the statistic that the adoption of six percent of IEG's recommendations was negligible. This member also warned against the three-year expiration criterion acting as an incentive for delaying implementation of the recommendations. *IEG explained that this figure reflected the relatively high number of new recommendations entered in the MAR in 2006.* Members took note that a summary of the MAR will be posted in the Bank's external Web site for disclosure.

Pietro Veglio
Chairman

Chapter 1: Evaluation Highlights

- The Bank has strengthened the results focus of its operations.
- This strengthening has implications for the way the Bank monitors and evaluates its operations.
- Results frameworks at the project and country levels form the basis for monitoring and evaluating results.
- Independent Evaluation Group evaluations have strengthened their assessments of results-oriented monitoring and evaluation.



Introduction

The *Annual Report on Operations Evaluation* (AROE) fulfills the mandate of the Independent Evaluation Group-World Bank (IEG-WB) to assess the progress, status, and prospects for monitoring and evaluating the development effectiveness of World Bank activities. The objective of the 2006 AROE is to assess the extent to which the Bank's monitoring and evaluation (M&E) systems provide Bank staff with the information they need to better manage for results.

Using M&E Information for Results

The Bank has been strengthening the results focus of its operations in recent years. The 2002 International Conference on Financing for Development in Monterrey and the joint statement issued by the heads of the multilateral development banks (MDBs) highlighted the need to better measure, monitor, and manage for results.

In 2003, the Bank started implementation of its Managing for Development Results Action Plan. Within the Bank, a key element is the Results-Based Country Assistance Strategy (RBCAS), which was piloted and then mainstreamed in 2005. The RBCAS includes a results framework that links country development goals with Bank operations and forms the basis for Country Assistance Strategy (CAS) M&E. The Bank also introduced results frameworks into projects and sector strategies: staff outline their expected outcomes and explain how they are to be achieved.

The introduction of *The results-based Country Assistance Strategy has been mainstreamed since 2005.*

results frameworks at the project and country levels has implications for the way the Bank monitors and evaluates its operations. Monitoring would include outcomes as well as outputs and inputs. The purpose of these results frameworks and M&E systems is to provide information that can drive management decisions. That is, what effect is the project or country program having on the planned outcomes? This would affect the way the Bank collects, analyzes, and uses monitoring information.

IEG's evaluations have been strengthening their assessments of results-oriented M&E. More targeted attention is being given to building evaluation capacity among in-

Results frameworks are now used at the country and project levels.

ternal and external clients. Effective communication of IEG's knowledge is a critical factor in increasing the use and usefulness of IEG products, and IEG is making progress in increasing awareness of its products.

Overview of the 2006 AROE

The 2006 AROE assesses the Bank's results focus in M&E. Chapter 2 discusses the implications of the Bank's results agenda on monitoring and self-evaluation of Bank operations. Chapter 3 reviews IEG-WB's effectiveness, presents a results framework for IEG, and examines how IEG can increase the use and usefulness of its evaluations. Chapter 4 summarizes the findings and presents recommendations for Bank management and IEG-WB.

The 2006 AROE draws on reports from Operations Policy and Country Services (OPCS) and the Quality Assurance Group (QAG), as well as IEG evaluations and a review of recently approved RBCAS. The evaluation team also conducted focus group interviews with Bank sector managers, task team leaders, and country coordinators; structured interviews with country directors and members of the Board's Committee on Development Effectiveness (CODE); and a client survey of key internal and external audience groups for IEG's recently completed evaluation studies. The findings are included as appendixes to this report.

Chapter 2: Evaluation Highlights

- RBCASs include country results frameworks that clarify the Bank's contribution to achieving country goals.
- Results frameworks are being used in projects and RBCASs, but their effectiveness as a monitoring tool remains in doubt.
- Challenges to results-based monitoring include the following:
 - Perception that monitoring is about inputs and outputs rather than results
 - Conceptual difficulties in developing results chains
 - Lack of client capacity and interest
 - Competing priorities in the Bank
 - Lack of appreciation of the utility of performance information and evaluations in management decision making.
- A greater focus on results increases the need for self-evaluation during implementation.



Monitoring and Evaluating Results in Bank Operations

This chapter reviews the implications of the Bank's results agenda on monitoring and for the self-evaluation of Bank operations. A descriptive overview of the Bank's evolving M&E framework at the activity (lending and analytical and advisory activities), country, sector, and global levels is provided in appendix A. This chapter focuses on Bank lending operations and the CAS—two areas where the Bank has systematically introduced a strengthened results focus over the past few years.

To the extent that global programs with Bank involvement produce results at the country level, they can be considered one of the instruments in the Bank's country programs. Evaluations are done both by Bank operational staff (self-evaluations) and by IEG. IEG reviews and validates the Bank's self-evaluations and conducts its own independent evaluations. This chapter covers the Bank's self-evaluation; IEG activities are discussed separately in chapter 3.

Recent Developments in Monitoring

The Bank has taken a number of steps over the past two years to strengthen the links between country development goals and individual Bank operations. At the project level, the Bank introduced results frameworks for its investment projects in 2004. A results framework focuses on the project development objective and intermediate outcomes, both of which are to be supported by performance indicators. These indicators would be used to track progress

toward meeting the development objective and to make changes in the project, if necessary, during implementation. The Bank introduced results frameworks into its adjustment operations when it issued its new policy on Development Policy Loans in 2004.

The Implementation Status and Results (ISR) Report was introduced in 2005 to replace the Project Supervision Report (PSR). One of the objectives in introducing the ISR was to strengthen the focus on outcomes. Monitoring is now given more prominence by making the performance indicators part of the main report; previously it was an annex.

Task teams are required to specify the expected development outcomes for economic and sector work (ESW), nonlending technical

A results chain links inputs to outputs and then to intermediate and final outcomes.

assistance, and other analytical and advisory activities (AAA), although they are not required to establish results frameworks for them.

At the country level, RBCASs were adopted as policy in 2005. Critical elements of the RBCAS are: (a) the country results framework, which specifies country development goals and outlines the results chains that link individual goals with CAS outcomes, intermediate outcomes and milestones, and Bank operations;

Results frameworks focus on objectives and intermediate outcomes and are supported by indicators that track progress.

and (b) the CAS Completion Report (CASCR), a self-evaluation of the CAS that is validated by IEG. The country results frameworks attempt to clarify the Bank's contribution to achieving country development goals, thereby addressing a weakness found in previous CASs, which discussed the country's

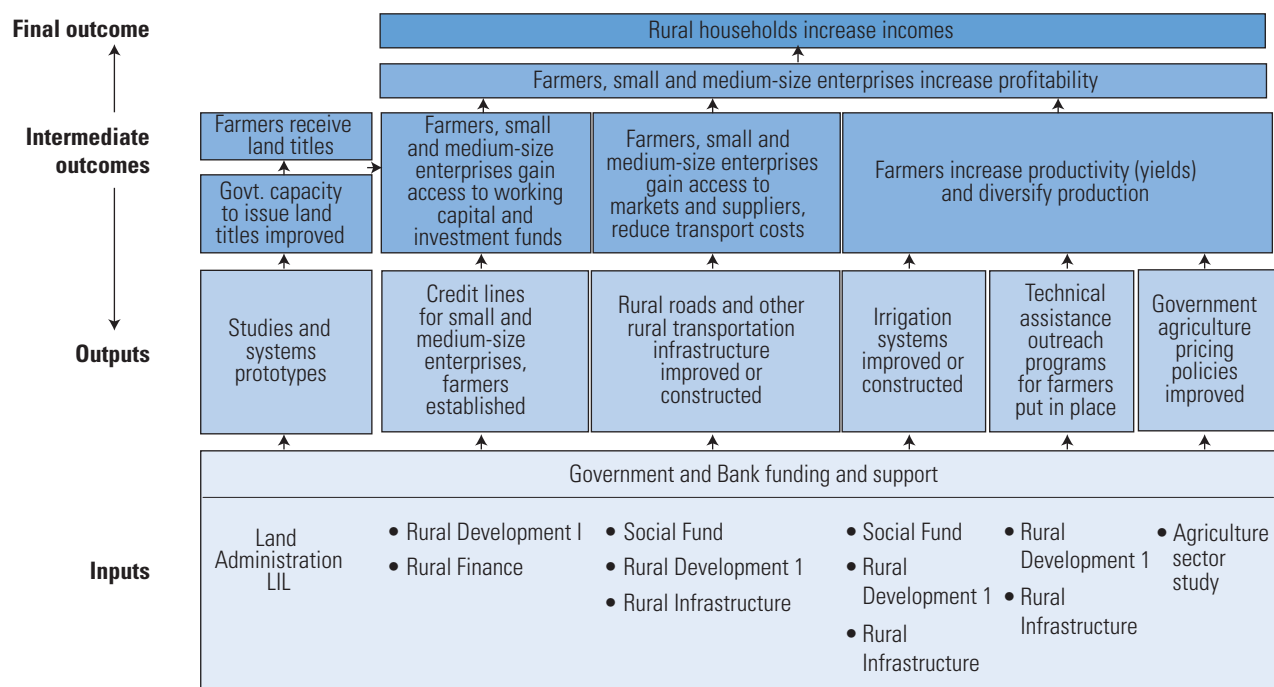
objectives and Bank operations in detail but did not establish the links between the two.

How a Results Chain Works

Figure 2.1 shows a sample country-level results chain for agricultural development, outlining the expected links between Bank projects, project outputs, intermediate outcomes, final outcomes, and the country development goal.¹ It also shows the results chain for individual operations.

For example, Bank funds channeled through a rural development project (input) would produce improved or new rural roads (output). This, in turn, would improve farmers' access to markets and suppliers (intermediate outcome 1), increase profitability (intermediate outcome 2), and improve rural household incomes (final outcome). The results chain outlines how the project is expected to contribute to the overall goal of improving rural incomes and reducing poverty (the country development goal).

Figure 2.1: Sample Results Chain for Agricultural Development



Source: Adapted from material from the Bank East Asia and Pacific Region.

Note: LIL = Learning and Innovation Loan.

The example in figure 2.1 shows that a single project could have several outputs and contribute to multiple intermediate outcomes. For example, the same rural development project could include a financing component that would provide farmers and small and medium-size rural enterprises with access to credit through formal financial institutions at terms more favorable than they would have been able to obtain from informal lenders.

Two or more projects could share common outputs and outcomes. For example, the Bank's country portfolio may include a rural credit project that would address wider issues of financial sector reform and financial institution building nationwide. The rural credit and rural development projects may or may not overlap, because the latter would be focused on specific areas of the country. But it does point to the need to ensure that both projects are well coordinated—for example, by sharing common outcomes and performance indicators. It also provides opportunities to compare approaches and learn about what works and what does not and to identify “good practices.”

The example in figure 2.1 also shows possible synergies between Bank projects that contribute to a common intermediate outcome. For example, financial institutions may require a formal title as collateral to provide financing to farmers and small and medium-size enterprises. The land administration project, which aims to strengthen the government's capacity to issue titles, would complement the rural finance operations. Alternatively, if it proves difficult to provide titles, other options for rural financing may need to be explored.

Figure 2.1 shows another case of synergies where different projects provide target beneficiary communities (in different geographic areas) with access to support for infrastructure development and extension services. Together with an agriculture sector study, they would contribute to increasing farmer productivity (one of the intermediate outcomes).

Developing results chains for individual operations often proves to be a complex undertaking, because there may be multiple causal chains

and sequencing patterns between Bank operations and CAS outcomes, with many intermediate steps. But simplified country-level results chains, such as the example shown, allow a country team to identify synergies and gaps in its operations and to establish common intermediate outcomes to which team members may jointly contribute. Where other development partners may be involved, the results chains also help identify synergies and areas of common interest.

Possible Uses of Results Frameworks

Country directors and task team leaders (TTLs) can use the results frameworks to establish the extent to which the Bank and individual Bank operations can be held accountable for producing results. That is, the country directors and TTLs make their accountabilities explicit through their choices of CAS outcomes and intermediate outcomes in the CAS results frameworks and the project development objectives and intermediate outcomes in the project results framework.

There are trade-offs. If the outcomes set are too modest, country directors and TTLs may be seen as setting the bar too low. If the outcomes are too ambitious, they may be criticized for lacking realism, because more factors come into play in

achieving the expected outcome, which makes attribution and Bank accountability unclear. Country directors and TTLs have to balance the ease of achieving successful outcomes with the risks involved in addressing important but complex issues, such as governance, or working in weaker policy and institutional environments, such as fragile states.

Nonlending activities also define expected outcomes but are not required to have results frameworks.

Results chains can be used to identify synergies between Bank projects.

The frameworks can be used to establish accountability.

The results frameworks include performance indicators—with baselines and targets—for each outcome that can be used to monitor progress and assess success or failure on completion of the Bank operation or country program. At the project level, the Bank has increased the importance and visibility of these performance indicators by giving them more prominence in the ISR (which replaced the PSR).²

The focus on outcomes increases the amount of information needed for M&E.

The focus on intermediate and final outcomes increases the amount of information to be collected during implementation and to be used for

M&E. In addition to inputs (such as Bank loan disbursements) and outputs (such as the length of roads constructed), the borrower will be required to monitor progress toward achieving agreed outcomes through the results frameworks (for example, the effect the roads have had on farmers and small and medium-size rural enterprises). While information on inputs and outputs may be available in house, collecting outcome information may require additional expenditures for surveys or other methods to assess the effect of the project on beneficiaries.

Information on what is happening or not happening to expected outcomes would help the government and Bank managers and staff determine where to focus their efforts. Decision making in these situations is often complex.

A lack of progress could lead to the provision of additional resources to correct problems, or to a reduction in resources to cut losses. Faced with such challenges, managers and staff may need to use evaluative approaches, such as rapid assessment methods, to get a

Such information is useful in determining where to focus additional effort.

better understanding of why progress toward outcomes is not being achieved as planned. This could lead to more

decisions, such as to restructure a project or program.

In sum, managers and staff in the borrower countries and the Bank could use the results frameworks as a management tool to do the following:

- Outline how their operations would contribute to achieving country-level objectives.
- Communicate priorities and clarify expectations and accountabilities with superiors and external stakeholders.
- Motivate their staff and teams by setting challenging but realistic goals.
- Improve coordination by identifying common outcomes to which other Bank operations and development-partner programs could contribute.

In addition, Bank country directors could use results frameworks to identify synergies and gaps in the Bank's country program. They could explore better integration between country and project results frameworks by sharing common outcomes and performance indicators where appropriate.

With baselines and targets, the results frameworks would provide a monitoring system that could act as an early warning system to alert managers and staff to the risk of not achieving project and country outcomes. When expected outcomes are not being achieved, self-evaluations could be used to explore the reasons for problems and lead to corrective actions. The implications of the Bank's results agenda for M&E are discussed in more detail in appendix B.

Current Status of Monitoring Results

Results frameworks are being incorporated into projects and RBCASs, but their effectiveness as a monitoring tool remains in doubt. At the **project level**, QAG found that consistency with country and sector strategies was strong in projects approved in fiscal 2004–05, but there was scope for improving the clarity and realism of development objectives in about 40 percent of the projects reviewed (World Bank 2006d). In addition, many projects (40–50 percent of the sample) lacked well-articulated results frame-

works with baselines and systems for tracking and reporting progress.

Similar weaknesses were observed in **the results frameworks of ISRs**. An OPCS review of the quality of results frameworks in ISRs prepared in fiscal 2005 found that 40 percent did not contain any outcome indicators with baseline data for either the project development objective or intermediate outcome indicators, and were not considered to be acceptable (World Bank 2006b). The review noted that there was little evidence that the ISR is used as a document of record for implementation or results reporting. A review of ISRs by selected networks found that in fiscal 2005 many lacked baseline data for outcomes—only 52 percent in rural development had such data, 50 percent in public sector governance or rule of law, 47 percent in health, and 32 percent in water supply and sanitation—which limited their usefulness for monitoring progress. The ISR was introduced in 2005, so these issues may reflect problems of the transition from the PSR to the ISR.

At the **country level**, 25 CAS results frameworks prepared during fiscal 2005–06 were reviewed in preparation for this AROE to assess their quality and usefulness as a tool to manage Bank country programs. A four-point rating system, with level 0 the weakest and level 3 the strongest, was used to assess:

- The strength of the links among country development goals, CAS outcomes, milestones, and the Bank program
- The usefulness of the performance measures for M&E (see box 2.1 for an example).

The analysis confirmed weaknesses similar to those at the project level: poorly articulated results chains and lack of indicators with baselines and targets. These weaknesses would lessen the usefulness of the results frameworks as country program monitoring and management tools. The results of the analysis are presented in appendix C.

The review found that results chains are better developed in International Development Associ-

ation (IDA) countries than in International Bank for Reconstruction and Development (IBRD) countries. Overall, about one-quarter (24 percent) of the CASs reviewed have strong (level 3) and another 44 percent have moderate results chains (level 2). While 85 percent of CASs in IDA countries have moderate to strong overall results chains (levels 2 and 3), only 50 percent of CASs for IBRD countries do (figure 2.2).

Countries with stronger results chains have CAS outcomes and milestones that are better formulated. One factor that may contribute to this difference is that IDA countries have experience working with Poverty Reduction Strategies (PRSs), where country-level performance measures are set and tracked as part of the overall development assistance program. In IBRD countries, where the Bank may account for a smaller share of the overall investment toward a country goal, it may be more difficult to articulate the Bank's contribution in a results framework and to set performance measures and milestones. It should be noted that better plans (results frameworks) do not necessarily translate into better achievement of results without government buy-in.

Lack of adequate baselines and targets continues to be a weakness, as identified in last year's AROE.³ Only 9 percent of the CASs reviewed had baselines for more than half of their CAS outcomes (levels 3 and 4; figure 2.3). Just under half (45 percent) had set targets for more than half of their CAS outcomes. This raises questions about the availability of data and the usefulness of these CAS outcomes for M&E.

The Bank does better at identifying milestones⁴ that are relevant to achieving CAS outcomes,

Frameworks with baselines and targets could act as an early warning system for projects.

The effectiveness of results frameworks remains in doubt.

Frameworks could be improved, and indicators are often inadequate.

Box 2.1: What Constitutes a Strong Results Framework?

A strong country results framework, which would score at least 3 on the four-point rating scale for all criteria, would establish a robust results chain, identifying clear links among:

- (a) Specific country development goals
- (b) CAS outcomes (identifying the nature and extent of the Bank's contribution to achievement of the goal)
- (c) Intermediate outcomes, actions, or outputs (that is, milestones) that serve to mark progress toward achievement of the CAS outcomes
- (d) Bank inputs (its lending and nonlending services).

In addition to these identified links, country development goals and CAS outcomes would be accompanied by performance indicators that are measurable and include baselines and specific, time-bound targets in a strong results framework. Milestones could be actions (such as passage of key legislation), in which case they should be defined in such a way that their relevance to the CAS outcome is clear. Finally, the CAS outcomes and indicators would allow the government and the Bank to focus on key elements of the Bank's country program and conduct a regular dialogue on progress made, based on the performance indicators.

The review identified several CASs with well-developed results frameworks, although none scored high in all categories.

The Burkina Faso CAS results framework, for example, had a strong results chain that described how the CAS outcomes were supported by the activities. It also had a few key milestones to allow for monitoring and management against each CAS outcome.

Performance measures were included for country goals and CAS outcomes. Most were well constructed and included baselines and targets. For example, a CAS outcome of "reduced risk and increased revenues for rural households" is supported by four performance indicators, one of which is "land tenure: 50 percent of male applicants and 50 percent of women applicants obtained land titles or equivalent in six pilot provinces." A milestone of land tenure is "revised implementing decrees for property management are adopted."

However, the Burkina Faso CAS results framework had 54 country goals and 65 CAS outcomes. These numbers could limit its usefulness in managing the country program.

Poor results frameworks—those scoring at level 1 in many of the categories—had poorly developed results chains. For example, milestones were included as country or CAS outcomes or, conversely, higher-level CAS outcomes were included as milestones. Country and CAS outcomes were poorly defined and not measurable, such as "involve all sectors at all levels to scrutinize government projects." Where performance indicators were identified, many did not include baselines or targets, so they could not be used to track progress.

although their precision could be improved (that is, they could describe more clearly what is to be achieved; see figure 2.4). Only 36 percent of the CASs, however, had specific, time-bound targets for more than half of their milestones.

Large numbers of indicators, currently common, pose problems for data collection and analysis.

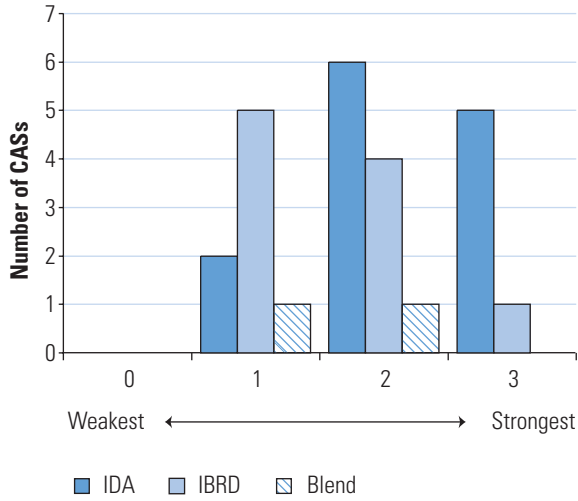
The number of indicators remains high. On average, there were 24 country development goals, 41 CAS outcomes, and 60 milestones per country. Although there is no correct number of indicators, large numbers pose problems for data collection and analysis. A smaller number of indicators would not only be more manageable

but could also help focus the dialogue with the government on the critical elements of the country program.

A comparison of the number of CAS outcomes with the fiscal 2006 country budget (as a proxy for program size) showed that the larger programs had more indicators (figure 2.5). But it also identified some outliers that would be candidates for review and simplification. A large number of CAS outcomes relative to the overall size of the country program may indicate a need for greater selectivity.

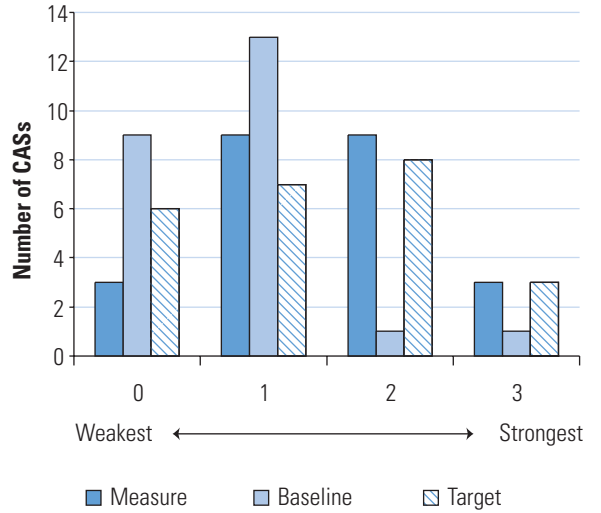
When results frameworks are developed, either at the project or the country level, a conflict may arise between *managing for results*, which

Figure 2.2: Results Chains Better Developed in IDA Countries than in IBRD or Blend Borrowers



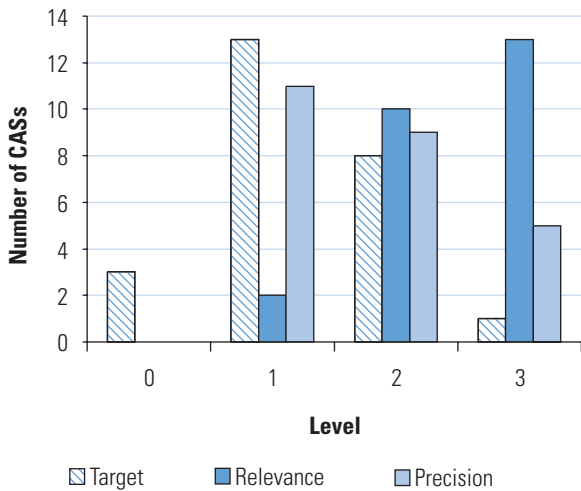
Source: IEG data.
Note: CAS = Country Assistance Strategy.

Figure 2.3: Country Assistance Strategies Tend to Lack Adequate Baselines and Targets



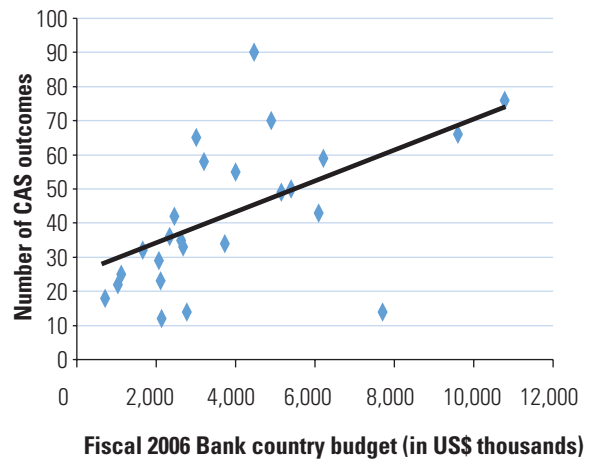
Source: IEG data.
Note: CAS = Country Assistance Strategy.

Figure 2.4: Precision of Milestones Could Be Improved



Source: IEG data.
Note: CAS = Country Assistance Strategy.

Figure 2.5: Larger Programs Tend to Have More Indicators, Outliers May Need Review



Source: IEG data.
Note: CAS = Country Assistance Strategy.

emphasizes learning on what results are being achieved and why and using this information for decision making, and **accountability for results**, where performance information is used to hold agencies, managers, or staff accountable (Binnendijk 2001). A balance needs to be struck between risk taking—tackling complex, challenging issues such as governance—and the likelihood of success, that is, focusing on outcomes that are achievable.

Similarly, too strong an emphasis on measurement may lead to projects and country programs focusing on areas that can be measured. For example, the ready availability of infant mortality data should not be a determining factor in where the Bank focuses its intervention in a country's health sector.

These challenges tend to be country-specific. Data availability would differ significantly between a middle-income country and a fragile state. Similarly, what could be considered an easily achievable outcome in a middle-income country may pose a significant challenge in a fragile state. This reconfirms the importance of the country as the unit of account in Bank operations and the need to strengthen CAS results frameworks.

Challenges to Developing Results-Based Monitoring

The 2006 AROE conducted focus groups and structured interviews with 66 Bank country directors, country coordinators, sector managers, and TTLs to examine why Bank staff find it difficult to develop a results-based monitoring system (results framework) and use it to manage for results. The key findings are summarized in box 2.2, and additional information on the focus groups and interviews is available in appendix D.

Focus group participants reported difficulties in developing results chains. Often they set ambitious project development objectives and had difficulty linking

Focus groups reported difficulties in developing results chains.

them to the projects: *“Once you are unrealistic in project conception, everything that follows*

from it is also overstated.” Participants also felt overburdened by the emphasis on quantitative data: *“There is a tyranny of the measurable here that we have to be very careful of.”* QAG also noted that M&E problems are frequently related to difficulties in defining and monitoring a set of outcome/impact indicators.⁵

Participants in the 2006 AROE focus groups mentioned lack of resources as a constraint to developing M&E systems during project preparation, establishing baselines, and systematically collecting data during project implementation. Insufficient budgets were often cited as an issue. In a previous QAG assessment of supervision in fiscal 2001–02, only 31 percent of TTLs responding felt that they had sufficient resources to adequately implement an outcome-focused system for supervision (World Bank 2003b).

A review of Regional budgetary expenditures for fiscal 2002–05 found that Bank budget expenditures for supervision within the Regions had increased in constant terms in line with the increase in the Regional budget (table 2.1). Supervision received more resources than either lending or country and ESW and accounted for about 27 percent of total client expenditures within the Regions. Supervision expenditure per project increased in real terms during the period.

This suggests that the issue may not be the lack of Bank budget for M&E per se, but the many demands on both financial resources and the TTL's time and attention. Participants explained that they often postponed the preparation of M&E arrangements until after project approval because of the need to reduce project preparation time:

You need to do it (develop monitoring and evaluation system) quickly, you need to do it before (appraisal), and then you are running out of time because there are all these pressures to deliver.

Some focus group participants noted that monitoring gets less priority during implementation:

Box 2.2: What Are We Hearing from Managers and Staff?**Managing for Results and Monitoring**

Judging by the focus groups, managers and staff seem to understand the overall objectives of the Bank's Managing for Development Results agenda and its implications for Bank operations. Managing for results means aligning project and country program outcomes with country objectives, focusing on outcomes rather than outputs, and measuring and being accountable for outcomes.

Based on interview findings, policy makers (Board members, country directors) see the results frameworks as a way of setting overall strategic directions and monitoring progress toward these strategic goals. TTLs accept the results frameworks as part of their work, but disagree about their usefulness. Country directors, sector managers, and TTLs struggled with technical aspects of managing for results—developing results chains, positioning Bank operations appropriately in the chain, and coming up with performance indicators.

The focus group identified the following challenges to managing for results:

- **Perceptions:** Many people still see implementation and monitoring as being about inputs and outputs. That is, results will come at the end of an operation. M&E is still seen as generating data, especially outcomes, whose use and usefulness are not necessarily clear or seen as a high priority.
- **Conceptual difficulties:** Managers and staff struggle to develop results chains—that is, to link higher-level objectives to Bank operations and identify appropriate performance indicators. They also have difficulty positioning a project within the chain. They may come up with overly ambitious project development objectives (PDO) or position a PDO too high up the results chain, and then struggle to develop a results framework around that PDO. They also had difficulty coming up with numerical indicators, commenting that not all important results could be measured.
- **Client capacity and buy in:** Lack of client capacity and interest is often given as a constraint to carrying out M&E. It is linked with lack of clarity on the use and usefulness of monitoring data, especially of outcomes, during implementation. This has led to difficulty in collecting data, including baseline data.
- **Bank pressures:** TTL overload, time pressures, and lack of resources were cited as constraints. TTLs tend to see M&E as a lower priority than operational activities (and also more problematic, given difficulties with results chains, indicators, lack of clarity on use, and client buy-in). When developing a project, M&E was often left until the last minute, given the pressure to reduce project preparation time.

Source: Focus groups and structured interviews (appendix D).

The rationing of resources begins to happen here [during implementation]. Monitoring and evaluation is an expensive thing, so the monitoring and evaluation guy gets bumped off the train.

The Bank's fiscal 2005 budget document⁶ identified safeguards and fiduciary requirements as cost drivers that lead to increased resource needs within the Bank.

A closely related issue is the lack of capacity and the cost to the borrower of developing systems to collect the necessary performance data and the borrower's reluctance to bear that cost. Increasing the focus on monitoring outcomes during implementation may require new ways of collecting data and add to borrower costs.

Participants in the focus groups saw lack of borrower buy-in as one of the key constraints to developing and using an effective outcome-based monitoring system:

Focus group participants pointed to a lack of resources as a constraint.

All clients want is to get the job done. They should do monitoring and evaluation, but it is not their priority because they all know that their job ends with the successful completion of the project, and then they want to go to the next project.

Participants also recognized the need to create demand for monitoring information and to build borrower capacity:

The issue may be more a matter of competing demands on staff time and attention.

Table 2.1: Regional Client Services

	Fiscal year			
	2002	2003	2004	2005
	Amount (US\$ millions, constant fiscal 2005 prices)			
Project supervision	153.5	159.7	163.3	168.7
Lending	124.2	118.6	151.8	138.9
Country, economic and sector work	131.6	140.0	148.2	149.1
Other country services	59.0	61.7	69.9	66.0
Sector and global services	81.5	87.0	91.9	91.6
Total client services	549.9	566.8	625.2	614.2
	Share (percent)			
Project supervision	27.9	28.2	26.1	27.5
Lending	22.6	20.9	24.3	22.6
Country, economic and sector work	23.9	24.7	23.7	24.3
Subtotal	74.4	73.8	74.1	74.4
Supervision/project ^a (US\$ thousands)	75	78	79	82

Source: World Bank data.

Note: Direct costs. The Bank moved to Systems Applications and Products, a new accounting system, in fiscal 2000, which resulted in some data problems for fiscal 2000 and 2001, so only four years (fiscal 2002–05) are presented.

a. Supervision cost for IBRD/IDA and special (West Bank and Gaza, East Timor) projects. Excludes Global Environment Facility and other projects.

I have managed to convince clients to monitor. I have one client who became creative about coming up with new indicators and using them to convince the prime minister to back her on certain things.

Finally, lack of incentives is often given as a constraint to managing for results. The 2004 AROE (IEG 2004b) noted that “changes in guidelines and templates are important, but not sufficient for ensuring a systematic adoption of a result-oriented approach in an organization.” It recommended, among other measures, that Bank management “review and strengthen incentives such as managerial recognition, career advancement, and reward systems to promote results-oriented monitoring and evaluation, managing for results and lesson-learning.”

Formal incentives to pursue outcomes have improved, but more could be done informally.

The Bank has taken steps to incorporate results explicitly into its management policies. The new staff-promotion criteria for level GH

require a “sustained record of achieving operational impact/results on the ground.” Performance on portfolio implementation is an element in the evaluation of all operational managers. The achievement of results is one of several criteria that go into assessing performance. Further analysis is required to establish the extent to which these policies have affected staff and manager behavior.

While formal incentives may be in place, informal incentives may wield stronger influence. Focus group participants mentioned the pressures to deliver projects as quickly as possible. OPCS also noted that “the consistent message to task teams is that delivery implies delivering a project to the Board rather than delivering results for clients” (World Bank 2006b).

Messages from Bank senior management are still seen as reinforcing this perception. For example, an internal Bank news article on fiscal 2005 operational results highlighted the increase in IBRD/IDA loans, grants, and guaran-

tees (which were 10 percent higher than in fiscal 2004 and reached a six-year high) and compared the numbers of ESW products delivered and technical assistance activities completed between fiscal 2004 and 2005.

In addition, a lack of understanding of the possible use and usefulness of performance information may be acting as a disincentive to use monitoring to achieve project and program outcomes. Many focus group participants still see results as something that comes at the end of an operation, and implementation as primarily about project administration (or focusing on ensuring that inputs are made to produce outputs according to the agreed schedule). To many, the Bank's Managing for Development Results agenda is seen as an add-on or "overhead" that requires the generation of more information, and the utility of that information is unclear:

Some things are not easy to measure, and we try to force things into a measurement box that isn't always appropriate, so it becomes the "feeding the beast" syndrome—you are doing it because the institution demands it.

Self-Evaluation and the Bank's Results Agenda

Traditionally, monitoring has been concerned with implementation, compliance with schedules, and the tracking of inputs and outputs, while evaluation has focused on assessing whether the project achieved its stated objectives. Focus groups held as part of the 2006 AROE found that many Bank staff share this view of M&E. Many equate monitoring with the ISR and evaluation with the Implementation Completion Report (ICR)—in other words, an event that is scheduled at the end of project when the "results" are clear.

The increasing focus on intermediate and final outcomes could change the way the Bank views and conducts self-evaluations. Bank managers and staff could use self-evaluations to better understand why progress may not be made in

achieving certain project or program outcomes during implementation. Evaluative approaches such as program chain logic assessment⁷ could

also be carried out prior to the start of a project or program to determine the strength and logic of the underlying conceptual model.

Although self-evaluations would continue to be carried out at the completion of a project, it may not be sufficient to examine whether a project was able to meet its objectives. It may also be necessary to address the broader question of what factors contributed to changes at the beneficiary level and to what extent those changes are attributable to the project. This would require more sophisticated methods, such as impact evaluation.⁸

Some teams have begun to use a heightened focus on project outcomes coupled with self-evaluation to improve performance during project implementation. A Rapid Results Approach, for example, calls for the formation of multistakeholder teams to agree on and achieve critical outcomes within a short (usually 100-day) period. This approach has led to a significant increase in the use of a Voluntary Counseling and Testing Center in Eritrea. In an Agricultural Technology Project in Nicaragua, a Rapid Results Approach led to a tripling in the production of high-grade milk over a short period in one municipality (with the approach replicated in others) and to the jump-starting of a program to promote agricultural productivity in another; this latter result was accomplished by the signing of 180 contracts between farmers and technical service providers within 100 days.

Other donor agencies are using evaluative methods to improve implementation of their programs. For example, the independent evaluation unit of the Deutsche Gesellschaft für

Many see the results agenda as an add-on, and its usefulness is not fully understood.

Some project teams have begun using self-evaluation to improve performance during implementation.

Technische Zusammenarbeit (GTZ) has developed an electronic interview and evaluation procedure, e-VAL, that allows evaluators to obtain and analyze stakeholder assessment on the likely success or failure of a project or program and to provide quick feedback to teams

and stakeholders, who can then take action where necessary (box 2.3).

The Bank is working to assess the effect of its interventions on beneficiary outcomes through impact evaluations. Impact evaluation gained

Box 2.3: GTZ’s e-VAL System: Using Computer-Supported Evaluation Tools for Rapid Progress Reviews

In 2003, the GTZ signed a new performance contract with the German Federal Ministry for Economic Cooperation and Development (BMZ) that strengthens results orientation by putting outcomes at center stage.

Under this contract, GTZ assumes joint responsibility with the national partner to achieve outcomes. In turn, implementation is made significantly more flexible: the project contract specifies budget, outcomes, and expected impact, but neither activities nor outputs. This flexibility requires continuous results-oriented adjustments of activities and outputs—and places a premium on the collection and use of real-time evaluation information for decision making.

GTZ’s evaluation system is based on the two pillars of self- and independent external evaluation and fosters both accountability and learning.

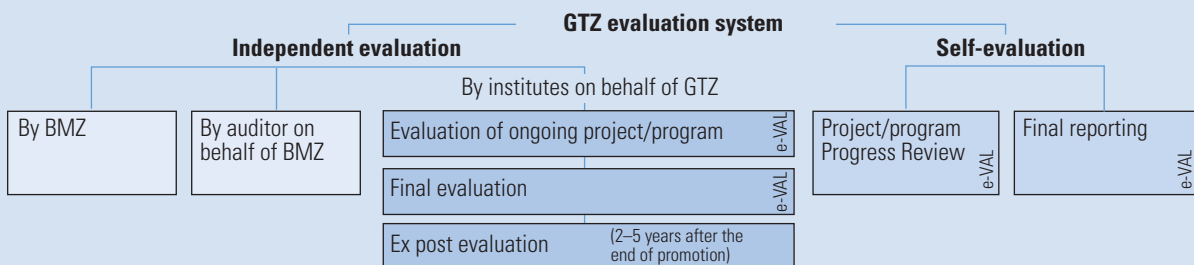
GTZ’s evaluation approach complements the standard Organization for Economic Cooperation and Development–Development Assistance Committee (OECD-DAC) criteria for assessing development results against objective indicators with a new computer-supported evaluation tool (e-VAL). e-VAL captures key qualitative dimensions of ongoing projects/programs, such as *ownership, participation, and target group orientation*. Specially trained interviewers conduct interviews with the project team, government partners, and civil society organizations or target group representatives.

The computer-supported evaluation procedure processes

these interview data and provides quick feedback at relatively low cost. The interviewees are encouraged to describe how they see “success,” “failure,” and the current situation of the project/program against a set of defined elements. The e-VAL software then converts these interview findings into a graphic summary to assess and compare the impressions among the various groups.

Sharing these results in real time increases ownership by fostering a common understanding of the achievements and failures of the program. Moreover, by aggregating the (anonymous) interview data at Regional and sectoral levels, e-VAL allows instant cross-cutting analysis beyond the level of an individual program and enables GTZ’s management to take corrective actions. To date, GTZ has carried out e-VAL assessments in almost 500 projects and programs worldwide.

The e-VAL tool was used as part of the assessment of the Armenia PRS in 2005. Representatives from the Armenian government, the donor community, and nongovernmental organizations (NGOs) were interviewed using e-VAL. The assessment found disagreement among the three groups on the achievements being made under the PRS. The government was most optimistic and the NGOs most pessimistic, but common ground was found on elements related to the implementation of the PRS. These evaluation findings were immediately fed back into the design of the next PRS stage, prompting a rethinking of the roles of the government, donors, and NGOs; greater participation by local NGOs; and emphasis on the social dimensions of poverty reduction.



Source: Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH.
 Note: BMZ = German Federal Ministry for Economic Cooperation and Development.

formal recognition through its establishment as a new AAA product in 2005.⁹ The DIME (Development Impact Evaluation Initiative) by the Bank's Development Economics (DEC) Vice Presidency and Poverty Reduction and Economic Management Thematic Group for Poverty Analysis, Monitoring, and Impact Evaluation provide Operations with training and advice on impact evaluations. An ongoing stocktaking of Bank impact evaluations by DIME has identified 60 ongoing evaluations. Of these, close to two-thirds are in social sector and poverty programs. The Africa and the Latin America and the Caribbean Regions represent 80 percent of the evaluations.

It would be helpful to clearly define the intended uses of impact evaluations, which range from ongoing evaluation during project implementation to extensive ex post meta-evaluations, and deepen the understanding of which interventions work under what circumstances at the beneficiary level. Impact evaluations can be costly and complex to administer. Alternative evaluative approaches, such as cost-benefit analysis, could be complimentary to impact evaluation in informing decisions. IEG evaluations have found that there is a need for more—and more rigorous—project economic analyses.¹⁰ Further research could be done to differentiate which methods could apply in what circumstances.

In sum, the increasing focus on project outcomes during implementation is expected to increase the need for self-evaluation during (and even before) implementation. Some parts of the Bank and other donor agencies are already taking this approach, but it is not yet widely used within the Bank.

Implications for Future Work

The Bank is taking steps to strengthen results frameworks at the project level. QAG and OPCS continue to review and point out issues with project results frameworks. Regions are taking steps to help their staff increase the results orientation of their work.

The Africa Region established an M&E team within its quality assurance unit, and other

Regions are bringing staff who are knowledgeable about managing for results into their quality assurance functions (World Bank 2006b). The Latin America and the Caribbean Region was awarded an IEG Good Practice Award in 2006 for its efforts to promote managing for results by disseminating good practices, conducting impact evaluations in several sectors, and providing technical assistance and support to project teams in designing and improving M&E of project components.

There are opportunities to further strengthen the CAS results framework as a management tool. The framework could be a useful way for the country director to identify and exploit synergies across Bank operations and manage the country program. That is, country directors could look to align projects and AAA around common CAS outcomes through the CAS results framework (the vertical alignment in box 2.4). Bank TTLs would continue to focus on how well their project or program is progressing, but with more attention to outcomes.

A recent review reached a similar conclusion. It recommended, among other measures, that the Bank organize client work into programs (clusters of tasks) focused on country or multicountry results. But incentives continue to discourage working across sectors. The review identified the current emphasis on own-unit deliverables and higher budget predictability for own-managed deliverables as key disincentives to working in multisector teams and recommended replacing the Bank's task-based approach to budgeting and management.

A stronger results focus would require strengthening teamwork within the country team. Country team members would need to come together, identify and agree on a common CAS and intermediate outcomes to serve as the focus of their efforts, determine how best to measure and monitor progress toward these shared

Country directors could make more use of the CAS results framework to identify and exploit synergies among activities.

Box 2.4: Integration of Objectives in Results Frameworks

One way to address the differences between country- and project-level outcomes would be to focus more on the integration of project objectives with country objectives through the results framework. The diagram below offers a schematic of how this might be done in the CAS.

	Country outcome 1	Country outcome 2	Country outcome 3	Country outcome 4
Project A	X			X
Project B		X		X
Study X		X		
Project C	X		X	X
Technical assistance activity				X
Study Y			X	

Efficiency (inputs and outputs) will continue to be a key concern. But more periodic assessment of outcomes is needed.

Focus will be on effectiveness (progress toward meeting CAS objectives and outcomes). More collaboration is needed across task teams and with borrowers and other development partners to develop M&E systems and monitor progress toward country goals.

outcomes, and set up a system to collect the necessary performance information. For example, progress toward a health outcome may require not only health projects, but also transport projects to improve access to health facilities, water and sanitation projects to provide access to safe water and a healthier home environment, and stronger community outreach to improve sanitary habits. Operations that share a common outcome may consider sharing performance indicators.

A focus on country-level outcomes would address concerns about disconnects in outcome ratings between country and sector policy outcomes and project outcomes. An IEG evaluation of Bank country programs (IEG 2004c) found that in a third of the country programs reviewed, aggregate

More teamwork within the country team would also help the results focus.

project outcomes were satisfactory, but the outcomes of the overall country assistance programs were unsuccessful (box 2.5).

Similar disconnects have emerged at the sector level. For example, most primary education projects met their stated expansion targets but failed to look at how well students were learning. Bank advice and lending for trade contributed to systematic reductions in protection and in opening up economies, but had less impact on employment and poverty reduction than expected. And, finally, a fragmented approach to capacity building may have resulted in missed opportunities for more systematic improvements. These disconnects could be addressed by conducting evaluation during and upon completion of Bank activities to improve understanding of critical factors contributing to success. This understanding should help refine the results framework at the country, sector, and project levels.

Buy-in by the client will be essential for better outcome-based M&E at both the project and country levels. The countries see lack of capacity and the additional cost of data collection at the beneficiary level as constraints to monitoring.

Box 2.5: Reducing the Gap Between Country- and Project-Level Outcomes

Several IEG evaluations found that satisfactory outcome ratings at the project level do not necessarily add up to satisfactory outcomes at the sector and country levels. This finding points to a gap in understanding the factors that contribute to country and sector outcomes. To overcome this problem and improve its development effectiveness, the Bank needs to focus more on country-level outcomes and how its interventions align with these outcomes and conduct systematic evaluation during and upon completion of projects to better understand the critical factors contributing to success.

IEG documented the gap between aggregate project outcomes and overall country outcomes in 2004. It found that for 33 percent of Country Assistance Evaluations (CAEs) prepared between fiscal 2001 and 2003, the aggregate rating of all project outcomes in a country was satisfactory, while the outcome of the CAS was rated unsatisfactory (IEG 2005d).

An update of this analysis for CAEs from fiscal 2001–06 confirmed the earlier findings with a gap of 32 percent. CAEs assess the overall country outcome against the objectives stated in the CAS, including the relevance of the strategy in the country context. Projects are much narrower in scope and are evaluated as free-standing products. The discrepancy indicates that the Bank faces challenges in translating good project outcomes into successful country strategy outcomes.

At the sector level, IEG has produced similar findings:

- Its power sector evaluation found that sector outcomes have fared worse than project outcomes (IEG-World Bank, -IFC, -MIGA 2003).
- Its evaluation of natural disaster assistance found that satisfactory outcome ratings for disaster response projects concealed shortcomings in the achievement of durable outcomes and long-term vulnerability reduction at the country level, because not enough attention was paid to prevention (IEG 2006d).
- A review of the primary education portfolio found that most projects failed to look beyond increased enrollment and completion to how well students were learning, although the Bank's primary education portfolio performed above Bank average and met stated expansion targets (IEG 2006c).
- An IEG evaluation of trade found that Bank advice and lending contributed to systematic reductions in protection and in opening up economies. But the impact on employment and poverty reduction was less than expected, and the Bank did not conduct sufficient analysis to inform its policy and lending on this issue (IEG 2006a).
- Its evaluation of capacity building in Africa found that most support for capacity building is designed and managed operation by operation. Overarching capacity-building objectives were left unarticulated, and most projects embedded capacity-building activities in other program components without specifying the capacity-building objectives. The lack of integration made it difficult to capture cross-sectoral issues and opportunities and to learn lessons across operations (IEG 2005b).

A 2004 IEG review of the PRSs (IEG 2004d) found that the development of country-specific indicators and monitoring systems to track them is still at an early stage in most countries, and the available information is generally not linked to decision making. The study noted that PRSs must be more effective in enabling countries to select and monitor relevant indicators that can support domestic decision making on a sustained basis. An IEG self-evaluation of its Evaluation Capacity Development (ECD) Program noted that strong demand for M&E, across all levels of government, is critical to sustaining ECD work (IEG 2004c).

An IEG review concluded that substantive demand from the government is a prerequisite to building effective M&E systems (Mackay 2006). It also observed that there was a “chicken-and-egg” problem:

Government demand for M&E is often inhibited by lack of understanding regarding its use and importance.

A lack of government demand for monitoring and evaluation because of lack of understanding of monitoring and evaluation and what it can provide; a lack of understanding because of lack of

experience with it; and lack of experience because of weak demand.

Demand for performance information may come from government agencies involved with core governmentwide functions such as planning, budgeting, or government administration. It may also come from within government agencies interested in strengthening their performance.

The Bank has been supporting governments' efforts to strengthen data collection. The Trust Fund for Statistical Capacity Building (TFSCB) provides grant funding to build statistical development strategies and master plans that

lay out the overall goals, actions, and resources needed to strengthen the national statistical systems in the recipient countries.

These master plans provide a basis to mobilize donor support and funding through the Bank's Statistical Capacity Building Multicountry Adaptable Program Loan (STATCAP). Both TFSCB and STATCAP focus on national statistical systems. While individual task teams have been working with line agencies to develop outcome-based M&E systems, often with funding from the Institutional Development Fund, no systematic support is currently available outside the national statistical agencies.

Chapter 3: Evaluation Highlights

- IEG evaluations have strengthened their assessment of results-oriented M&E at the country and project levels.
- IEG is giving greater emphasis to applied learning and real-time use of evaluation.
- IEG has targeted ECD to focus on specific countries and has taken a lead in harmonizing evaluation practices.
- IEG's results framework should continue, and even enhance, its focus on the use of evaluation findings.
- The communications and outreach strategy is increasing the awareness of IEG evaluations, but clients want more summaries of findings and greater accessibility of reports online.
- The Board and higher-level staff report substantial use of IEG products, while TTLs make less use of them.
- More operations-oriented IEG products are needed.



Improving IEG's Effectiveness

This chapter provides an overview of IEG's role in advancing the Bank's Managing for Development Results agenda. It includes a results framework that links IEG's mandate and objectives to its operations. Within this framework, the chapter updates IEG's efforts to increase its evaluation focus on results. The chapter includes findings from the annual client survey, an update on the communications and outreach strategy, and the status of the Management Action Record (MAR).

IEG's Results Framework

IEG has three functions in the Bank's results framework. First, it provides accountability by independently reporting on the results achieved by Bank operations. Second, it distills the Bank's operational experience into knowledge of *what works* and *why* and makes that knowledge widely available to the Bank and to the global development community. Third, it supports client governments with its technical knowledge of M&E through its ECD activities.

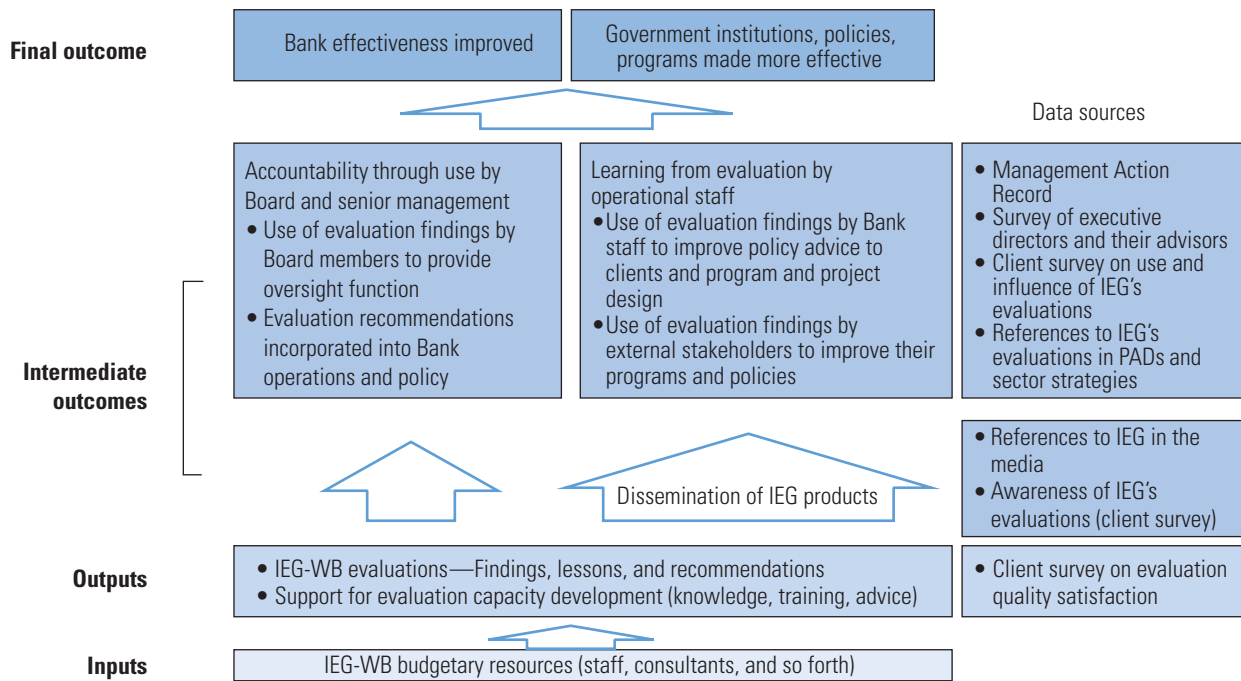
IEG has a mandate to assess "whether the World Bank Group's programs and activities are producing the expected results, including global, regional, and other programs in which the World Bank Group is a participant" (IEG 2005f). By reporting the results of its evaluations to the Board of Directors and communicating the findings and lessons from its work to Bank management, operational staff, and the development community, IEG expects to increase the Bank's

effectiveness and to influence Bank and client country decisions on policies, programs, and procedures. While fully independent,¹ IEG is placed within the Bank to maximize its operational effectiveness and to provide operational staff and strategic decision makers with knowledge that helps them work more effectively.

Figure 3.1 summarizes IEG's results chain and relates it to measures of performance that have been collected for this report. IEG's outputs are the findings, lessons, and recommendations from its evaluations and evaluation capacity support in client countries. Dissemination efforts are the intermediate step between outputs and outcomes.

For IEG's accountability function, the intermediate outcomes of these outputs are the Board's use of evaluations to fulfill its oversight function and the incorporation of IEG recommendations in Bank internal policies and procedures.

Figure 3.1: IEG Results Chain



Note: PAD = Project Appraisal Document.

For IEG’s function as a knowledge provider, the intermediate outcome of its outputs is the influence and use of these outputs to improve the Bank’s policy advice and program and project design. It also includes use of these outputs by external partners to improve their development work. Figure 3.2 illustrates the uses of evaluations by IEG’s key clients and stakeholders.

The final outcome for IEG outputs is the use of IEG knowledge about what works and why to lead to improved effectiveness of Bank operations in reducing poverty. For example, early findings of the recently completed evaluation of World Bank support to primary education (IEG 2006c) have been incorporated

The final outcome for IEG outputs is the use of knowledge it produces to improve the effectiveness of Bank operations.

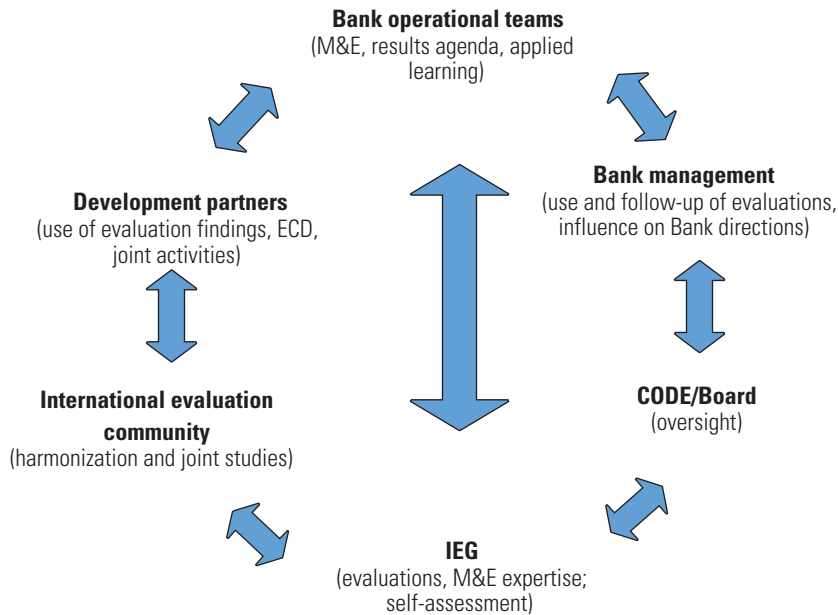
in the 2005 Education Sector Strategy Update. The trade assistance evaluation (IEG 2006a) is being used to recalibrate the Bank’s approach to trade, and pilot initia-

tives are getting under way in several Regions. The Bank’s operational staff, management, the Board, and external clients use IEG’s outputs to strengthen actions taken in client countries. Measures that attribute achievement of the final outcome to IEG, however, are difficult to construct. The remaining sections of this chapter describe each of these links in IEG’s results chain.

IEG’s Outputs: Increasing the Focus on Results and Impact in Independent Evaluation

Focusing on results has several implications for IEG’s work. IEG can contribute to the Bank’s efforts by strengthening the focus on M&E in its evaluations. It can also improve the use and influence of evaluations by producing quick-turnaround products for applied learning and decision making. Outside the Bank, IEG can affect results by continuing the development of evaluation capacity and providing leadership in initiatives to harmonize evaluation standards among MDBs and bilateral donors.

Figure 3.2: Uses of Evaluation Knowledge



Note: ECD = evaluation capacity development; M&E = monitoring and evaluation.

Strengthening the Focus on M&E in IEG Evaluations

The most significant implication of the Bank's Managing for Development Results agenda is that IEG is expected to assess the results orientation of Bank operations. In the 2004 AROE, IEG committed itself to do the following:

Strengthen attention to results-oriented M&E by requiring evaluations at the country, sector, and project levels to address systematically the three key criteria for results-oriented M&E² and increase the weight that IEG ratings give to the extent of result-oriented M&E and managing for results in Bank programs.

IEG has been **strengthening its assessment of results-oriented M&E**. At the country level, CAEs are now organized around the Bank's country objectives and sub-objectives and link Bank outputs with outcomes. In addition, IEG is interacting more with country teams during the preparation and drafting of the Country Assistance Evaluations (CAEs). IEG's emphasis on strategic partnerships is leading to further

collaboration with partners on country evaluations.

The adoption of CASCR Reviews³ represents a major shift in the Bank's evaluation system, because it ensures systematic evaluation coverage of CAS results. In CASCR Reviews, IEG looks at the relevance of the CAS to the country's development priorities, implementation of the country program, achievement of CAS objectives, and the quality of the CASCR itself.

As of May 2006, IEG had reviewed 28 CASCRs. It must be noted, however, that the current process limits the effectiveness of the CASCR Review as a learning tool. IEG receives the final CASCR for validation when the preparation of the new CAS is nearly complete. It may then be too late for the country team to take advantage of IEG's findings by incorporating much—or perhaps any—of the material directly into the CAS document. But lessons learned on the

IEG contributes to the Bank's results agenda by focusing on M&E in its evaluations.

IEG has strengthened its assessment of M&E.

implementation of the CAS should be considered by the country team as it moves to implement the new CAS.

As part of IEG's new strategic direction, the three IEG units in the World Bank Group—in the Bank, the International Finance Corporation (IFC), and

Country, sector, thematic, and project evaluations now have a more in-depth focus on M&E and results frameworks.

the Multilateral Investment Guarantee Agency (MIGA)—will work more closely than in the past on country evaluations. For example, IEG now includes evaluative content from IEG-IFC in all

IEG country evaluations, except where there is a separate Country Investment Review, and from IEG-MIGA where it has conducted evaluation work. Similarly, the ongoing evaluation of middle-income countries is being carried out jointly across IEG.

Sector and thematic evaluations now have a more in-depth focus on M&E and results frameworks. Greater emphasis

IEG should sharpen the distinction between outputs and outcomes in its project ratings.

is given to assessing the policies and programs that have led to positive outcomes and deriving lessons for their replicability, especially in meet-

ing the Millennium Development Goals (MDGs). This emphasis can be found in IEG's recent reviews of trade assistance, community-driven and community-based development, HIV/AIDS, and pension reform.

At the project level, ICR reviews have included a separate section on the design, implementation, and use of M&E since August 2005. Project Performance Assessment Reports (PPARs) are also giving greater attention to M&E issues. A desk review showed that, as of February 2006, 53 percent of fiscal 2006 PPARs reviewed the

IEG has also started evaluating global programs.

design of monitoring, which is a marked increase from 15 percent in fiscal 2005 and 6 percent in 2004. PPAR

formats vary, however, and a separate M&E section has yet to be uniformly adopted; only 20 percent of PPARs from fiscal 2006 had a separate section on M&E, compared with 15 percent in fiscal 2005 and none in 2004.

In light of the results agenda, it is noteworthy that the language in the Bank's project-rating system differs from that of the results framework. In the results framework, outcomes can be defined as a "change in the condition or behavior of a target group resulting from... outputs" (World Bank 2005c). In the Bank's project ratings, however, a broader meaning is attached to outcome. Outcome ratings in ICRs, ICR reviews, and PPARs assess overall project performance, summing up the ratings for relevance, efficacy, and efficiency; they do not distinguish between outputs and outcomes. Moreover, neither project documents nor IEG evaluations clearly differentiate between project objectives that constitute outputs and those that constitute outcomes.⁴ It would be helpful for the Bank and IEG to apply clearer language and sharpen the distinction between outputs and outcomes in project ratings.

IEG is implementing the recommendation made in its 2004 evaluation of global programs. It has completed the first three pilot Global Program Reviews (GPRs) and plans to undertake an additional four to six GPRs during each of the next two to three fiscal years. IEG will take stock of its experiences with this new product, a review tentatively scheduled for fiscal 2008, which will be similar to the recently completed retrospective analysis of CAEs. The lessons learned from the pilot GPRs have been incorporated in a revised set of guidelines for GPRs. IEG is also working with international evaluation partners to develop a common methodology for evaluating global programs.

Evaluation methodology

In addition to increasing the focus on M&E, IEG has made refinements in the methodologies for several of its evaluation products. The country evaluation methodology has been refined based on the CAE retrospective study from fiscal 2005

(IEG 2005d). IEG has worked with OPCS and QAG to harmonize the Bank's project rating systems for full consistency. Finally, IEG has developed a rating system for CASCR Reviews.

IEG Initiatives for Applied Learning and Real-Time Use of Evaluation

IEG is giving greater emphasis in its products to applied learning and real-time use of evaluation to improve Bank performance. The real-time use of IEG knowledge in the education sector in Peru is an example (box 3.1).

IEG has also introduced a new line of quick-turnaround products that address immediate needs for evaluative findings and lessons of experience, in the form of notes, presentations, or briefing papers. The quick-turnaround work with the Pakistan Country Office on lessons for dealing with natural disasters in the wake of the 2005 earthquake is an example of this increased focus on applied learning.

IEG's evaluation of the LICUS (low-income countries under stress) Initiative (IEG 2006b) assessed implementation experience rather

than outcomes, given that many of the LICUS operations were still under way. The focus was on how the Bank's approach to LICUS has been implemented, what has been learned about effectiveness, and how the Bank can do better in the future. This study represents a new approach to providing timely assessments that can help inform strategic and operational decisions.

To create incentives within the Bank for good performance in design, implementation, M&E, and development effectiveness, IEG has been giving annual Good Practice Awards to operations that exemplify strong performance in these areas.⁵ In addition to providing incentives for high performance, the awards heighten the profile of operations that offer examples of good practices.

IEG is giving more attention to applied learning and real-time use of evaluation.

Deepening Strategic Partnerships

IEG's success in achieving greater focus on results and learning will require strategic partner-

Box 3.1: Real-Time Use of IEG Knowledge—Accountability through Education Indicators in Peru

An example of the generation and dissemination of new knowledge in real time can be found in the education sector. For several years, field visits during PPAR missions to primary education projects have found high enrollment increases but very limited learning outcomes. In lower-income countries, children often read haltingly or drop out illiterate. Also, instructional time is poorly used. Classes are often cancelled for a variety of reasons.

Government and operations staff did not have a clear idea of when children ought to read fluently and tended to attribute low outcomes to poverty and malnutrition. IEG staff, with the help of Danish trust funds, surveyed the research and found studies that helped establish criteria and monitoring indicators for reading fluency. And on the basis of neurocognitive research, a rationale was established to clarify the importance of instructional

time and textbook use with respect to achieving development goals.

The results were put to a test in Peru, where nearly half the lower-income second graders were unable to read a single word. IEG worked with task managers in the Latin America and the Caribbean Region to outline learning parameters and M&E indicators that were introduced into the country dialogue and are now being adopted by the government and donors.

One important output was the production of a public service video to help Peruvian parents understand how well their children should be learning and to encourage schools to set and achieve clear learning standards, particularly for the lower grades. The video provided a tool of accountability that any parent can understand.⁹

Source: World Bank 2004a.

a. This research and its application by IEG and Bank staff has spawned a book by Helen Abadzi (2006) that integrates neurocognitive, educational, and social psychology research on issues identified by IEG PPARs in education.

ships. For example, to capitalize on the impact evaluation expertise available in other groups in the Bank, IEG has been collaborating with the Thematic Group for Poverty Analysis, Monitoring, and Impact Evaluation and has generated two joint working papers (IEG/Thematic Group for Poverty Analysis, Monitoring, and Impact Evaluation 2006a, b). In addition, IEG is currently identifying ways to collaborate with more groups, such as DEC and the Bank’s networks, on impact evaluations to ensure that evaluation findings are used.

Evaluation capacity development

IEG is strengthening its support of results-oriented M&E capacities in client countries. As a result of its 2005 self-evaluation of ECD, IEG has refocused its high-intensity support in a few targeted countries. At the same time, IEG has maintained low-intensity support to a broader range of countries. These efforts have been particularly evident in middle-income countries in Latin America, where IEG, the Latin America and Caribbean Region, and the Inter-American Development Bank are supporting a network of M&E system managers.

IEG has refocused its high-intensity support for capacity development on a few countries.

IEG will also offer the International Program in Development Evaluation Training in collaboration with Carleton University for the sixth time in mid-2006. The four-week course draws broad interest from evaluation professionals and policy makers worldwide. In addition, IEG is providing condensed versions of the training in Trinidad and Tobago, India, and Africa. The training will be provided in China in October 2006.

IEG is supporting conferences on evaluation capacity to broaden the reach of capacity building through opportunities such as the African Evaluation Association Conference and the International Development Evaluation Association Conference. IEG has produced ECD products on M&E methodology and influ-

IEG has been a leader in harmonization.

ential evaluations consistently draw strong interest among practitioners.⁶

Harmonizing development evaluation

As the largest and oldest of the evaluation units in the MDBs, IEG has taken a leadership role in harmonization efforts within the international evaluation community. IEG has actively promoted harmonization of development evaluation methods through the MDBs’ Evaluation Cooperation Group (ECG), the DAC Evaluation Network, and the UN Evaluation Group.

ECG has developed good practice standards for evaluation of both public and private sector development work. Member institutions have been benchmarked against private sector standards and now are being benchmarked against the public sector standards. Beyond that, IEG has proposed—and the ECG has agreed to—a joint paper synthesizing findings and lessons on the linkages between infrastructure and environment operations. If this pilot cooperation is successful, ECG may undertake a program of joint evaluation syntheses in the future.

As part of the DAC Evaluation Network, IEG has engaged actively in developing a system for evaluating progress on the 2004 Paris Declaration. It participates in the multidonor evaluation of general budget support and is exploring the foundation of a “network of networks” on impact evaluation.

Strengthening IEG’s Results Framework

To further strengthen IEG’s contribution to the workings of the results agenda in the World Bank Group, IEG’s own results framework, as monitored by the AROE, should continue, and even enhance, its focus on the use of evaluation findings:

- By the **Board** for oversight
- By **management** through the incorporation of recommendations into Bank policies and strategies
- By **Bank staff** through the use of evaluation findings for policy advice to country partners and in project design

- By **external partners** through the use of evaluation findings to improve programs and policies.

The results framework can be used as a management tool to identify first how IEG's activities contribute to achieving results and second where to focus efforts. For example, as IEG increases its emphasis on quick-turnaround products, it can measure changes in the use and perceived usefulness of its products by operational staff. It can also measure changes in clients' perceptions of the timeliness of IEG products. To this end, this year's client survey has further refined client feedback on influence and use of evaluations.

Quality and Relevance of IEG Outputs

IEG measures the quality and relevance of its outputs as part of its annual internal and external client surveys. This year's surveys queried target audiences for IEG products prepared during 2005, including 4 sector and thematic studies, 3 corporate reports, 7 CAEs, 37 PPARs, a synthesis report, and an impact evaluation. IEG surveyed 4,285 internal clients, consisting of Bank staff and executive directors and their advisors. The response rate was 22 percent, compared with 31 percent last year.

The survey of external clients regarding evaluations that had been disseminated externally was sent to a targeted sample of 2,759 individuals and had a 12 percent response rate. Given these response rates, it has to be noted that the survey results are indicative for respondents, but cannot be generalized to the surveyed population. Detailed survey results and graphs are presented in appendix E.

Readership and awareness

Fifty-six percent of Bank staff who responded to the annual client survey were aware of the evaluation for which they were surveyed. This is comparable to an awareness rating of 58 percent in 2005 and 39 percent in 2004. Among respondents to the external survey, 76 percent were aware of the evaluation for which they were surveyed.

Every year the survey contains questions on the quality and influence of IEG evaluations. This year, a section on the use of evaluations was added and the section on influence refined, to better understand the use of evaluations by Bank staff and to measure the achievement of IEG's second intermediate outcome.⁷

Quality of IEG evaluations

As shown in figure 3.3, Bank staff were asked to rate their satisfaction with IEG's evaluation for 10 attributes of quality on a 6-point scale. Bank staff respondents reported the highest satisfaction with the relevance of IEG's evaluations to their work, with 76 percent of respondents rating it 4 or higher, the highest rating in the past three years. Bank staff were least satisfied with the incorporation of all available information and the depth of analysis, with 58 percent and 59 percent of respondents expressing their satisfaction, respectively.

Executive directors and their advisors value the relevance of IEG evaluations to their work most, with 88 percent of respondents rating it 4 or higher. They assign the lowest rating (69 percent) to timeliness.

External respondents generally were highly satisfied with the quality of IEG evaluations.

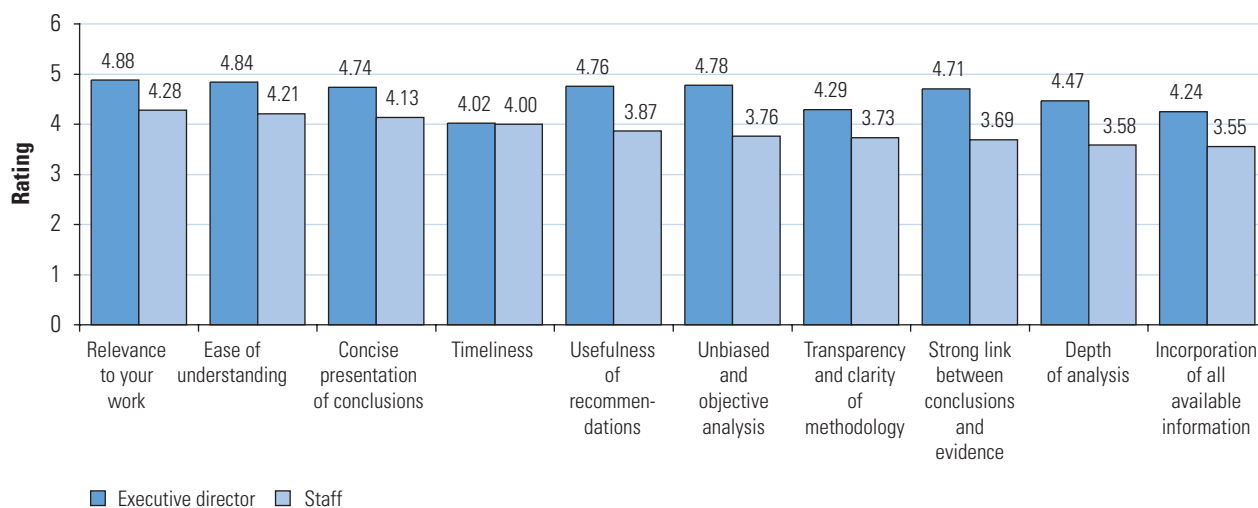
Eighty-five percent rated their satisfaction with the relevance of evaluations to their work at 4 or higher. External respondents ranked depth of analysis and incorporation of all available information last, with 80 percent rating both aspects 4 or higher.

In response to the question of whether evaluations influenced their understanding of the subject area, 79 percent of Bank staff respondents rated this 3 or higher. Ninety percent of executive directors and their advisors rated this

IEG's results framework should continue and even enhance its focus on the use of evaluation findings.

Satisfaction was highest for the relevance of IEG reports and lowest for depth of analysis and incorporation of all available information.

Figure 3.3: Bank Staff and Executive Directors and Advisors Generally Satisfied with Quality of IEG Evaluations



Source: IEG data
 Note: Mean on a six-point scale, N = 292 for staff and N = 50 for executive directors.

aspect at 4 or higher. Among external clients, 81 percent rated the influence of evaluations on their understanding of the subject area at 4 or higher. Overall, IEG does better at compiling lessons and good practices than at providing staff with specific knowledge about results chains and how outcomes are linked to outputs.

Executive directors and their advisors’ **use of evaluations** for assessing the Bank’s policies and procedures is high among those who responded: 90 percent assigned ratings of 4 or higher. Bank staff use evaluations mostly for commenting on the work of others, making a case for a particular course of action, and providing advice to clients; they use evaluations less for modifying strategies or operations or designing new projects or programs.

About one in four (27 percent) respondents acknowledged using evaluations to modify ongoing operations, assigning a rating of 4 or higher to this factor. External clients use IEG evaluations mostly for research (69 percent), and 62 percent use them to refocus ongoing strategies or programs. In open-ended questions, respon-

dents pointed to two obstacles to the usefulness of evaluations: lack of specificity of recommendations suggesting how to put them to practice, and lack of reflection on the country context within which the Bank operates, which would make findings relevant to Bank operations in the field.⁸

When asked about how the evaluation for which they had been surveyed could be improved, 57 percent of staff respondents recommended that IEG make its findings more operational. Last year, 72 percent of Bank staff respondents made this recommendation. This year, more respondents recommended that evaluation teams obtain more evidence (49 percent), broaden consultation with Bank staff (48 percent), and improve depth of analysis (45 percent), compared with last year (36 percent, 43 percent, and 33 percent, respectively). Executive directors and their advisors made similar suggestions.

Communicating Knowledge from IEG Evaluations

Effective communication of IEG’s knowledge to Bank staff, governments, other donors, and the

international community is a crucial link between evaluations and outcomes. The 2005 AROE recommended that IEG “review the results of pilot initiatives to promote more effective outreach and communication of IEG products and consider lessons learned in the full implementation of the communications strategy.”

To strengthen communication and outreach, IEG has undertaken pilot initiatives in e-mail marketing, Web promotions, media outreach, Web site enhancements, multilingual outreach, and IEG Help Desk enhancements.

IEG has begun to track the effectiveness of its communication efforts. Awareness of evaluations, as reported by Bank staff who were polled this year, has increased from 39 percent in 2004 to 56 percent in 2006. The number of follow-up inquiries to IEG's Help Desk has increased tenfold during the past year, reaching between 2,000 and 2,400 each month. Help Desk inquiries are increasingly concerned with evaluation methodology, advice on M&E systems, and IEG's product portfolio, suggesting that IEG's outreach campaigns are triggering follow-up questions among key audience segments. For a comprehensive list of improvements in IEG's outreach methods, see appendix F.

In spring 2006, IEG organized two media outreach campaigns to coincide with the release of its reports on assistance for trade reform and on natural disaster assistance. Both campaigns generated extensive media coverage.⁹ The growing number of references in the media to IEG evaluations indicates strong interest in the Bank's performance on key development initiatives.

In addition to electronic and media dissemination, IEG reaches key audiences with customized dissemination events. Twice-yearly conferences coincided with Bank–International Monetary Fund meetings and featured panel discussions on the community-based and -driven development (IEG 2005e) and HIV/AIDS

(IEG 2005c) studies in October 2005 and on the natural disaster evaluation (IEG 2006d) in April 2006. IEG also has been conducting in-country events, most recently for the trade evaluation and studies on community-based and -driven development. The latter event was a videoconference with practitioners and policy makers in six West African countries.

IEG's communication and outreach can be strengthened if IEG provides more summaries of its findings. In the client surveys, about two-thirds of all client groups made this recommendation. About half of all client groups would like to see better notification regarding events and the availability of new documents, and 52 percent of executive directors and their advisors would like IEG products to be more accessible online.

IEG Intermediate Outcomes: Evaluation Recommendations Incorporated into Bank Operations and Policy

IEG influences the Bank's Managing for Development Results agenda through recommendations to management as part of sector, thematic, and corporate evaluations, as well as CAEs. Management is accountable to the Board for follow-up. One of the intermediate outcomes for IEG is the extent to which management incorporates IEG recommendations and findings in policy advice and project and program design.

IEG's 2004 evaluation of lines of credit (LOCs) offers a good example of corrective action by management following IEG recommendations.

IEG found poor results of Bank-funded LOCs and widespread lack of compliance with operational policies governing LOCs. It recommended

More than half of staff respondents suggested that the findings of IEG evaluations needed to be more operational.

Further strengthening can be achieved by providing more summaries of IEG findings.

Incorporation of evaluation recommendations in operations and policy is a major intermediate outcome for IEG.

that Bank management update Operational Policy (OP) 8.30 to cover all LOCs across all sectors and Regions. The Consultative Group to Assist the Poorest’s (CGAP’s) self-evaluation of LOCs for compliance with OP 8.30 confirmed IEG’s findings. To address IEG’s recommendations, management subjected all identified LOCs to a quality assurance system to foster compliance with OP 8.30. Comments from a staff survey indicate that the evaluation has affected projects; for example, one comment was: “*evaluation of Bank LOCs has provided guidance on our recent design of an LOC in the area of micro and small lending. Lessons learned in that report have proved very enlightening.*”

The MAR allows IEG to track its recommendations from sector, thematic, and corporate evaluations and to monitor the degree of their

Two-thirds of IEG recommendations were adopted at a high or substantial level in 2005.

adoption by management. The MAR tracks two indicators: the level of adoption¹⁰ and the status of individual recommendations.¹¹ It

presents management’s ratings for these two indicators and IEG’s validation of the same.

The MAR for 2006 tracks management’s progress on 79 recommendations. These include 21 new recommendations from the 6

IEG studies (excluding CAEs) presented to the Board in calendar year 2005 and 58 recommendations carried forward from 2002 to 2004.

For the 2006 MAR, IEG rated 66 percent of its recommendations as having been adopted by management at high or substantial levels (figure 3.4), compared with 75 percent in 2005 and 62 percent in 2004.¹² The trend of high or substantial adoption of about two-thirds of IEG’s recommendations by management has continued in 2006.

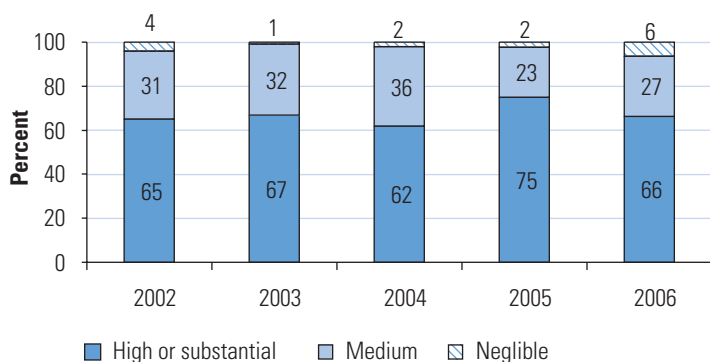
IEG and management agreed on the rating for level of adoption for 65 percent of recommendations (50 out of 77).¹³ This figure represents a decline from 70 percent in 2005 and 78 percent in 2004. Where there was disagreement in level of adoption, IEG’s ratings were lower than those of management for all except one of the recommendations.

The 2005 AROE included recommendations to improve the MAR. A pilot three-year expiration criterion has been introduced for MAR recommendations from fiscal 2006 onward. Recommendations issued in the previous three calendar years (2002–04) are being retained in the MAR. Recommendations from earlier years (calendar years 1998–2001) are being archived, subject to review by the Committee on Development Effectiveness (CODE) and the Board. It should be noted that the three-year retirement rule reduces the comparability of the 2006 MAR to those of prior years because the 2006 MAR has proportionally more recent evaluations than earlier MARs. A summary of the 2006 MAR is included in appendix H and posted on the Bank’s external Web site.

IEG Intermediate Outcomes: Use of Evaluation Findings by Bank Staff at the Operational Level

IEG’s performance on this intermediate outcome can be measured by whether Bank staff use IEG findings to improve Bank policy advice and program and project designs. For the 2006 AROE, IEG sought to understand better how Bank staff use IEG findings. The internal client

Figure 3.4: Level of Adoption of IEG Recommendations Similar to That in Previous Years



Source: IEG data.

survey on the use and influence of IEG evaluations indicates that IEG provides good services to its principal client, the Board. More than 80 percent of this respondent category rated use of evaluations to assess sector and country strategies and Bank policies and procedures at 4 or higher.

Bank staff make less use of IEG products in their day-to-day work, which consists of designing and implementing projects and programs. Although IEG is seen as enhancing the understanding of subject areas and lessons learned, it is rated lower at helping to identify results chains and providing knowledge that can be readily applied in ongoing operations. Thus, IEG provides high-level knowledge that is useful for assessing programs, giving advice to clients, and making comments, but it needs to focus on influencing ongoing and future operations.

Improving the timeliness of its products and giving more attention to presenting the context of evaluated operations is likely to improve IEG's impact on Bank operations. Moreover, IEG could do more in disseminating PPARs. Though they contain valuable project lessons, PPARs are fully disclosed on IEG's Web site but are not being disseminated to a wider audience. Consistently high ratings for PPARs for influence and use suggest that they represent an untapped opportunity for applied learning at the operational level.

In addition to the client survey, findings from a desk review of selected operational documents for explicit references to IEG reports suggest that IEG evaluations are used as a reference source more frequently at the sector-strategy level than at the project level. A review of 234 Project Appraisal Documents (PADs) and 60 ESW reports from calendar year 2005 and 15 sector strategies from 2000 to 2005 found few explicit references to IEG evaluations in project design documents and ESW reports. Only about 10 percent of these documents refer to the content of IEG reports. In contrast, sector strategy papers cite IEG documents much more frequently. Thirteen of 15 sector strategy

documents contain multiple references to IEG. It should be noted, however, that this cursory analysis does not capture the full extent of IEG's influence.

The Management Action Record is being simplified and improved.

Focus groups conducted for the 2006 AROE generated information on how Bank staff and management view the use and usefulness of IEG's evaluation results, products, and information. Appendix D presents a summary of the focus group findings.

TTLs are least likely to report that they incorporate evaluation findings into planning and design, a finding consistent with the results of the desk review of PADs. Their operational context requires information on how to monitor, establish indicators, and prepare projects for evaluation. Respondents in management positions reported greater use and usefulness of IEG products. CODE members rely heavily on IEG's reports, advice, and recommendations. Nearly all levels of Bank staff respondents acknowledged the challenge of bridging the gap between knowledge and application. In addition to IEG conducting retrospective evaluations of previous projects and programs, these findings suggest the value of IEG pilot testing short summaries of findings that are directly targeted at operational staff and monitoring their use and perceived usefulness by TTLs.

IEG's role in a results-focused context has been evolving to strengthen the focus on M&E in IEG evaluations and to enhance evaluation capacity among its internal and external clients. To improve Bank performance, IEG will need to continue giving greater emphasis in its products to applied learning and real-time use of evaluation findings. IEG's success in increasing the

Bank staff use evaluation findings less than executive directors and their advisors.

IEG findings are more often used at the sector-strategy level than at the project level.

focus on results will require strategic partnerships, both internally and externally. Continued emphasis is needed on increasing the awareness of IEG products and on developing quick-turnaround products to provide just-in-time knowledge to decision makers at the strategic and operational level.

The use of IEG evaluation findings is evident with strategic decision makers. The MAR is a useful tool to assess the implementation of IEG recommendations. Increased attention is needed to make IEG's findings more relevant to ongoing operations. IEG's work program for fiscal 2007 indicates that it is working to meet these emerging needs.



Conclusions and Recommendations

The World Bank has been strengthening the results focus of its operations, starting with the emphasis on improving project development outcomes in the early 1990s, the move to a stronger country focus in the late 1990s, and the heightened attention given to managing for results, in concert with other development partners, in the early 2000s.

In 2003, the Bank put in place a Managing for Development Results action plan, which included actions to do the following:

- Strengthen the countries' capacity and demand to manage for results (Country Pillar).
- Enhance the Bank's relevance and effectiveness in contributing to results (Bank Pillar).
- Harmonize results-based approaches across development agencies (Global Pillar).

The Bank is making progress in carrying out the plan, particularly at the project and country levels. OPCS reported (World Bank 2006b) that PRSs are focusing government, donor, and Bank programs around desired results in low-income countries. Within the Bank, the RBCAS, CAS Progress Report, CASCR, and IEG validation provide the basis for managing for results at the country level. In addition, results frameworks have been introduced for Investment Loans and Development Policy Loans. The newly adopted ISR gives more prominence to monitoring. The Bank also continues to work with other MDBs and other donors to harmonize approaches,

policies, and procedures to reduce the costs to countries dealing with multiple partners in managing for results.

On the sector, thematic, and global levels, efforts to focus on results are at an early stage. Sector results frameworks are being developed and the Bank is reporting on progress in selected sectors through Sector Strategy Implementation Update reports. Following up on an IEG review that identified the need to strengthen results frameworks for better monitoring and evaluation in global programs, some changes have been made in the management of global programs (appendix A). Further work is needed to report achievements at the sector and global program levels and to assess more systematically and strategically what has worked well and why in different sectors and themes across countries.

In its 2006 progress report, OPCS identified a list of possible actions to accelerate the implementation of the results agenda over the coming years and laid out next steps in its results framework (annex B of the report). The list of proposed

action items is summarized in appendix G. At the country level, OPCS proposes enhancing governments' capacity to use performance information for policy implementation at the sector and cross-sector levels.

Internally, the updated action plan calls for Regions and networks to provide leadership for advancing the results agenda in collaboration with OPCS under the umbrella of a newly constituted Results Steering Group. New training initiatives and stepped-up support for Bank teams are proposed to encourage increased use of M&E information at the operational level. While the report provides a credible list of action items, OPCS does not identify priority action items or funding sources of this next phase.

The emphasis on results is hardly new. Monitoring and evaluation have long been considered to need improvement.¹ The most recent Monitoring and Evaluation Action Plan (fiscal 2001–03) introduced results-based M&E in pilot countries and provided training and technical support to increase the Bank's own capacity to manage for results (World Bank 2000). Although the country pilots encouraged the adoption of results-based country portfolio reviews and influenced the development of RBCASs, the efforts to strengthen the Bank's internal M&E capacity were below expectations because of competing priorities, fragmented efforts, and lack of adequate funding.

The 2004 AROE identified culture, incentives, and lack of a funded action plan as constraints to improving the results orientation of the Bank's M&E system. It recommended that Bank management (a) develop, maintain, and communicate a sustained vision of a results-oriented Bank and define how organizational culture, incentives, and structure need to change; (b) develop a phased and costed plan to implement results management; and (c) review and strengthen incentives to promote results-oriented M&E.

This year's AROE finds that the Bank has made significant progress in putting in place revised policies and procedures to better manage for

results. However, it notes that these have not yet translated into improved incentives and practices at the operational level. Poorly developed results chains and indicators without adequately defined baselines and/or targets call into question the likely effectiveness of these results frameworks for M&E at the project and country levels.

The analysis identified the following constraints to outcome-focused M&E:

- Conceptual difficulties in establishing meaningful results chains
- Possible crowding out by other priorities (safeguards, fiduciary, and the like)
- The continuing perception that "delivery of new projects" is what matters most
- A lack of appreciation of the possible uses of performance information and evaluations for management decision making.

The analysis reconfirmed the importance of focusing on country outcomes and aligning Bank operations around them, as well as the need to promote more cross-sectoral collaboration and multisector teams. The Bank needs to continue to strengthen results frameworks, at both the country and project levels. While it is critical to come up with a manageable number of outcomes and performance indicators, the focus should not be exclusively on what is measurable and achievable, and staff should be encouraged to take on difficult but important challenges.

With retrofitting of internal policies and procedures largely completed at the country and project levels, and under way at the sector, thematic, and global levels, the Bank now faces a key challenge: to change the culture and perceptions within the organization that act as informal incentives and disincentives. These are difficult challenges that will take time to resolve. The basic principles of the recommendations from the 2004 AROE remain valid: the Bank needs to maintain and communicate its vision for a results-oriented Bank; address incentives; and develop a cohesive, funded action plan. Management could send a strong signal to staff that it remains committed to the results agenda by

proposing, communicating, and implementing such a funded action plan.

IEG's role in advancing the Bank's results agenda is evolving to meet the need for timely and operationally useful information. To increase IEG's effectiveness, greater use and influence of IEG evaluations can be achieved by emphasizing learning aspects, embedding evaluation functions earlier and more systematically into operations, developing quick-turnaround products to support decision making based on evaluative findings, and continued deepening of communications to raise awareness of evaluation findings and lessons.

Recommendations for Management

- Building on progress achieved so far in advancing the results agenda, agree on a three-year action program supported by appropriate funding for the implementation of the next stage, differentiating new allocations and re-deployed resources. Critical action items would be the following:
 - Support country directors and country teams in their efforts to refine and use results frameworks at country and sectoral levels to manage country programs. Similar efforts are needed to strengthen the results frameworks of thematic and global programs.
 - Assess the effectiveness of (self-) evaluation approaches during project and country program implementation and provide guidelines to staff on their use.
 - Strengthen incentives and accelerate a results-oriented training and communication program for management and staff to encourage use of M&E information.
- Identify and support in-depth learning opportunities to develop and use results-based approaches with task teams, particularly in challenging country cases and in complex multisectoral settings.
- Provide support to task teams and technical advice to countries that intend to institutionalize M&E systems to strengthen the collection and use of performance information in order to enhance their capacity and demand to manage for results.

Recommendations for IEG-WB

- To further strengthen IEG's contribution to the workings of the results agenda in the World Bank Group, IEG should continue to follow its own results framework and monitor it through the AROE. Its focus on the usefulness of evaluation findings for its core audiences should be enhanced: for the **Board** for oversight, for **management** through the incorporation of recommendations into Bank policies and strategies, for **Bank staff** through the use of evaluation findings for policy advice to country partners and in project design, for **external partners** through the use of evaluation findings to improve their programs and policies, and for the **countries** more broadly. In playing this role, IEG should:
 - Improve the timeliness of its evaluations.
 - Strengthen the operational relevance of the findings.
 - Increase access to and exchange of the lessons.

APPENDIXES

APPENDIX A: OVERVIEW OF MONITORING AND EVALUATION IN THE WORLD BANK

Monitoring

Monitoring is a joint responsibility of the country and the Bank at the project and country levels. It requires: (a) defining the expected outcomes; (b) identifying monitoring indicators for final and intermediate outcomes, as well as outputs that contribute to achieving the outcomes; (c) ensuring that baseline data are available and targets are set to assess progress; (d) making sure a system is in place to collect, analyze, and report the data; and (e) monitoring progress. Monitoring information should be used to assess progress toward achieving objectives and to inform adjustments or other actions needed to ensure that the objectives are met.

At the activity level

Bank task teams work with the client to set up the monitoring and evaluation (M&E) arrangements for Bank activities before implementation begins.¹ The Bank's current requirements for different kinds of activities are as follows:

- **Investment Loans and Credits:** All projects are required to have a results framework that includes the Project Development Outcomes (PDOs) for which the project can be held reasonably accountable at completion, and intermediate outcomes that can be used to assess progress toward the PDOs during implementation. Each outcome should include indicators, along with baselines and targets to measure the degree of success in achieving the PDOs and intermediate outcomes. The results framework is included as an annex to the Project Appraisal Document (PAD). It is meant to specify how the indicators will be used to assess performance and how the data will be collected.
- **Development Policy Loans (DPLs):** A focus

on results is one of the key features of the new policy on DPLs.² The program document should specify the expected outcomes of the operation and include measurable indicators for monitoring progress during implementation and evaluating outcomes on completion. As with investments, the program document has a results framework that includes a definition of program objectives and outcomes that are directly influenced by the operation, actions that are expected to cause the planned outcomes, and outcome indicators with baselines and targets to measure progress and assess the program at completion. The Bank is taking a more programmatic approach to DPLs, requiring clear definition of expected medium-term outcomes and the initiation of monitoring systems. The program document for the first operation incorporates measurable indicators for monitoring progress and evaluating results upon completion of the program. The program documents of each subsequent operation would need to report progress on these indicators under the previous operation(s). Even single-tranche operations where all conditions are fully met prior to presentation to the Board are expected to include end-of-program outcomes and a medium-term program with (intermediate) outcomes (even if these were to occur outside the time frame of the single-tranche operation) that could be monitored separately.

- **Analytical and Advisory Activities (AAA):** Task teams are required to outline the development objectives and establish results indicators for each objective when starting economic and sector work (ESW), nonlending technical assistance, or other AAA. The objectives and indicators should be outlined in the

concept paper and recorded in the Activity Initiation Summary (AIS). The task teams should select their objectives and indicators from a list available in Systems Applications and Products (SAP) or supply their own. These indicators should be used to evaluate success at the completion of the activity.

Task teams are responsible for periodically recording progress using a standard reporting format in SAP. For loans and credits, the Implementation Status and Results (ISR) Report was introduced in 2005 to replace the Project Supervision Report as part of the Bank's business simplification process. One of the objectives in introducing the ISR was to strengthen the focus on likely outcomes. Monitoring is given more prominence by making the project performance indicators part of the main report.³ For AAA, process milestones are recorded in SAP in the Activity Update Summary as they are achieved. The system, however, is seen by some as "user unfriendly" and may not be uniformly used.⁴

At the country level

The Country Assistance Strategy (CAS) provides the basis for monitoring implementation of operations at the country level. The Results-Based CAS (RBCAS) approach, which was started on a pilot basis in 2003, has been adopted as Bank policy and is incorporated in the revised Bank Procedure (BP) 2.11, Country Assistance Strategies, issued in June 2005. The revised BP now includes a separate section on CAS implementation that points out the importance of the results framework, the use of national systems, and links with project-level systems for M&E.

A key element of the results-based approach is the CAS results framework, which lays out the logical relationships between Bank operations and outcomes and includes measurable performance indicators to track progress. It links country development goals, CAS outcomes, milestones, and Bank operations, and it clarifies the Bank's contribution to achieving country development goals. A CAS Progress

Report is prepared mid-way through CAS implementation to present key developments and introduce corrections as needed.

The Bank is attempting to assess the impact of its operations at the country level when conducting Country Portfolio Performance Reviews (CPPRs). The 2005 *Annual Report on Operations Evaluation* (AROE) identified countries where the Bank was using the CPPR to raise awareness of and carry out a dialogue with the government on the results of Bank operations. The Operations Policy and Country Services (OPCS) review of CPPRs (World Bank 2006c) found that country teams have been proactively incorporating a results approach into the CPPR process.⁵ The review concluded that an RBCAS and a government strategy with specific country objectives and (ideally) targets were needed to properly link Bank operations with the overall country program.

At the sector level

At the sector level, the annual Sector Strategy Implementation Update (SSIU) presents a consolidated view of progress in implementing the sector and thematic strategies adapted by the Bank. The report is in two parts: the first provides a Bank-wide overview of strategy implementation progress and a reflection of the strategies in CASs, as well as the Bank's sector and thematic inputs (resources), outputs, and development outcomes. It also discusses progress on managing for results at the sector level.

Part two of the SSIU provides a more in-depth analysis of four strategies that are selected in rotation. Each sector presents its Bank-wide results framework, which establishes final and intermediate country outcomes for the sector; identifies performance indicators; outlines how the Bank could contribute to these results; and presents process indicators to determine whether implementation is on track. One of the challenges in tracking implementation of the sector strategies is the quality and availability of data and the variability across countries. For example, the SSIU reports that two of the three

country outcome indicators on labor market relevance in the education sector results framework are not yet available. The Bank will work with the International Labor Organization and the United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics to develop these data.

At the corporate (Bank-wide) level

Monitoring at the corporate level is carried out through different instruments, each with a specific purpose, although there are overlaps. Corporate-level performance monitoring is summarized in an attachment to this appendix. Performance reporting and performance indicators are divided into two tiers: country-level outcomes and Bank performance outcomes.

- **Country-level outcomes:** A higher tier of outcomes to which the Bank contributes, but for which attribution is difficult to assess—for example, the Millennium Development Goals (MDGs) in the Global Monitoring Report (GMR) and the 14 country outcome indicators in the IDA 14 Results Measurement System (RMS).⁶ Data on these country-level outcomes are compiled by Development Economics (DEC).
- **Bank performance outcomes:** A second tier of performance outcomes and indicators where the Bank's responsibility and accountability can be more clearly defined. For example, the Quality Assurance Group (QAG) reports on the effectiveness of the Bank's current lending portfolio⁷ and AAA programs through its Annual Report on Portfolio Performance (ARPP). The IDA 14 RMS contains a second tier of nine indicators to monitor IDA's contribution to country outcomes in three categories: (a) at the country level, where the cumulative introduction of RBCASs in International Development Association (IDA) countries will be monitored; (b) at the project level, with four indicators relevant to project quality⁸ that will be monitored; and (c) outputs for completed projects in four sectors (health, education, water supply, and rural transport).

As part of its efforts to advance the results agenda at the operational level, the Bank is

proposing a two-tiered Results Monitoring and Learning System (RMLS) with a first tier of country indicators, which is expected to include the IDA 14 RMS plus indicators for growth, macroeconomic management, social protection, and trade to accommodate IBRD countries; and a second tier of Bank indicators consisting of QAG and Independent Evaluation Group (IEG) scores on project quality and outcomes, plus country performance indicators based on Country Assistance Strategy Completion Report (CASCR) ratings. An Operational Performance and Results Report would build on the existing work carried out in the ARPP, SSIU, and CAS retrospectives and include an analysis of the RMLS information. A key objective of the report will be to extract lessons learned that can be used in operations.

To allow benchmarking of the multilateral development banks (MDBs), the Common Performance Assessment System (COMPAS) is being developed to provide a common source of information on the results orientation of the MDBs under the GMR. COMPAS focuses on processes and specific actions to enhance a focus on results that are within the control of each MDB; it includes a set of process and results indicators to compare progress.⁹ The results of the first COMPAS exercise will be available late in 2006.

As part of its efforts to better align resources with strategic directions, the Bank introduced Strategic Performance Contracts (SPCs) with the Corporate Budget Reform of 2005. The SPC has four parts: (a) Strategic Direction (Are we focusing on the right things?); (b) Strategic Choices (Are we deploying resources effectively to meet our objectives?); (c) Risk Management (How does risk impact our work programs?); and (d) a Performance Contract (How well are we performing?). The vice presidential units (VPUs) outline their strategic objectives and present outcome indicators under Strategic Choices and share a common set of key performance indicators for strategic alignment, unit outputs, resources, and structure. Most of these indicators are taken

from Bank systems or are included in the two-tier corporate monitoring system described above. The SPCs allow the VPUs to lay out and assess their strategy, link it with their overall resource allocation, identify risk, and provide key performance indicators to monitor performance.

Monitoring of global programs and partnerships

In addition to International Bank for Reconstruction and Development (IBRD) loans, IDA credits, and AAA, the Bank engages in Global Programs and Partnerships (GPPs), which have grown from about 20 programs in 1990 to about 110 global and 40 Regional programs that disbursed around \$3 billion in fiscal 2005. GPPs are partnerships, typically with a governance structure and secretariat located in one of the partner organizations or as an independent legal entity in its own right. Partners contribute resources, which are pooled in support of activities that are global, Regional, or multicountry in scope.

Within the Bank, the Global Programs and Partnership Council oversees the strategic framework and policies for GPPs. Bank task managers are assigned to implement individual GPPs within the Bank. GPP activities in the Bank are funded through the Bank's budget (for staff costs), Development Grant Facility (DGF) grants, and trust funds. Three large programs—the Global Fund to Fight AIDS, Tuberculosis, and Malaria; the Global Environment Facility; and the Consultative Group for International Agriculture Research—accounted for about three-quarters of global program expenditures in fiscal 2005.

IEG evaluated the Bank's involvement in GPPs in two phases: Phase 1 (IEG 2002) carried out a portfolio review of the 70 programs that existed at the time, and Phase 2 (IEG 2004a) presented an in-depth analysis of 26 programs. Both reports identified the need for improvement in the management of GPPs within the Bank. The evaluations also found that many programs had weak M&E systems, that is, lack of clarity in

program objectives; weak results frameworks, which lacked a structured set of indicators; and inadequate arrangements for data collection and analysis.

The Bank has taken steps to strengthen its oversight of GPPs. The Global Programs and Partnership Council was established in response to a recommendation made in the Phase 1 evaluation. The Bank issued a strategic framework paper in 2005, which took steps to strengthen the accountability of the VPUs for overseeing GPPs, including more effectively linking GPPs to the Bank's country and Regional strategies and revising the choice of issues for engagement.

To improve quality at entry, QAG is reviewing a sample of DGF-supported GPPs, typically within six months of Board approval of the DGF allocation. Monitoring and reporting is expected to be strengthened through the Grant Reporting and Monitoring (GRM) system for trust funds, which was introduced in 2006 to replace the previous Trust Fund Status Report. The GRM is more focused on outcomes and aligned with the ISR. As with the ISR, however, the GRM needs to be underpinned by programs with robust results frameworks to function effectively as a monitoring instrument.

Evaluation

The World Bank evaluates its operations on their outcome (including relevance, efficacy, and efficiency), sustainability (resilience to risks), and the performance of the Bank and the borrower. Evaluations are done both by Bank operational staff (self-evaluations) and by IEG, which reviews and validates the Bank's self-evaluations and conducts its own independent evaluations.

At the activity level

Investment and development policy loans/credits.

The evaluation framework at the project level is well established. The task team prepares an Implementation Completion Report (ICR) for each project (both investment and DPLs), which

is supposed to be submitted to the Board within six months of project closing. The ICR assesses the degree to which the project achieved its development objectives and comments on Bank and borrower performance. It discusses lessons learned from the success or failure of the project. The borrower, in turn, provides comments on the ICR. IEG then conducts desk reviews of all ICRs to independently validate their ratings.

During the past 18 months, Bank management and IEG have introduced several adjustments to ICRs to improve their effectiveness:

- **Rating outcomes of projects with formally revised objectives:** In the past, operational staff and IEG often disagreed on how to evaluate projects where the Bank had made mid-course adjustments to accommodate changing country circumstances. Guidelines were issued in January 2005 that provided for projects with formally revised objectives¹⁰ to be assessed against both the original and revised project objectives, using disbursements to determine the weight of each.
- **Harmonizing evaluation criteria:** In October 2005, OPCS and IEG jointly announced a revised and harmonized set of ratings for ICRs and IEG's ICR reviews. With the new rating scheme: (a) a common six-point rating scale was introduced for outcome, Bank performance, and borrower performance to allow for greater nuance in judgment; (b) the assessment of the project's institutional development impact was made part of the rating for outcome, precluding the need for a separate rating (and reducing the total number of ratings from five to four); and (c) the sustainability rating was replaced by a new rating of risk to development outcome (on a four-point scale), which specifically assesses the uncertainties faced by the project's development outcome at the time of evaluation. Additional clarification was issued on outcome, Bank performance, and borrower performance ratings. The system will be implemented in fiscal 2007.

- **ICRs for programmatic DPLs:** OPCS and IEG jointly announced a simplified procedure for programmatic DPLs in March 2006. Under this new procedure, a single ICR is prepared after the closing of the last operation in a pre-specified series of programmatic DPLs. The ICR would include and justify individual ratings for each operation in the series; simplified ICRs for each operation are no longer required. The new procedures took effect in April 2006.

In addition to the ICR reviews, which systematically review all lending operations, IEG conducts ex post evaluations of one-quarter of completed projects through its Project Performance Assessment Reviews (PPARs). Projects subject to PPARs are selected using a number of criteria, including the potential to learn important lessons (especially in new areas of Bank work); lack of adequate data in project documentation to ensure reliable validation of ratings; and support for planned IEG country, sector, or corporate evaluations. A PPAR usually involves a field visit and is generally carried out several years after project completion to get more information on actual outcomes than is available shortly after project closing.

AAA. An Activity Completion Summary (ACS) is prepared for all AAA activities, irrespective of cost.¹¹ The ACS consists of two parts: a rating of each results indicator selected in the AIS and a summary of major conclusions and recommendations resulting from the task. The ACS is supposed to be filed in SAP within six months of the delivery to client date.

IEG does not carry out regular evaluations of AAA. However, AAA is reviewed as part of sector and thematic studies, Country Assistance Evaluations (CAEs), and other IEG studies. IEG has recognized the lack of a systematic review of AAA as an evaluative knowledge gap and is currently conducting an evaluation of the impact of ESW and nonlending technical assistance on client policies, programs, and institutions. The evaluation is scheduled for delivery in fiscal 2007.

At the country level

All CASs are now required to include a self-evaluation of the Bank's previous country program, the CASCR. The country team reviews the effectiveness of the previous program in achieving stated objectives and outcomes based on the results framework (where available) set out in the previous CAS; derives lessons learned, which are incorporated in the new CAS; and summarizes the findings in the CASCR. A draft CASCR is ready for review at the beginning of CAS preparation (upstream review), and the final version is made available when the draft strategy is near completion (downstream review).

IEG reviews the CASCR to independently validate the country team self-evaluation. IEG's assessment of the final CASCR is made available to the Board before presentation of the CAS. IEG reviews the relevance of the CAS to the country's development priorities, implementation of the country program, achievement of CAS objectives, and the CASCR.

In addition, IEG carries out an independent assessment of the Bank's country program over a longer period of time—typically 10 years, which covers multiple CAS periods—through its CAEs.

At the sector level

The Bank provides a report on the status of some strategies, which are selected on a rotating basis, in the SSIU. For each strategic or thematic area, the SSIU provides information on the key achievements in the fiscal year, progress made in intermediate and final outcomes, improvements made in output quality, and progress toward improving M&E within the sector or thematic area. No self-evaluations are being carried out at

the sector/thematic strategy level. IEG conducts sector and thematic reviews that examine Bank performance and experience over 5–10 years and evaluate the relevance and effectiveness of Bank policies and operations.

Evaluation of GPPs

Self-evaluations of GPPs are being carried out as part of the Bank's evaluation of its DGF grants and trust-funded activities. For activities funded through the DGF, Grant Completion Reports are to be prepared within six months of the closing date of the grant. For trust-funded activities, an ICR is to be prepared for all trust funds exceeding \$1 million within six months of activity completion. The current ICR format is being revised to provide a more complete assessment of the achievement of trust fund objectives, and a more systematic review process of the content of the ICRs is being put in place.

Following the Phase 2 Global Program Evaluation (IEG 2004a), IEG plans to carry out Global Program Reviews (GPRs) for four to six programs over each two- to three-year period. The GPRs are similar to PPARs for projects and CAEs for CASs. Three pilot GPRs were completed in fiscal 2006.¹² Unlike investment projects and Bank country programs, GPPs are partnerships with their own governance structures and the Bank is only one partner in these programs. In response, IEG is working with the Development Assistance Committee of the Organisation for Economic Co-ordination and Development (OECD/DAC) Network on Development Evaluation, the United Nations Evaluation Group, and the Evaluation Cooperation Group of the MDBs to develop a common evaluation framework for GPRs.

Attachment: Corporate-Level Performance Monitoring Summary

System/performance indicators	Reporting frequency	Bank's intended user/decision maker	Data collection	Responsibility for data collection
Global Monitoring Report				
MDGs	Annual—joint with partners	Report to development committee (little direct Bank attribution)	From national statistical systems/WDI	DEC
COMPAS (process indicators and results indicators on MDBs' results orientation)	Annual—joint with partners	Report to development committee on the performance of IFIs (direct Bank attribution for its own performance)	Compiled from submissions from MDBs Some results indicators available, other need to be developed Initially focus on process indicators	MDB chairing the MDB Working Group on Managing for Results; currently ADB, next World Bank Within World Bank—OPCRX handling
IDA 14 RMS				
Tier 1—Country outcome indicators (income/poverty, health, gender, education, infrastructure, private sector development, governance, growth)	Complied annually Reported every 3 years in line with IDA replenishments	Reporting out to IDA donors (collective action, little direct attribution to Bank)	From national statistical systems/WDI	DEC + OPCRX
Tier 2—IDA's contribution (RBCASs, project quality/ outcomes, specific outputs)	Complied annually	Reported every 3 years in line with IDA replenishments.	Reporting out to IDA donors, external stakeholders (direct attribution to Bank)	Some available from SAP/Bank database, some indicators to be developed, systems to be put in place to capture outputs OPCRX
RMLS (proposed)				
Tier 1—Country indicators (IDA 14 plus growth, macro-economic management, social protection, and trade)	Annual—(proposed) Operational Performance and Results Report	Reporting progress to Board, senior management. Board, management, and staff expected to draw lessons learned and implications for operational work (collective action, little direct attribution to Bank)	From national statistical systems/WDI	DEC + OPCRX
Tier 2—Bank indicators (CASCR assessment by sector/theme + IDA Tier 2 but Bank-wide)	Annual—(proposed) Operational Performance and Results Report	Reporting progress to Board, senior management, staff, and drawing lessons learned and implications for operational work (Direct attribution to the Bank)	Some available from SAP/Bank database, some indicators to be developed, systems to be put in place to capture outputs	OPCRX

(Continues on the following page.)

Attachment: Corporate-Level Performance Monitoring Summary (continued)

System/performance indicators	Reporting frequency	Bank's intended user/decision maker	Data collection	Responsibility for data collection
Sector Strategy Implementation Update				
Sector performance indicators (outputs, quality, inputs, costs, client responsiveness, results management, and selected country outcomes)	SSIU produced annually, but specific sector would most likely only be reported on every 4–5 years	Reporting progress to Board, senior management (Direct attribution to the Bank)	Bank database, human resources data, staff estimates	Networks
Annual Report on Portfolio Performance				
Overall lending (commitments) Portfolio performance (percent DO satisfactory, projects/commitments at risk, etc.) AAA (deliverables) Implementing the Results Agenda (direct attribution to the Bank)	Annual	Reporting to the Board and senior management on the effectiveness of the Bank's lending portfolio and AAA	SAP/Bank database, QAG assessments	QAG
Strategic Performance Contracts (Regions)				
Part 1—Strategic directions (outcome indicators—varies by Region)	Based on frequency of strategy review	Used as basis of annual dialogue between VPs and senior management during budget preparation	Not clear—various sources depending on indicator	Regions
Part 4—Performance contract (key performance indicators: [1] results on the ground, [2] strategic alignment, [3] unit outputs, and [4] resources and structure)	Fiscal 2006 QBR2 (most recent available) discusses traditional deliverables (number of projects approved, commitment amounts, number of ESW completed, number of CASs completed, expenditure utilization, and so forth) or items that are measurable and worth measuring on quarterly basis	Used as basis of dialogue between vice presidents and Bank senior management as part of QBR	Depends on indicator Most available from SAP/Bank database, but some are new and indicated as “to be developed”	Regions

Note: AAA = analytical and advisory activities; ADB = Asian Development Bank; CAS = Country Assistance Strategy; CASCR = CAS Completion Report; DEC = Development Economics Vice Presidency; ESW = economic and sector work; IDA = International Development Association; IFIs = international financial institutions; MDGs = Millennium Development Goals; OPCRX = World Bank Results Secretariat; QAG = Quality Assurance Group; QBR = Quarterly Business Review; RBCAS = Results-Based Country Assistance Strategy; RMLS = Results Monitoring and Learning System; RMS = Results Measurement System; SAP = Systems Applications and Products; WDI = World Development Indicators.

APPENDIX B: IMPLICATIONS OF THE WORLD BANK'S RESULTS AGENDA FOR MONITORING AND EVALUATION

Monitoring and evaluation traditionally have been considered separate activities. This view is evident in World Bank Operational Directive (OD) 10.70, Project Monitoring and Evaluation (1989), which defines M&E at the project level as follows:

- **Monitoring** is the continuous assessment of project implementation in relation to agreed schedules, and of the use of inputs, infrastructure, and services by project beneficiaries. Its main objectives are to provide continuous feedback on implementation and to identify actual or potential success and problems as early as possible to facilitate timely adjustments to project operation.
- **Evaluation** is the periodic assessment of the relevance, performance, efficiency, and impact (both expected and unexpected) of the project in relation to stated objectives.

Monitoring and evaluation were also considered the responsibility of the borrower's project management team. The Bank team was expected to assist the borrower in setting up and using the M&E system. The OD acknowledges that an interim evaluation can be carried out during implementation, but this is not necessarily encouraged. The OD states that "supplementary data collection and special studies required for interim evaluations should be kept as simple as possible, and planned to *minimize interference with regular project operations*" (emphasis added). Evaluation was expected to be carried out at the completion of the project.

Focus groups conducted for the 2006 AROE found that this view of M&E is prevalent among Bank operational staff—that monitoring is about

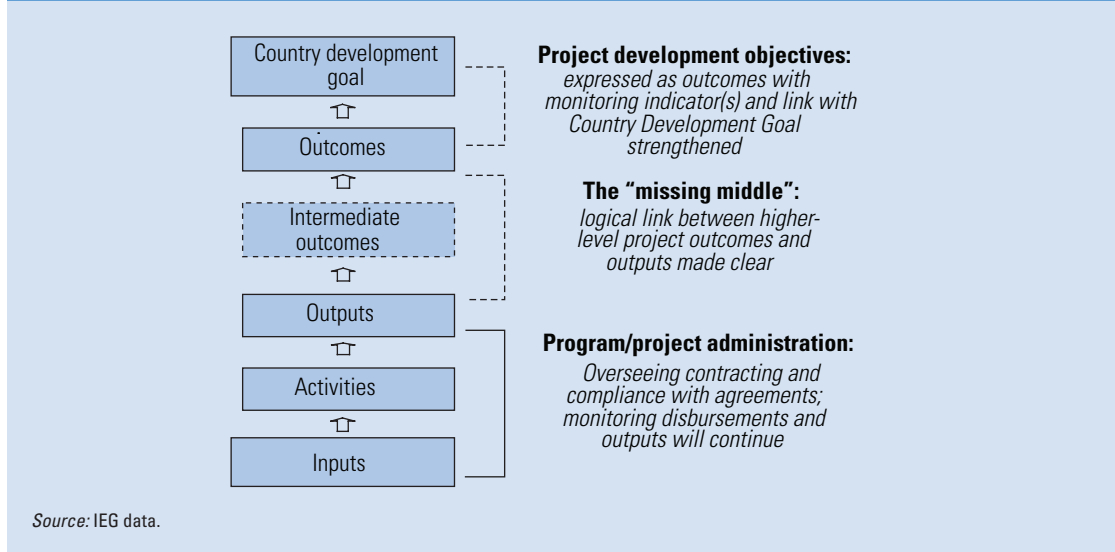
tracking inputs (disbursements), and possibly project outputs, while evaluation is done at the end when the "results" are clear. At the project level, evaluation has meant the ICR, which is then validated by IEG.

The Bank has been increasing its focus on the achievement of results, specifically the impact its activities have on beneficiaries. This has several implications for Bank operations.

First, the Bank will need to be clearer about what it expects to achieve through its operations. To accomplish this, project development objectives must be clearly defined and the changes expected at the beneficiary level clearly articulated. Indicators need to be found to measure and assess progress toward achieving not just the objective, but the expected changes as well.

Second, strengthening the focus on results underscores the importance of strategy. The link between outputs, outcomes, and higher country development goals will need to be strengthened and the results chain articulated. Intermediate outcomes that bridge higher-level outcomes and outputs will need to be identified, to address the gap often found between broad overall objectives and specific Bank operations (the "missing middle" in box B.1). New Bank procedures now require project results frameworks to specify intermediate outcome indicators and describe how they will be measured and used to track progress toward the achievement of project objectives.

Third, individual Bank operations need to be aligned with a country's development goals. Many actors and numerous factors contribute to

Box B.1: Implications of Results for Bank Operations

achievement of these higher-level country development goals. This increases the need to reach out to partners, clarify roles and contributions to the outcomes, and jointly monitor progress.

The move toward a stronger results focus has implications for how M&E is done in the Bank. These will be discussed in more detail in the following sections.

Monitoring

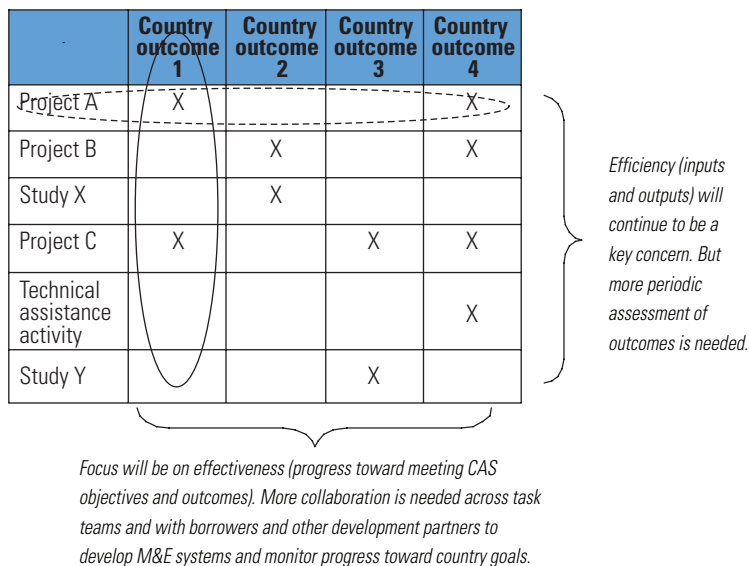
The increasing focus on results will require changes in the way the Bank approaches monitoring of its operations. At the activity level (projects and AAA), this will require an increase in the amount of information collected and monitored. Inputs (contracts, disbursements) and outputs will continue to be critical to project implementation, and the bulk of a team leader's time is expected to be taken up by traditional project and AAA administration. However, team leaders are now also being asked to monitor and track intermediate and final outcomes and to use that information in the ISRs.

The stronger focus on outcomes will likely change the way monitoring data are collected. As monitoring increasingly focuses on the

changes activities (projects and AAA) have on beneficiary behavior, implementing agencies may need to turn to third parties or to different instruments for information. For example, a road agency may no longer just report on the length of roads being built but may need to monitor the effect the roads have on potential beneficiaries (such as increases in the use of clinics and other social services in urban areas). This would require the road agency to commission a survey or work with other agencies to collect the necessary information.

Managers may need to revisit the design of their programs or projects during implementation. By asking what is working and not working with regard to outcomes, managers will be looking at which elements of the program or project are contributing to producing the expected impacts on the beneficiary and which are not. For example, schools could be constructed and textbooks delivered, but the children may not be learning because of poorly trained teachers, which may or may not have factored into the original design. During implementation, managers will be looking not only at the efficiency of the program or project (the rate at which inputs gets translated into outputs), but also at its effectiveness (the rate

Figure B.1: Implications of Managing for Results for Bank Country Operations



Source: IEG data.

at which it is producing the desired outcome for beneficiaries).

At the country level, the focus on results would strengthen Bank management of its overall country program. The Bank has now fully adopted RBCASs, which require results frameworks at the country level that link Bank activities (projects, studies, technical assistance) with the country’s development goals. RBCASs also include performance indicators (for CAS outcomes and intermediate outcome indicators or milestones to measure progress toward these outcomes) to track progress. An activity can contain more than one objective; for example, a health project may contribute to reducing child mortality and to improving public sector governance in the health sector. At the same time, more than one activity can contribute to a common CAS objective. Figure B.1 shows a typical RBCAS results framework, where several Bank activities with multiple objectives align against a common outcome.

A results approach would require greater collaboration and coordination among task teams

within the Bank and development partners at the country level. Task teams have been working—and will continue to work—toward meeting their project objectives, ensuring that inputs are provided and outputs delivered to achieve the desired outcome. That is, they will look horizontally across the matrix in figure B.1.

However, the overall effectiveness of the Bank’s country program will depend on the extent to which these activities work together toward meeting a common country-level (CAS) objective. That is, the vertical links between the activities become operationally critical for the Bank’s effectiveness and would be the focus of monitoring at the country level.

Coordination at the country level around results is already happening between the Bank and other donors through the Poverty Reduction Strategies (PRSs). For IDA countries, the PRSs provide a framework to link country development priorities with donor assistance around specific country development outcomes. The Bank and the other donors have committed to harmonizing their monitoring and reporting requirements

and to working through national systems to lessen the burden on the developing countries.

Rather than have individual project units collect outcome information, it may be possible to work with other interested parties to systematically collect the information at the country level. Some implementing agencies are reluctant to pay the additional costs (for example, to commission surveys of road users) required to collect outcome information, preferring to use their budgets to produce outputs instead (such as building more roads).

However, other parts of the government may be interested in obtaining outcome information. For example, the planning and budgeting unit of a ministry, local government, or government agency may want to compare the effectiveness of programs for budget allocation purposes. Rather than relying on the implementing agencies, they could organize surveys and other data collection efforts to produce a uniform set of performance information that could be used across all projects (both local and foreign assisted) under their control. This would reduce the burden of data collection and provide uniformity and transparency to program implementa-

tion. Several Latin American countries are developing governmentwide M&E systems in line with their own circumstances and needs (Zaltman 2006).

Monitoring systems should be designed with both the use and user in mind. Performance information can be used for communications, management, and learning (box B.2). The uses—and hence information needs—differ by function for both the borrower and the Bank.

Operational staff (Bank team leaders and government project and program managers) may consider outcome information useful for management, to ensure that an activity will have the expected impact on the beneficiaries, but their day-to-day focus will be on the input and output information required for administration. Operational staff also need to feed input and output information up the chain of command to address demands for accountability.

Those responsible for overseeing the overall direction of an agency or assistance program (Bank staff at the country level; staff in government agencies; and staff within agencies concerned with planning, budgeting, and

Box B.2: Possible Uses of Performance Information

To Communicate

- **Upward within an organization**—With superiors and external stakeholders, communicate priorities, clarify expectations and accountabilities, and report on results that are or are not being accomplished to address their demand for accountability.
- **Downward within an organization**—Clarify priorities and set expectations and accountabilities; motivate team/staff by setting clear, challenging, but realistic goals; and provide staff/team with periodic sense of accomplishment to examine lack of progress to focus future efforts or to celebrate success.
- **Sideways within an organization and with clients and other development partners**—Communicate priorities, clarify expectations and accountabilities, promote coordination by identifying higher-level outcomes that multiple units may have in common.

To Manage

- Monitor progress toward achieving expected objectives and outcomes.
- Identify problem areas that need attention/improvement.
- Determine staffing and financial resource needs.

To Learn

- Compare results (benchmark) across units/organization and identify good practices.
- Identify approaches that are working well (and could be replicated).
- Raise questions about why outcomes are not meeting expectations and trigger in-depth examination of performance problems.

Sources: Behn 2003; Harvard University Kennedy School of Government 2001; Hatry 2003.

human resource management) would be more interested in outcomes (what is being achieved) than in specifics of implementation (how it is being done). They would use performance information for resource allocation and program management. They too need to feed performance information up the chain of command to meet accountability requirements.

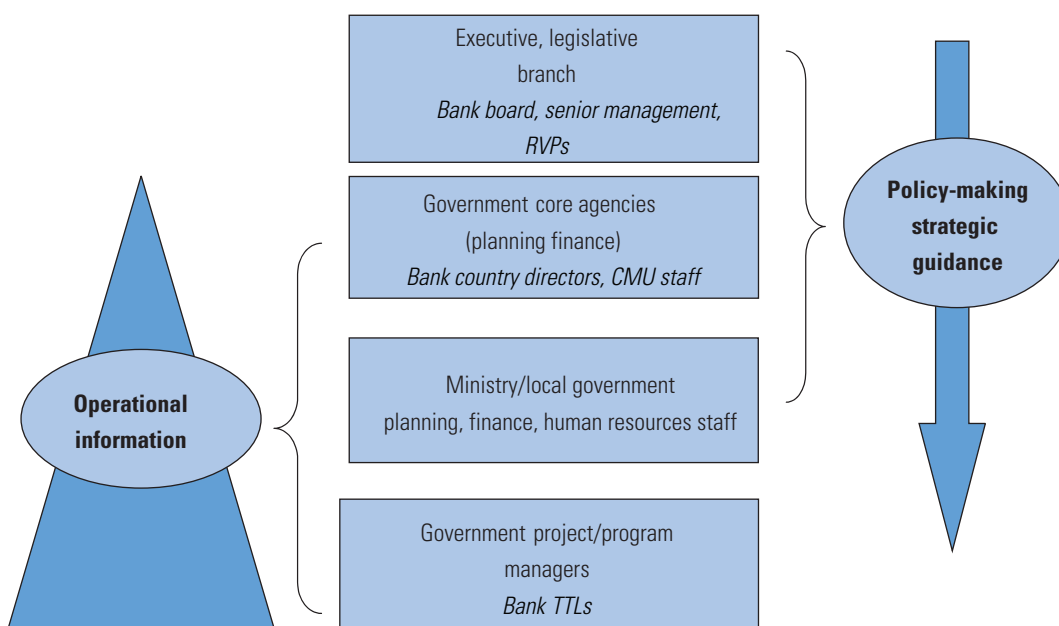
Finally, senior managers (in both the government and the Bank) would be interested in performance information, but largely to address accountability concerns of their enabling environment (the legislature, the Bank Board, and so on). Those higher up in the Bank or government hierarchy will be less concerned with operational information and more interested in strategic performance information (information that will tell them the results of the policies and programs they authorized); see figure B.2.

Learning, a critical use of performance information, often gets lost, especially when accounta-

bility is emphasized. There is an inherent trade-off between “managing for results,” which emphasizes learning what results are being achieved and why and feeding that information back into decision making, and “accountability for results,” where performance information is used for reporting and fulfilling accountability (Binnendijk 2001). Learning takes place at all levels. The operational staff are interested in information (such as good practices) that would be relevant to their situation, while senior managers would be more interested in benchmarking (comparing performance across projects or program, agencies, and the like) and more strategic issues.

In sum, the increased focus on the outcomes of Bank operations will affect what information is collected, how it is collected, and how it is used. It has implications for the way the Bank and the borrower approach monitoring. The main changes this will bring about are shown in table B.1.

Figure B.2: Audiences for Performance Information



Note: CMU = Country management unit; RVPs = Regional vice presidents; TTLs = task team leaders.

Table B.1: Moving from a Traditional to a Results-Focused Approach to Monitoring

Traditional approach	Greater results focus
Monitor progress on inputs and outputs.	➡ Monitor progress on inputs, outputs, and outcomes.
Carry out data collection primarily in house.	➡ Come up with alternative ways to collect information and/or work with other agencies.
Focus on activities, mainly investment projects.	➡ Focus on combinations of activities that share common outcomes at the sector, country, and/or global levels.
Use monitoring systems specific to individual activities.	➡ Monitor systems common across activities and development partners. Use institutional approach to M&E, working with client that may need similar information.
Have limited use of monitoring information beyond routine reporting.	➡ Use monitoring as a learning opportunity.

Evaluation

With the focus on results, managers and staff are expected to track not only inputs and outputs but also outcomes. They are expected to use progress or lack of progress toward achieving project outcomes to focus their efforts during implementation. Project and program outcomes will no longer be something to be determined and validated at the end through an evaluation. Increasingly, evaluative approaches are expected to be used not just at the end, but also during implementation of a project or program.

While both monitoring and evaluation will be concerned with achieving objectives and expected outcomes, evaluation would be needed to analyze why certain outcomes were achieved or not achieved. That is, monitoring would inform managers and staff about what is being achieved or not achieved, but evaluation would provide information on why.

A greater focus on results in Bank operations will change the timing and type of evaluation carried out. *When* managers and staff need evaluative information would change. To date, evaluations (ICRs, Activity Completion Reports, and so on) have been carried out after the completion of a project or activity (study, technical assistance). They have provided ex post information about what *worked* and what *did not work*, and *why*. With a stronger results focus, managers and staff

may want information on what *is working*, what *is not working*, and *why* in achieving outcomes during implementation and use that knowledge to make adjustments to achieve the expected results.

Evaluations would be carried out more frequently, not only during implementation but also before the start of a project or program. For example, program logic chain assessments could be done before a project or program is started to determine strength and logic of the causal model behind the design. Focus groups, community interviews, and other rapid assessments could be used during implementation to get a better understanding of the changes that the project or program is causing among target beneficiaries. These methods are summarized in table B.2.

Bank activities may be only one of many factors—including the role of other donor agencies—that contribute to an outcome. Questions of contribution and attribution would need to be addressed through evaluation. This would require more sophisticated methods, such as impact evaluations (table B.2). A greater focus on outcomes would require closer collaboration with the government and other development partners, including conducting joint evaluations.

In sum, the increasing focus on project outcomes during implementation will blur the

Table B.2: Some Uses of Evaluation to Better Manage for Results

Question	Issue	Evaluation approach
Do we have the right strategy and project/program design?	The results chains in the projects are weak. Task teams struggle to link the project with higher-level country development goals and CAS outcomes, especially to determine key intermediate outcomes that could be used for M&E.	Program Logic Chain Assessments determine the strength and logic of the causal model behind a policy, program, or project. The model addresses the deployment and sequencing of the activities, resources, or policy initiatives that can cause the desired change in an existing condition. The assessment would address the plausibility of achieving the desired change, based on the record of similar efforts and on the research literature. The intention is to avoid failure from a weak design that would have little or no chance of success in achieving the intended outcomes.
Do we fully understand what is happening at the beneficiary level?	Project/program managers may not have a full understanding of the changes that the project/program may be causing to the beneficiaries. This could be caused by a lack of appropriate beneficiary data or difficulties in comprehending quantitative monitoring data.	Rapid appraisal methods provide quick assessment data from beneficiaries and stakeholders on the progress of a given project, program, or policy. It is a multi-method evaluation approach that would involve (a) key informant interviews, (b) focus group interviews, (c) community interviews, (d) structured direct observation, and (e) surveys. Rapid appraisals are quicker and less costly than formal surveys, but are also less valid and reliable.
To what extent were the outcomes a result of the project/program?	Many factors affect the target beneficiary groups, so it is difficult to distinguish the impact of the project from that of other factors and to understand how the project/program affected the beneficiaries (positively/negatively, intended/unintended).	A rigorous impact evaluation identifies a counterfactual to analyze the situation, with and without the project/program. It establishes the impact of the project/program and address questions of causality and attribution.

Note: CAS = Country Assistance Strategy; M&E = monitoring and evaluation.

traditional distinction between monitoring and evaluation—that evaluation is carried out at the end to assess the results of a project or program. A results focus will influence when and how evaluations are done. Evaluations can take place during the preparation of a program and during

implementation, as well as at the end. Evaluation would function more as a management tool, so that more self-evaluations are expected to be carried out by the Bank. Table B.3 summarizes the differences between traditional and results-focused approaches to evaluation.

Table B.3: Moving from a Traditional to a Results-Focused Approach to Evaluation

Traditional approach		Greater results focus
Monitoring seen as an internal management activity, while evaluation carried out by external, independent entity.	➔	Greater use of self-evaluation by management Independent validation needed to verify self-evaluations.
Evaluations usually carried out (ex post) at the completion of the activity to determine what worked, what didn't work, and why.	➔	More evaluations (self- and independent) carried out before start of project/program to examine what would work and what would not work and during implementation to determine what is working, what is not working, and why.
Evaluations focused primarily on single activities (such as a project) and carried out separately or in parallel with other government and/or development partner activities.	➔	While the Bank would continue to conduct evaluations of its own operations, evaluations increasingly carried out jointly with government and other development partners.

APPENDIX C: REVIEW OF CAS RESULTS FRAMEWORKS

Background

The Bank is strengthening the results focus of its CASs. The results-based CAS approach, which started on a pilot basis in 2003, has been adopted as Bank policy and is incorporated in the revised Bank Procedure (BP) 2.11, Country Assistance Strategies, issued in June 2005. Two key elements of the new approach are the preparation of a CAS Completion Report (CASCR) to review the effectiveness of the previous program and a CAS results framework (results matrix), which outlines the outcomes expected from Bank support during the CAS period.¹ The revised BP has a separate section on implementation, which discusses M&E, use of country systems, and the links of Bank operations with M&E.²

The change was instituted following the finding of a review of CASs that the treatment of results was the weakest element (World Bank 2003a). The review found that CASs described in detail their objectives and the Bank instruments to be used, but explicit links between the two were often missing. The CASs contained self-evaluation frameworks (program matrices) that contained large sets of performance targets and indicators, but only about half were specific, reasonably quantitative, and time bound. Many did not differentiate between overall country goals and outcomes influenced by the Bank program or articulate the outcomes that the Bank or the country wanted to achieve. The results-based approach is intended to overcome these issues.

The CAS Results Framework

The CAS results framework provides the basis for monitoring and evaluating the Bank's

country program. The framework is expected to lay out the logical relationships between Bank operations and outcomes and the measurable indicators to track key links. It contains the following elements:

- **Country development goals** are the longer-term or higher-order development objectives identified by the country, which may not be achieved during the CAS period or are not solely addressed by the CAS.
- **CAS outcomes** are highest-order results achievable during the CAS period; the Bank expects to influence these directly through its operations.
- **Milestones** are the outputs, actions, or outcomes that are expected to be realized during CAS implementation and can serve as progress markers toward CAS outcomes.
- **Bank (and partner) programs** represent the ongoing and planned lending and non-lending activities of the Bank (and key partners) that are expected to be influenced by CAS outcomes.

The results chains linking the country development goals, CAS outcomes, milestones, and Bank operations should be visible in the results framework, clarifying the Bank's contribution to achieving country development goals.

The CAS results frameworks are intended to be used as a program monitoring and management tool during CAS implementation and to provide the basis for evaluating the effectiveness of the Bank's program. As a monitoring and management tool, the CAS results framework will need to focus on a manageable number of performance indicators to allow for regular monitoring

of progress toward CAS objectives. As an evaluation tool, the CAS results framework will need to contain sufficient performance information to allow for a thorough assessment of the program. In both cases, the performance indicators must be specific, measurable, baselined, and targeted. CASs could include a comprehensive results framework but identify a smaller set of core outcomes and indicators that can be used to manage the program during implementation.

In addition to the CAS results frameworks, the Bank is also developing sector results frameworks that link the expected final outcomes, intermediate outcomes, and expected Bank contributions as part of Bank-wide sector strategies. The intent is to establish the links between final outcomes, intermediate outcomes, and the Bank's contribution to achieving these results within a sector. This, in turn, would allow the Bank to track progress on these outcomes and its contributions to these outcomes. Sector strategy and performance information will need to fit with the country priorities, plans, and needs. This raises the question of the extent to which a global framework can be adapted across multiple countries with different priorities and levels of Bank engagement.

Objectives, Methodology, and Scope

Objectives of the review

The objective of the review is to assess the quality and usefulness of the results framework in managing the Bank's country program. The CAS results frameworks were examined to assess (a) the strength of the links between country development goals, CAS outcomes, milestones, and the Bank program; (b) the usefulness of the performance measures for M&E; and (c) the extent of integration between CAS and sector results frameworks.

Methodology and scope

A maturity model³ was used to evaluate the results frameworks and performance measures. Assessment criteria were established, and four levels of maturity, from level 0 (lowest) to level 3 (highest),

were defined. The evaluation framework is presented as an attachment to this appendix.

First, the overall results frameworks were examined to establish the extent to which CAS results frameworks specify the expected linkages between Bank interventions and the country's long-term goals. Results frameworks with the strongest links were rated level 3. The assessment criteria for each level are outlined in the attachment.

Second, the performance measures for country development goals, CAS outcomes, and milestones in the results frameworks were assessed to establish their usefulness for M&E based on the following criteria:

- **Measure:** The measure is clearly identifiable as a performance measure, distinct from a goal or program activity.
- **Baseline:** The measure includes the current level of performance as a baseline.
- **Target:** The measure includes a specified, time-bound performance target.

In addition, milestones were assessed for the following criteria:

- **Relevance:** Relevance to the objective the program intends to achieve; can serve as good proxy for program outcomes
- **Precision:** Describes the specific achievement and when it will be achieved and provides enough information to assess program implementation.

Results frameworks with more robust measures (those that showed more of the above qualities) were rated level 3. The assessment criteria for each level are outlined in the attachment.

Twenty-five CASs were included in the review for the 2006 AROE.⁴ These included CASs from fiscal 2005 that were not included in last year's AROE and those prepared from that time to March 2006. An initial review of CAS results frameworks (results matrices) was done for the 2005 AROE. However, the results are not strictly comparable

because the analysis was carried out during a transition period when CASs were increasingly following the results-based format, but before RBCASs were formally adopted by the Bank

Third, a provisional assessment examined the linkages between country and sectoral results frameworks. The review was limited to the sector results frameworks for water and sanitation from the 2005 SSIU paper.

Findings

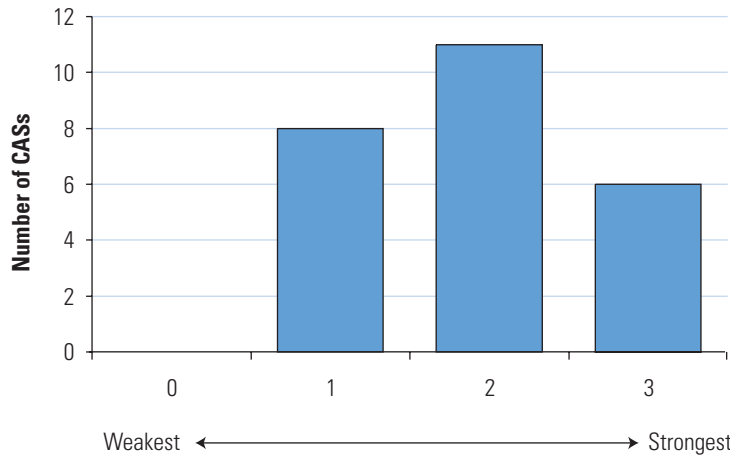
Results chain strength

The examination of the strength of the links among country development goals, CAS outcomes, milestones, and the Bank program (results chain strength) revealed scope for improvement. All of the CASs followed the RBCAS format. A few countries made some modifications—for example, the Uganda CAS did not distinguish between country development goals and CAS outcomes.

Of the 25 CAS results frameworks reviewed, six (24 percent) were rated at level 3, as having relatively stronger results chains (results chains are generally visible and most country goals are logically supported by Bank activities); 11 at level 2, as having modest results chains (some country goals are logically supported by Bank activities, but many goals are less effectively portrayed); and 8 at level 1, with poor results chains (the link between country goals and Bank activities was not clear for most country goals). No CASs were rated at level 0 (figure C.1).

Given the complexity of the results frameworks with multiple objectives, multiple sectors, and multiple indicators, an assessment of overall results chain strength for the entire framework is a matter of judgment. There was a general convergence of quality around the middle. In the least well-articulated results frameworks, it was very difficult to follow the logic among country development goals, CAS outcomes, and milestones because they often consisted of broad goals and bullet points. In the strongest results frameworks, the country teams had made an effort to link Bank operations

Figure C.1: Results Chain Strength

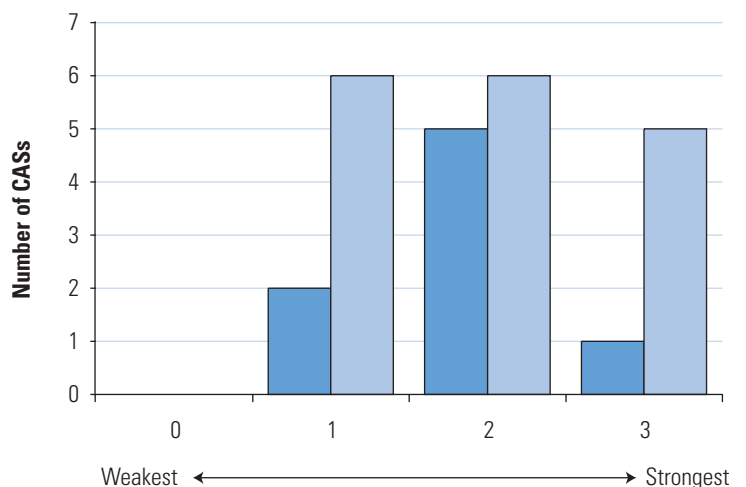


Source: IEG data.
Note: CAS = Country Assistance Strategy.

to CAS outcomes and attempted to identify the intermediate outcomes and milestones necessary to achieve the final (CAS) outcomes and defined baselines and set targets.

No clear differences were discernable by fiscal year (figure C.2). Of the eight fiscal 2006 CASs,

Figure C.2: Results Chain Strength by Fiscal Year

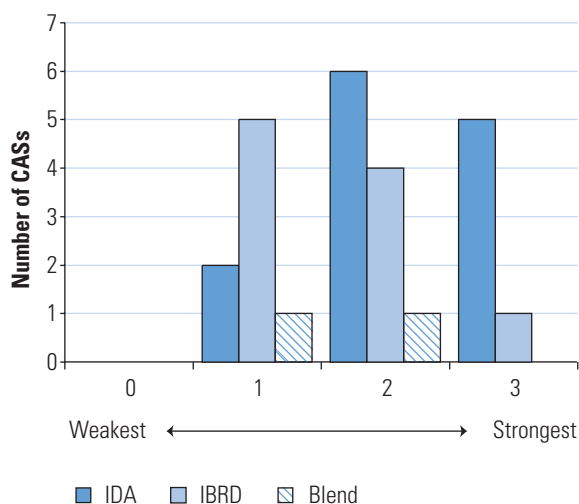


Source: IEG data.
Note: CAS = Country Assistance Strategy.

five (63 percent) were rated at level 2. The distribution of 17 fiscal 2005 CASs is quite even: 5 (30 percent) were rated level 3, 6 (35 percent) were rated level 2, and 6 (35 percent) were rated level 1. The analysis was conducted using CASs in the Bank’s publications database that were

presented to the Board from the fourth quarter of fiscal 2005 to the third quarter of fiscal 2006. It is very likely, given the time it takes to prepare a CAS and the factors that determine when it is presented, that most of these CASs were prepared around the time the RBCASs were being adopted. So it is probably too early to draw any conclusions based on fiscal year differences.

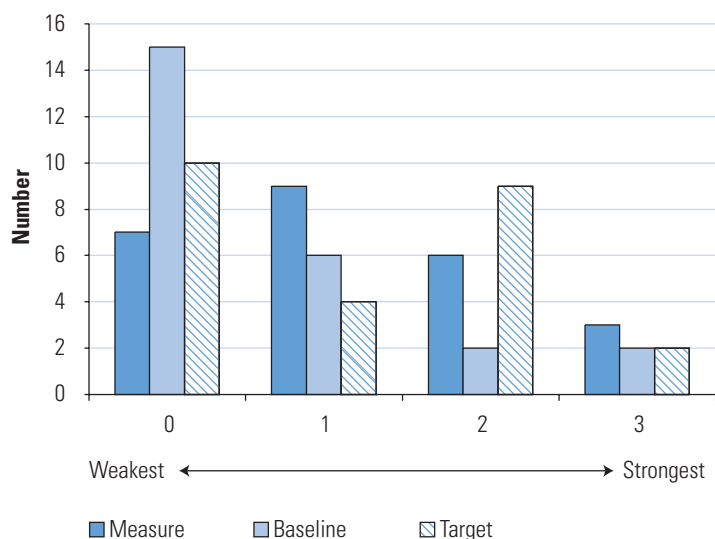
Figure C.3: Results Chain Strength by Country Type



Source: IEG data.
 Note: CAS = Country Assistance Strategy.

Significant differences are evident in the comparison by country type—IBRD, IDA, and IBRD-IDA blend (figure C.3). The linkages between country development goals, CAS outcomes, milestones, and the Bank program were rated more highly in IDA countries than in IBRD or IBRD-IDA blend countries. Of the 25 CAS reviews, 10 are IBRD, 13 IDA, and 2 IBRD-IDA blend. Among IDA countries, 11 (85 percent) were considered to have moderate to strong overall results chains (levels 2 and 3). Among the IBRD countries, nine (90 percent) were considered to have weak or moderate results frameworks (levels 1 and 2). This result is not surprising, because IDA countries have experience working with PRSs, which set and track country-level performance measures.

Figure C.4: Country Development Goals



Source: IEG data.

Performance measures

There is scope for strengthening the performance measures in the CAS results frameworks. In more than half of the frameworks reviewed, the majority of the country development goals were not clearly identifiable as performance measures; did not include baselines; and did not set a specific, time-bound target (figure C.4). That is:

- **Measure:** Seven CASs (28 percent) had few (level 0) and nine (36 percent) had less than 50 percent (level 1) of their country development goals clearly identified as performance measures, distinct from a goal or program activity.
- **Baseline:** Fifteen CASs (60 percent) had few (level 0) and six (24 percent) had less than 50 percent (level 1) of their country development goals include the current level of performance as a baseline.
- **Target:** Ten CASs (40 percent) had few (level 0) and four (16 percent) had less than 50 per-

cent (level 1) of their country development goals include a specified, time-bound performance target.

Overall, the review demonstrated that baselines are the area most in need of improvement.

Country development goals reflect the country's development program. Often these goals are given as strategic objectives. The availability or lack of specificity, targets, and baselines most likely reflect the level of detail available in the government plans linked to the Bank CAS.

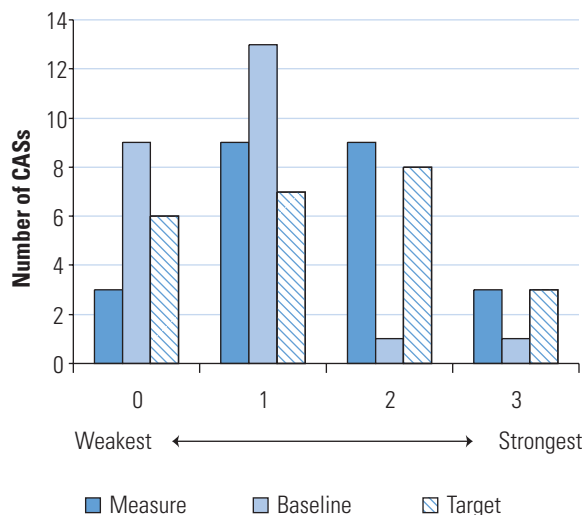
The assessment of CAS outcomes exhibits a similar pattern. In about half of the CASs reviewed, CAS outcomes could be identified as a performance measure. But in more than half of the frameworks, the majority of the CAS outcomes did not include baselines and did not set a specific, time-bound target (figure C.5). That is:

- **Measure:** Three CASs (13 percent) had few (level 0) and nine (38 percent) had less than 50 percent (level 1) of their CAS outcomes clearly identifiable as performance measures, distinct from a goal or program activity.
- **Baseline:** Nine CASs (38 percent) had few (level 0) and 13 (54 percent) had less than 50 percent (level 1) of their CAS outcomes include the current level of performance as a baseline.
- **Target:** Six CASs (25 percent) had few (level 0) and seven (29 percent) had less than 50 percent (level 1) of their CAS outcomes include a specified, time-bound performance target.

Again, baselines were the weakest area: baselines were provided for fewer than half the CAS outcomes in 92 percent of the CASs reviewed.

The results frameworks include milestones that are outputs, actions, or outcomes expected to be realized during the CAS implementation period, and they can serve as progress markers toward CAS outcomes. Most of the milestones included in the results frameworks were actions (such as regulation adopted) or outputs.

Figure C.5: Country Assistance Strategy Outcomes



Source: IEG data.

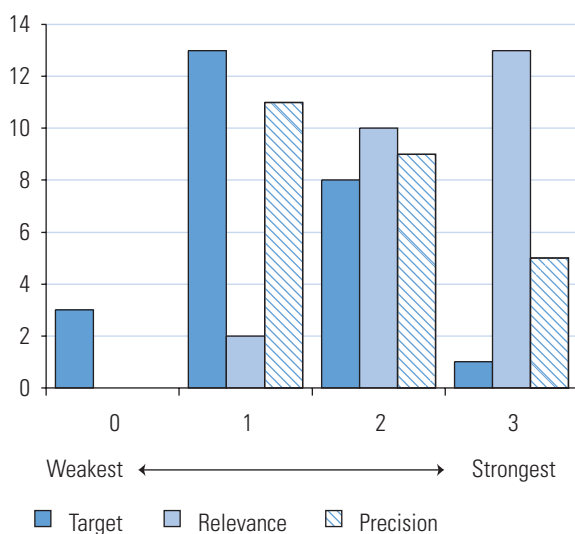
Note: CAS = Country Assistance Strategy.

Therefore, the review focused on whether they were well targeted, relevant to the CAS outcome, and precisely defined.

Overall, actions that contributed to CAS outcomes were well articulated. Most of the milestones were judged to be relevant, but precision of definition could be improved and targeting was weak (figure C.6). That is:

- **Target:** Three CASs (12 percent) had few (level 0) and 13 (52 percent) had fewer than 50 percent (level 1) of their milestones include a specified, time-bound target.
- **Relevance:** Ten CASs (40 percent) had 50 to 75 percent (level 2) and 13 (52 percent) had more than 75 percent (level 3) of their milestones considered relevant to the objective that the program intended to achieve.
- **Precision:** Eleven CASs (44 percent) had fewer than 50 percent (level 1) and another nine had 50 to 75 percent (level 2) of their milestones adequately describe the nature and timing of the specific achievement and provide enough information to assess program implementation.

Qualitative differences in performance measures were observed between the sectors.

Figure C.6: Milestones

Source: IEG data.

Measures tended to be strongest in macroeconomics, education, and health, where the Bank has been more active in using numerical indicators, or where, as in the social sectors, the MDGs have provided specific numeric targets.

Although there is no “correct” number of indicators for a country program, the numbers appear to be high. There were, on average, 24 country development goals for each country, and 6 countries had more the 30 goals (table C.1). The average for CAS outcomes was 41 goals for each country, and 5 CASs had more than 60 outcomes. The average for milestones was 60 for

each country; 10 CASs listed more than 60 milestones. Large numbers of indicators present a challenge for collecting information and monitoring progress. Moreover, without baselines and targets, the indicators would not be useful for carrying out an evaluation of a completed CAS.

Larger country programs can be expected to have a larger number of indicators. For simplicity, the Bank’s fiscal 2006 country budget was taken as a proxy for Bank country program size. Figure C.7 presents the distribution of the number of CAS outcomes by country budget and figure C.8 the distribution of the number of milestones by country budget. Overall, the number of both CAS outcomes and milestones increases with country budget size, indicating a need for additional measurement for larger programs. Some outliers can be observed in the middle range (\$4 million to \$6 million), indicating scope for examining and possibly reducing the number of indicators in individual countries.

Links between Bank-wide sector and country results frameworks

IEG conducted a cursory comparison of the water and sanitation sector results framework and CAS results frameworks to review the extent to which a Bank-wide sectoral framework could be incorporated into the CASs.

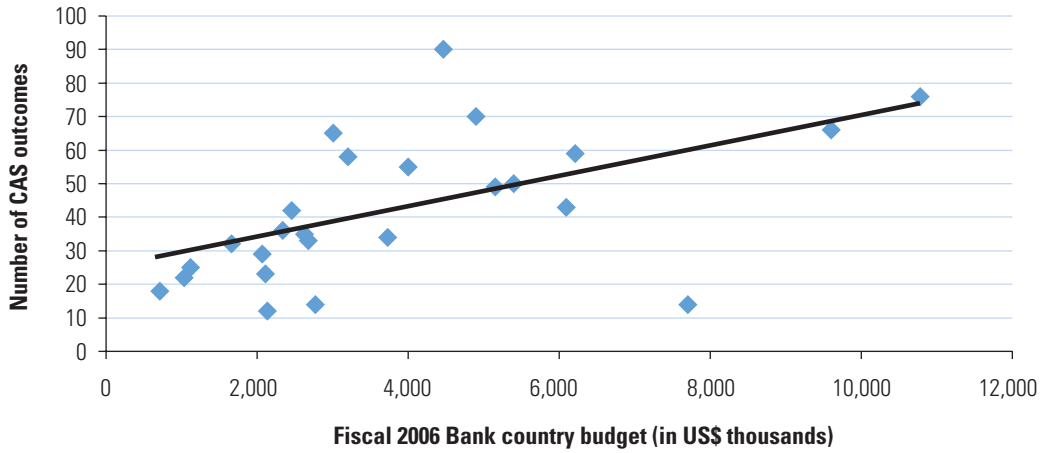
The water and sanitation results framework is structured around MDG Target 10, “Halve the number of people without sustainable access to

Table C.1: Number of Countries by Total Number of Indicators

	Country development goals	CAS outcomes ^a	Milestones
Fewer than 9	3		
10–29	16	8	4
30–59	6	11	11
More than 60		5	10
Total countries	25	24	25
Average number per country	24.4	41.3	59.5
Maximum number	59	90	125

a. Excludes Uganda, where CAS outcomes were combined with country development goals.

Figure C.7: Country Assistance Strategy Outcomes and Country Program Size



Source: IEG data.

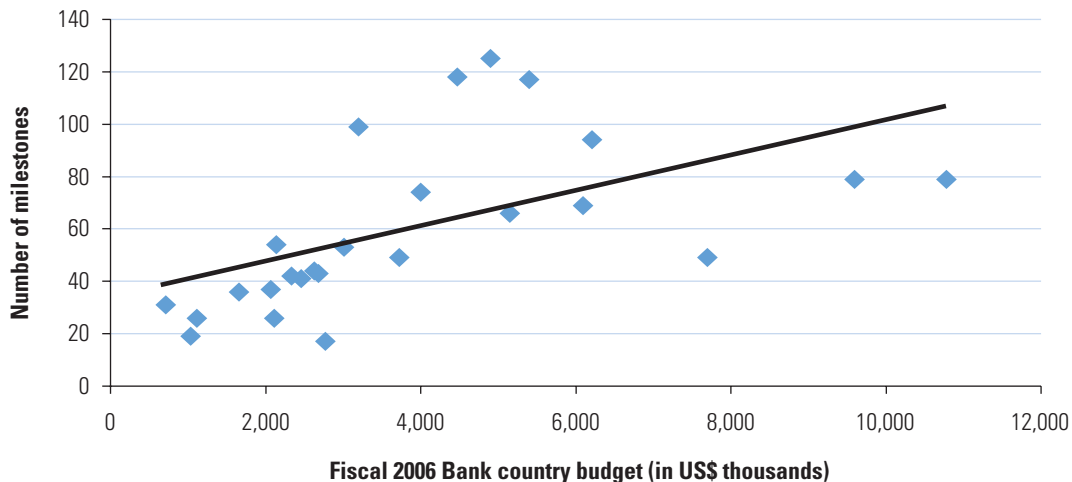
Note: CAS = Country Assistance Strategy.

safe drinking water and basic sanitation” and the IDA 14 RMS, which includes “Proportion of population with access to a secure water source” as one of the country outcome indicators and water supply as one of the areas for reporting Bank outputs. The results framework sets “the number of people with sustainable access to safe drinking water” as the final outcome and the “number of people with access to improved drinking water sources” as the intermediate

outcome. Performance measures specify the number of people and the proportion of the population that have met the outcomes.

Of the 25 CASs in the review, 15 contained direct references to water or sanitation. Others referred to improvement in municipal services where water and sanitation could be key elements. Of these 15, 8 included the number or percent of people serviced or with access to

Figure C.8: Milestones and Country Program Size



Source: IEG data.

improved water as a performance indicator for either for country development goal or CAS outcome. While the sectoral results framework focused on increasing access to water and sanitation, many CASs also emphasized the institutional aspects, such as increasing cost recovery and financial viability of a water utility, or reducing system losses and improving operational efficiency.

This analysis shows the difficulties of balancing simplicity and clarity that a results framework demands with the complexities of Bank strategy, which is adapted to different country situations. The water and sanitation results framework is based on clear international and corporate priorities, MDGs, and IDA 14. It provides a single performance indicator that reflects the strategic objectives to be achieved. But devising comparable data that can be aggregated across countries may not be straightforward. The SSIU reports that the Bank is working through the UNICEF-World Health Organization Joint Monitoring Program to harmonize household surveys to increase the comparability of data across countries. In addition, the results framework (focused on increasing access or supply) does not fully reflect the institutional aspects of the Bank's work, which is critical to sustain development. In many CASs, improving cost recovery, reducing operational losses, and strengthening management capacity were seen as key sectoral CAS outcomes.

Overall Conclusions

Comparing the CAS results frameworks reviewed this year with those reviewed last year, it appears that the creation of a single, Bank-wide structure for the presentation of results matrices has had a beneficial effect on the strength of the matrices.

The least well-developed frameworks this year appeared to be better structured than the least well developed last year. Overall, the consistent categories of measures and definitions appear to have formed a foundation for the logical structure of relationships between outcomes and activities.

Although the structure has improved, there is still wide variation between the quality of measures and the consistency of the contents of each one of the columns of the results framework. This appears to reflect a tendency to try to include too much at each level of the results framework, especially at the levels of CAS outcomes and milestones. Because of this, carefully defined and otherwise salient indicators of country progress and Bank success can be swamped by indicators that are less well considered.

In addition, many of the performance measures may not be useful for M&E because they lack precision regarding the nature of the expected change, the baselines, and the targets. Data alone are not useful unless they are put in the proper context—that is, do we know where we want to go and how far have we come? Although this was not a subject of this review, the lack of baseline data raises questions about whether the information is actually available.

Based on the evidence, country teams need to revisit their country results frameworks, establish priorities, and reduce the number of indicators, especially Bank CAS outcomes and milestones, and ensure that the remainder are properly defined, baselined, and targeted. The large number of CAS outcomes and milestones may reflect a need for greater selectivity.

Attachment

Country Assistance Strategy: Results Matrix Evaluation		
Country		
Date of CAS		
Type of country [IBRD/IDA/IBRD-IDA blend]		
Reviewer		
Framework level		
Criterion	Evaluated matrix maturity level	Criteria for evaluation
Results chain strength		Extent to which CAS results framework specifies the expected linkages between Bank interventions and the country's long-term development goals (includes how far along CAS results lie in the causal chain toward country development goals). The results chain should be visible in the results matrix such that the links among country goals (country development goals), Bank goals (CAS outcomes), and Bank activities (milestones) are clear.

Maturity levels:

Level 0: Absence of evidence in the criterion area

Level 1: Most Bank goals and activities seem misaligned with the country goals, or a profusion of goals and activities buries the logical, causal relationships.

Level 2: Some pillars or high-level goals are well organized and supported by succinctly stated Bank activities and goals, but other areas are less effectively portrayed.

Level 3: Most pillars or country goals are logically supported by succinctly stated Bank activities and goals.

Results / Measures Level				
Matrix element	Evaluated result/ measure maturity level			Criteria for evaluation
	Country development goals	CAS outcomes	Milestones	
Measure				The measure is clearly identifiable as a performance measure, distinct from a goal or program activity.
Baseline				The measure includes the current level of performance as a baseline.
Target				The measure includes a specified, time-bound performance target.
Milestone relevance				Milestones are relevant to the objective that the program intends to achieve; can serve as good proxies for program outcomes.
Milestone precision				Milestones describe the specific achievement and when it will be achieved and provide enough information to assess program implementation.

Maturity levels:

Level 0: Very few results/measures in the matrix demonstrate the defined qualities.

Level 1: Fewer than half of the results/measures demonstrate the defined qualities.

Level 2: More than half, but fewer than three-quarters, of the results/measures in the matrix show the defined qualities.

Level 3: More than three-quarters of the results/measures in the matrix exhibit the defined qualities.

APPENDIX D: FOCUS GROUP AND INDIVIDUAL INTERVIEWS WITH CODE MEMBERS, COUNTRY DIRECTORS, SECTOR MANAGERS, AND TASK TEAM LEADERS

This appendix summarizes the findings from a series of focus group discussions and structured interviews of staff and managers regarding their perceptions of the Bank's M&E system and the application of M&E information.

The study was conducted during March 2006 by Group Dimensions International to provide information for the 2006 AROE. A total of 66 people attended 13 sessions that were organized separately for task team leaders (TTLs), sector managers (SMs), and country coordinators (CCs). Individual interviews were conducted with 11 members of the Board's Committee on Development Effectiveness (CODE) and country directors (CDs). The research focused on four areas:

- The use and usefulness of M&E information
- Application of M&E information to implement operations and manage for results
- Challenges, incentives, and disincentives to using M&E information
- The use and usefulness of IEG evaluation findings, products, and information.

The nature and purpose of focus groups is described in more detail in the attachment to this appendix. Unless otherwise specified, the findings summarized here reflect themes that arose across all respondent types.

Use and Usefulness of M&E Information

Respondents were clear regarding the distinctions between monitoring and evaluation and the relationship between them.

Monitoring provides the shorter-term "process and progress" picture. TTLs use monitoring for

corrections in design, allocation of funds, resetting priorities, and guiding staff. Evaluation reviews monitoring and affords reflection across programs, sectors, and countries, and for a longer term. From the point of view of TTLs and SMs, monitoring means working with the ISR, and evaluation signifies working with the ICR.

- *The challenge is more on evaluation than on monitoring. Monitoring is a more natural exercise. It's part of supervision; you monitor input, processes, you count outputs; something you relate easily with. It is done in a formal way or not, but done in a much better way. The evaluation is a more tricky exercise if you don't design that well at the very beginning and you have means to collect all the information, prepare for the analysis, and all of that. This is something very, very tricky sometimes. (TTL)*
- *People should be aware that those evaluations also look backwards, but of course you derive from it recommendations for the future. It is an exercise based on past experiences, and in that sense the outcomes may be linked to the political priorities of that time. Sometimes people have expectations which are not clearly justified from an evaluation exercise viewpoint, but you should use this type of retreat for reflecting the issues the Bank has been engaged in and that were top priorities in our strategic framework. (SM)*

TTLs practice two kinds of monitoring.

The first is a less formal, ongoing collection of information as part of oversight and the second, systematic, continuous, and regular collection of information based on a predefined set of indicators.

- *During implementation and supervision, you are essentially in this two-dimensional world—the more formal and structured M&E that you are doing, which has been designed during project formulation, and then there is the informal stuff, supervision, etc. Right at the end of a project, you have the more bureaucratic form of M&E, which is the ICR, etc. That may be the only time you actually return to the benchmarks and indicators you set up in the preparation. (TTL)*
- *The Bank does a lot of monitoring. When you're out in the field, you're focused on problem solving. There are a myriad of issues and problems in a project after you thought everything was nicely set up at the PAD stage and it went to the Board. You have to deal with these—and the client has to deal with these. The biggest problem is dealing with the obstacles that hinder a project that might not have anything to do with M&E, because otherwise there will be no project. (TTL)*

Respondents saw evaluation results (and monitoring information) being used for both accountability and learning.

They saw IEG's role as providing both accountability and learning, but thought IEG may have shifted more toward accountability. They asked why IEG should not stress more learning.

- *You don't want to repeat the mistakes of the past. You do need to make sure you have some good information from monitoring of past activities and broader evaluations. (CD/CC)*
- *We would be more interested in themes that push the sector forward, not things that we've been working on for 30 years. (TTL)*

Managers and staff learn about evaluation results in both formal and informal ways.

A great deal of wisdom about lessons, best practices, and problem solving occurs informally through discussions with colleagues, supervisors, and clients about what one can reasonably expect to achieve. In addition, participants saw discussions among operational staff and between the evaluator and staff during an

evaluation as potential opportunities for identifying and conveying lessons.

- *The CD knows that next year he/she will produce a new CAS and will start with the team to reflect seriously on what is going on now. If there is a good internal debate and partners were consultants, and there were lessons learned that were clear, I don't care if the report was not out...it is impossible to require this. Most important are the learning processes. (CODE)*
- *It is much more useful during implementation to have an annual program review [Ministry together with Bank TTL] and you look at the data together and determine whether or not you are moving in the right direction. This process has much more impact on the operation and implementation than some report [does]. I have seen that we have to shift, or to focus on a certain issue that was revealed by the study. (TTL)*

Respondents said they learn about evaluation results and obtain knowledge from formal reports related to their work.

They also mentioned that knowledge from self-evaluations and IEG reviews of the CASCR and ICR often entered the discourse *too late in the process to have a definable, systematic influence on the future of Bank work.*

- *By the time the auditor report comes out, we have moved on to the next phase or the portfolio is finished. (SM)*

Evaluation results do not flow into future work smoothly or systematically through mechanisms set up in the project cycle.

The project cycle was used in the focus groups to discuss where and how M&E results entered into the process. Respondents challenged the usefulness of the notion (expressed in the cycle) that the ICR and the IEG review of the ICR lead to "lessons learned," which are then codified as a mandated "lessons learned" section of the PAD. Many saw learning as a continuous process rather than as something to be done at certain intervals.

- *The cycle bears no relationship to reality—it might in a more structured environment but not here. The project cycle is a business model designed by engineers and is very logical and linear; the actual operating environment in the World Bank and in the Regions is chaotic. (SM)*
- *The project cycle is like an Apollo moon launch, a very calculated progression from one step to the next; the actual work is not. The real project cycle looks more like an annual learning cycle, than the current project cycle. TTLs need M&E information on an annual basis. (TTL)*
- *If people can be aligned around a common set of indicators, they will be aligned around the agenda. Once you have that, you can speak to that agenda through the indicators, which is nonconfrontational and is more constructive. (SM)*
- *It is all about gathering information and making changes. You have different levels of managing tools. You keep an eye on where you want to go and use information about how well you are doing to influence how you use leverage to get results. We are a step back from the front line. (CD/CC)*

Application of M&E Information to Implement the Bank's Managing for Results Agenda

Respondents at all levels were clear that M&E information and knowledge allow Bank staff, clients, and others to obtain a clearer picture of how well goals are being achieved.

Respondents agreed that managing for results entails serious efforts to establish and measure indicators in the field to ensure that agreed-upon objectives are being met.

- *Be explicit about what you want to do, collect good information about how well you are doing, and use that information to impact the way you move forward. (CD/CC)*
- *One must think of a project in terms of paying someone to deliver an outcome (not inputs). Instead of buying cement for building a school, we pay for the increase in enrollment of girls in school. If you are really walking the talk, you pay for outcomes. (SM)*
- *We are interested in results on the ground, in the country. How are our projects working to decrease poverty? Are the projects working? We are not interested at looking at Bank outputs; we look beyond that to what is happening in the country. (CD/CC)*
- *Managing for results means that we are not only spending money, not only helping countries, but making sure that we have some results on the ground. Maybe the main goal is not achieved, but at least that there is some progress. The result for me is something that you have on the ground and you are progressing in the right direction. This is important for donor countries but also for recipient countries. Sometimes people are aware of the aid we receive, but we also have to show that this produces outcomes. (CODE)*
- *Ultimately, it is the reduction of the rate of poverty, the reduction or increase in education for all, decrease in gender equality—all these MDGs or other development aspects, cleaner energy, etc. are measurable. These are ultimate goals and there are some sub-goals that are more important. Even though the numbers are not good in the short term, the Bank assistance may be doing a better thing for now, which ultimately [will] be more helpful to the countries than the numbers. (CODE)*

Respondents saw “managing for results” slightly differently, depending on their vantage points.

For SMs, managing for results meant “making sure a program fits outcomes and objectives for individual projects” within the context of broad country/sector goal. For CDs and CCs, it meant country-level outcomes rather than outputs. For CODE members, it meant a way to manage an institution, program, or plan to achieve a goal with a (usually) quantitative target, but sometimes also a qualitative target.

Feedback about the usefulness of the results framework is mixed.

For some respondents, the results framework triggers thinking about the realism of indicators and brings a more strategic intentionality

in working with clients. Others saw the results framework as rigid in that it was difficult to integrate new indicators during implementation and to take into account the effects of external factors or risks that can delay or derail results. The results framework does not always reflect current country circumstances, which can mean that unintended and unexpected achievements remain undetected by formal M&E.

- *It is a useful tool, because it focuses on what we want to achieve and then to align the activities to achieve these results. It also brings more coherence, a more strategic view to the program. (CD/CC)*
- *The results framework and using results as the indicators is actually the right way to go. It makes more sense. It does make it more difficult to monitor; more difficult to identify outcomes and link them to certain actions. Evaluation becomes more difficult, because you are no longer saying that our success is there if we build so many schools. Now we are successful if the number of girls in schools increases. There are so many factors that it becomes impossible to determine whether the project was successful or not. I still think it is the right way to go. It also will make evaluation more difficult—by nature. (CD/CC)*
- *I don't think it is useful at all if the TTL has drawn a results matrix sitting behind his desk. The use of the results matrix is in building the ownership of the different stakeholders in the program, whether they are from the sector, internal stakeholders, or external stakeholders, including the government. They are usually done by looking at reading materials, talking to people.... There is a bit of a disconnect with what is actually happening in the country. There is a missing link. They are not specific enough. They don't go enough in depth in the country analysis. (CD/CC)*

Managing for results has implications for relations with clients and team building.

CCs and CDs, in particular, see that the results orientation of the Bank, in combination with the new tools developed for managing for results

(for example, the results framework), has changed the way the Bank interacts with the client around the CAS.

- *The results-based CAS is a novelty and stimulates dialogue; it was the beginning of being more explicit about what we are trying to do. It feels like the blind leading the blind, but it's getting better with each year. (CD/CC)*
- *Results management means structuring the work in a way that it is focused on achieving results. For the CAS, we first decide what the results are that we want to achieve, and then we structure the programs to achieve these results, and then build the right teams to work on these programs, to achieve the results. (CD/CC)*

Challenges, Incentives, and Disincentives in Using M&E Information for Results

All respondents repeatedly underscored the importance of conducting responsible M&E to document Bank work and to manage for results.

Professionalism, an “evidence-based” work ethic, Bank system requirements for data generation, and motivation from management drive the use of M&E information internally.

- *To use M&E data and information generally is good. Professionals see the benefit for that, and you don't need particular incentives. (CD/CC)*
- *In the public health field we value it very much. “It's indigenous, in our blood.” We cannot work without knowing what's happening and what's going on. (TTL)*
- *We have many groups looking over us to see if we are doing M&E; it's a prominent part of the ISR. (TTL)*

But respondents described organizational pressures that make M&E of results problematic.

Factors such as TTL overload, time pressures, and lack of resources for monitoring tend to produce monitoring plans and indicators that are put together quickly, just before the PAD goes to the Board.

- *Time is especially important in monitoring, because it takes a long time to reach agreement with the client on what is worth both-ering to monitor and how on earth they are going to do that. You can't do that within six months to prepare a project. (TTL)*
- *You need to do it quickly, you need to do it before, and then you are running out of time, because there are all the pressures to deliver. (CD/CC)*
- *Real action starts during implementation. That is where the challenge comes in, regarding the amounts of resources that actually have been allocated. The rationing of resources begins to happen here. M&E is an expensive thing, and so the M&E guy gets bumped off the train sometimes and gets pushed back. (TTL)*
- *In some respects basing your program on M&E is harder. It's easier to just implement something rather than to think about whether it is making a difference or having an impact. The disincentive is extra effort, extra time, and extra money. (TTL)*

Respondents experienced conceptual difficulties in developing results chains and identifying indicators.

They pinpointed difficulties positioning a project within the chain. That is, they come up with overly ambitious project objectives or position the objectives too high up the results chain and then struggle to develop results frameworks around them. They also had difficulties identifying appropriate performance indicators and were looking for support.

- *I prepared a project and I wanted to stop at a level of improving accessibility. The intervention was to rehabilitate some hospitals and improve the quality of services. The peer reviewers said that this was not an impact.... What is the impact on people? It would take more than the project life to show impacts. One of them said, "What about maternal mortality?" (TTL)*
- *One of the main causes of this is how ambitious development projects have become. That is where the whole thing begins. Everybody wants to have this best project that will change*

the entire face of the world—and not tone it down, make it more humble, more realistic. Once you are unrealistic in project conception [overambitious], everything that follows from it is also to be overstated—indicators, etc. That is where we should bring in more humbleness and modesty—because that is where the whole problem begins. (TTL)

Staff (and clients) feel burdened by the heavy emphasis on quantitative data.

Respondents in all groups agreed that not all results that matter for assessing the performance of the project can be measured. TTLs tend to be action-oriented people; some expressed frustration with the bookkeeping and shepherding of data required for monitoring, often at the expense, they believe, of operations.

- *There is a tyranny of the measurable here that we have to be very careful of. (SM)*
- *Some things are not easy to measure, and we try to force things into a measurement box that isn't always appropriate, so it becomes the "feeding the beast" syndrome—you are doing it just because the institution demands it. (SM)*
- *We try to use specific numerical targets for projects, but that doesn't always happen. In many projects, it tends to be more subjective, even though they should be more precise; not all things are quantifiable. (SM)*

Client buy-in to monitoring varies greatly between countries with reasonable to high M&E capacity and countries with low M&E capacity.

Governments often do not recognize the value of developing monitoring (and evaluation) capacity and are therefore unwilling to use borrowed funds to develop that capacity. Respondents pointed to three main factors that can reduce the usefulness of monitoring data for clients: (a) divergent data requirements of the Bank and the client; (b) a less-than-optimal process of designing the M&E system; and (c) uncertainty on the part of the clients regarding possible negative consequences of less-than-positive monitoring results. Lack of client buy-in to gathering data threatens effective monitor-

ing, which makes building client capacity, especially for monitoring, crucial.

- *We are asking them to collect something that will be used by us rather than by them. A lot of indicators I used for them as well. But they may not see it. (TTL)*
- *All clients want is to get the job done. They should do M&E but it is not their priority because they know that their job ends with the successful completion of the product, and then they want to go on to the next project. They do it because you ask them to do it. (SM)*

Creating demand for M&E and building client capacity are critical.

The most important thing to leave in the country is an effective monitoring system, network, or method of doing M&E.

- *When you have a sensible set of indicators, it is very useful for the client. Also, in terms of culture, we are forced to look into the results and make sure we don't get caught in implementation and forget about the end. In many client countries, you get bogged down in implementation and to force clients to be partners in tracking the results is useful. (TTL)*
- *I have managed to convince clients to monitor, but it is only because I represent the Bank, and because I am tenacious about it, that they finally give in and start monitoring these things. Some of them then find it useful. I have one client who became creative about coming up with new indicators and using them to convince the prime minister to back her on certain things. (TTL)*
- *There are other kinds of data that are simply difficult to do. If you want to do user satisfaction, you have to have a survey. That has to be programmed into the project, which means that it is not sustainable beyond the end of the project. (TTL)*

Use and Usefulness of IEG Reports

All respondents understood IEG's methods and appreciated its evaluation results, but their position in the organization influences the extent to which they find IEG evaluations useful

(the higher the level, the more useful the reports).

CODE members found evaluations and reviews to be important and used them in commenting on sector strategies and specific projects and in giving feedback to their country authorities regarding development policies. CDs and CCs appreciated the longer view represented in the CAE and saw IEG reviews as important in managing Bank country operations: "How well are our projects doing in achieving what they set out to achieve?" SMs, for the most part, found IEG evaluation products useful, particularly cross-sectoral studies. TTLs reported having little time to read IEG reports other than ICR Reviews and reports that appear to be directly related to their work.

- *IEG is struggling to make sure that their reports are more than reports, that is, not just a piece of paper. This is not an easy undertaking. Most evaluations far exceed our expectations. Every time, we learn a lot. All the evaluations are giving very good food for thought for management. (CODE)*
- *It is a problem for someone to produce a good evaluation. The team needs to spend lots of time to understand the circumstances—all sides, not just one side. When these reports are done in a superficial way, they don't really help. When they are done well and thoroughly, they are very helpful. (CD/CC)*
- *I cite IEG all the time because it is the best we have got. I do not like the attacks on the Bank from outsiders who do not have a nuanced view. (SM)*
- *By nature of their work, IEG ends up coming up with either irrelevant findings [for TTL] or misinterpretations—mainly because their audience of their evaluation is not TTLs. They're producing evaluation results for the Bank, to talk about what's happening in this sector, what's happening in this country, etc. (TTL)*

All respondent types expressed fairly high levels of confidence in evaluation results and analysis and thought that the vast majority of IEG reports are of good quality.

They perceive IEG's independence as central to generating confidence in evaluation results.

Involving evaluators with (sector/country) expertise was seen as a key determining factor for quality. Across groups, respondents defined a useful evaluation as one that does the following:

1. Provides systematic evidence on impacts rather than simply reporting on outcome data
2. Adequately reflects all the factors that presented roadblocks to achieving results
3. Highlights what works and what does not work and why
4. Gives practical recommendations for how to improve in the future.

- *As I am representing recipient countries, [IEG reports are] probably more important for me than to my colleagues. If I have to discuss with one of my governments, I know that the evaluation was made by an independent body and so the information I have is good information. I can say to my authorities on this project, an independent body has found this. (CODE)*
- *Often you find an evaluation report that doesn't tell you why a certain result was achieved. It's one thing to evaluate at one level and it's another thing to tie it back down to the business we are in. We need to have those reports that focus on our sector and then tell us why results were or were not achieved. Those are good reports for us to adjust our business. (SM)*
- *An evaluation of the open water fisheries component looked fantastic on fish output, but when IEG did the analysis of who was benefiting from the project, it was a case of elite capture. For inputs and outputs, it was successful, but not in outcomes; the project had to be redesigned. (SM)*

Respondents found some IEG evaluations too academic or retrospective to be useful.

CODE members, Bank managers, and staff are tackling current development issues. Poor evaluations lacked many of the qualities of useful evaluations cited above.

- *Recommendations may be very idealistically written. In those cases, management might*

want to fully take into account all this—and in normal cases, they do. But in terms of implementation, it may be very costly to apply all these recommendations. One has to strike a balance. In development you have to take into account the ability for implementation. (CODE)

- *It is very easy to say, okay, here are the results for outcomes that we would like, and three years later do an assessment where we realize we have not met any of them and the evaluation report essentially says “unsuccessful” and often does not put this into the broader context of what was achieved or the underlying context. (CD/CC)*

Respondents prefer shorter, timelier, more focused, and more accessible reports.

All respondent types would like shorter reports, briefings, and e-mails announcing reports that apply specifically to their project type, sector, and/or country (or on “hot” topics). CODE members would appreciate more focused, more current, and shorter reports that have sharper and timelier recommendations. Respondents need lessons that culminate in forward-looking recommendations. Although IEG has moved toward publishing more of its products in other languages, CODE and country directors/coordinators suggested translating even more documents into more languages and routinely disseminating them to counterparts in client countries. Results that do not reach clients and beneficiaries probably do not have as broad an impact, they noted.

- *The major problem for the Board is information overload, being flooded by e-mails and paper. We have to condense and crystallize how we present information. (CODE)*
- *Maybe IEG can make some choices: maybe [focus] just on the relationship with private structure... for other issues, deepening may not be necessary. There is some room for improvement here. I would like to see some evaluation reports that are less broad in coverage, but that go into depth on selected issues only—with a rationale behind how IEG chose the strategic issues. (CODE)*

- *IEG should go away from the approach that we need to evaluate everything. IEG needs to have the courage to evaluate areas where we have a clear indication that our performance is not so good... forget the areas where performance is satisfactory but not necessarily spectacular. This is not easy. The DG has to answer this in the end, as he is responsible. (CODE)*

Attachment: The Nature and Purpose of Focus Groups

Focus groups constitute a form of scientific social, policy, and public opinion research. As structured group interviews that proceed according to a careful research design and the principles of group dynamics, focus groups should be distinguished from discussion groups, problem-solving groups, buzz groups, or brainstorming groups. They are not designed to help a group reach consensus or to make decisions, but rather to elicit the full range of ideas, attitudes, experiences, and opinions held by a selected sample of respondents on a defined topic.

Through focused interaction on questions of interest to the client, respondents from target groups can provide a wealth of qualitative data that cannot be gained from surveys alone. Participants are chosen because of background characteristics of special interest to the client (in this case, Bank staff and management) and are given the opportunity in a guided interaction to discuss and debate issues surrounding a program, policy, service, plan, or product. Focus groups normally range from one to two hours in duration. The moderator can pursue ideas that are generated during the discussion. Motivations, feelings, and values behind reactions to products can be elicited through probing, restating questions, and eliciting opinion from others in the group.

Advantages of the focus group are that the client is brought closer to the target groups through observation of the session, listening to tapes, or reading transcripts and/or final reports (in this case, because of the in-house nature of the research and guarantees of respondent anonymity, clients did not observe, listen to tapes, or read transcripts). Participants stimulate each other in an exchange of ideas that may not emerge in individual interviews or surveys; ideas can be linked to areas of particular interest to the client for in-depth exploration.

Through focus groups, it can be learned what characteristics are most salient to participants, the level and nature of emotional value associated with those characteristics, and how participants differ on key issues. Focus group interviews are useful for identifying how target groups think and feel about the topic under discussion. Focus groups also work well in eliciting solutions to problems that participants describe; how systems do or do not work; and their own levels of satisfaction with organizational requirements, policies, and culture.

Especially for in-house focus groups, there is a possibility of bias when potential respondents who have what they perceive as negative organizational experience with the client might be over-represented among those who choose to respond to the invitation. Moderators are trained to ensure balanced participation, regardless of a respondent's point of view. Because focus groups are especially helpful in drawing out the full range of opinions and ideas, the presence of participants with more negative views is useful. A more serious bias would occur if respondents with only positive views (or only negative views) of the client were to attend.

The complexity of insights generated by focus groups extends far beyond the numbers involved or the cost of conducting them. Because of the small numbers involved, however, focus groups should typically not be used to generate quantitative data, nor are participants statistically representative of the target populations from which they are drawn. However, if the research design is well crafted and recruitment is carried out in a neutrally manner from a stratified set of respondent categories, then respondents should be conceptually reflective of their populations.

APPENDIX E: CLIENT SURVEY RESULTS

Since 1999, IEG has surveyed its clients annually to gauge their perceptions of the quality and impact of IEG evaluations as part of the continuous process of self-evaluation. The survey assesses IEG evaluations in four areas: readership and awareness; quality; influence of the report on Bank staff understanding, processes, and policies; and recommendations made to IEG for improving the quality and relevance of its reports. This year's survey also included questions on the use of IEG evaluations in Bank operations and sources of knowledge in the Bank.

This year, IEG conducted three separate surveys: internal Bank staff; executive directors (EDs) and their staff; and selected external audiences. This appendix reports on the findings from these three surveys.

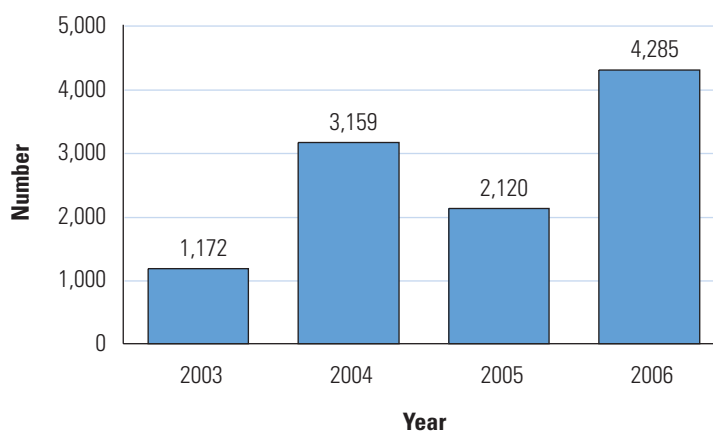
Results of Survey of Internal Clients and Executive Directors

This year's internal survey of 4,285 Bank staff and EDs and their advisors was the largest that has been done in this series (figure E.1). It sought to elicit the perceptions of respondents regarding IEG evaluations produced during 2005.

The following IEG products were the subject of the survey:

- CAEs for Albania, Bolivia, Mauritania, Moldova, Pacific Islands, Romania, and Turkey
- The 2004 ARDE
- The 2005 AROE
- PPARs (37)
- Sector, thematic, and corporate evaluation studies:

Figure E.1: Number of Staff Surveyed 2003–06



Source: IEG data.

- *Capacity Building in Africa: An OED Evaluation of World Bank Support*
- *Committing to Results: Improving the Effectiveness of HIV/AIDS Assistance*
- *Pension Reforms and the Development of Pension Systems: An Evaluation of World Bank Assistance*
- *The Effectiveness of World Bank Support for Community-Based and -Driven Development*
- *IEG Review of World Bank Assistance for Financial Sector Reform*
- *Improving the World Bank's Development Effectiveness: What Does Evaluation Show?*
- *Maintaining Momentum to 2015: An Impact Evaluation of Interventions to Improve Maternal and Child Health and Nutrition in Bangladesh*

Two primary audiences were targeted for this year's survey: Bank staff who could be defined as intended audience for a particular evaluation

based on their role or affiliation (4,095 surveyed) and EDs and their advisors (190 surveyed).

This appendix provides a summary of key findings. Where possible, this year's results are compared with results from previous years and similar types of respondents.

Response rates

The surveys were sent electronically to 4,285 staff over a four-week period in April 2006. The response rate averaged 22 percent over all evaluations (933 responses in total; figure E.2). For individual reports response rates ranged from 13 percent (for *Improving the World Bank's Development Effectiveness—What Does Evaluation Show?*) to 40 percent (for the Bangladesh impact evaluation) and averaged 22 percent across all survey products. As in last year's survey, reminder notifications were sent to nonrespondents. While survey response rates are directly comparable only to 2005,¹ response rates and sample sizes over the past four years show a correlation between larger samples and smaller response rates.

Given the response rate of 22 percent it has to be noted that the survey results are indicative

for respondents but are not generalizable for the overall surveyed population.

Readership and awareness

Overall awareness of reports was 56 percent among non-ED respondents (figure E.3). This includes respondents who had read, participated in a discussion, or attended an event on the evaluation for which they were surveyed (direct awareness) and respondents who were aware of the evaluation's findings even if they had not read the report or participated in an event (indirect awareness). Forty-two percent of respondents had read, participated in a discussion, or attended an event on the evaluation for which they were surveyed. This continues last year's trend of 45 percent and 13 percent, respectively, and marks an increase from 27 percent and 9 percent in 2004.

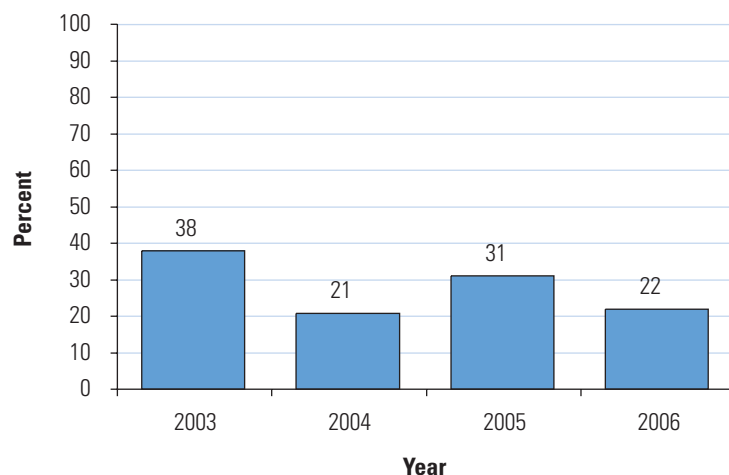
For non-ED respondents who were aware of the evaluation, about two-thirds (64 percent) heard about it through an e-mail announcement, 12 percent saw the evaluation Web site, and 20 percent received a hard copy. Seventy percent said they had shared information or findings from the evaluation with colleagues, 26 percent with colleagues and clients.

Quality of IEG evaluations

Survey recipients were asked to rate their satisfaction with the specific IEG evaluation they were queried about on a six-point scale for 10 attributes of quality: relevance to their work, ease of understanding, concise presentation of conclusions, timeliness, usefulness of recommendations, unbiased and objective analysis, transparency and clarity of the methodology, strength of the link between conclusions and evidence, depth of analysis, and whether all available information was incorporated. Figure E.4 shows the proportion of respondents that rated each aspect of the evaluations 4 or better.

This reflects a new approach to obtaining user views and differs from the method used in previous years, when timeliness, transparency of methodology, and incorporation of all available information were yes/no questions and other

Figure E.2: Response Rates



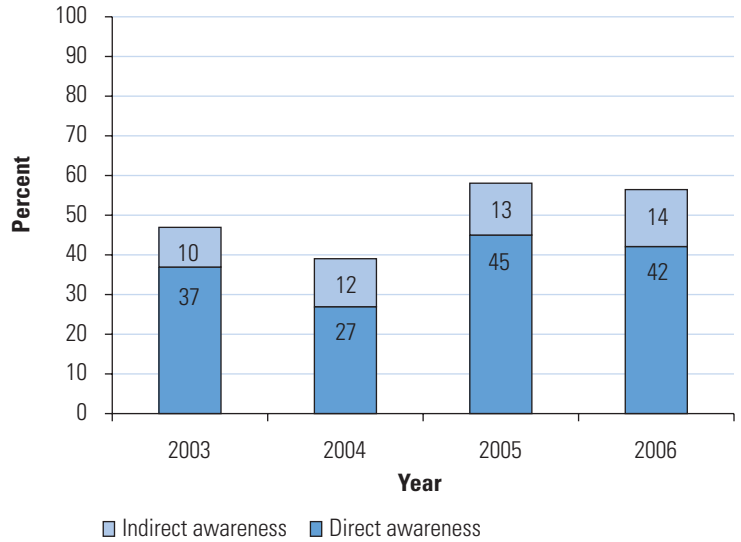
Source: IEG data.

attributes were to be rated on a four-point scale.² Due to this change in methodology, the survey data better reflect audience satisfaction and enable more meaningful comparisons of IEG’s client audiences’ satisfaction with attributes of evaluation quality. However, the change in the rating scale introduces some discontinuity in trends over time.

This year Bank staff survey respondents (292 Bank staff responded to this question) gave the highest ratings to the relevance of IEG evaluations to their work (average of 4.28, with 76 percent rating it 4 or better), closely followed by ease of understanding (4.21, or 76 percent) and concise presentation of conclusions (4.13, or 72 percent). Staff satisfaction was lowest for incorporation of all available information (3.55, or 57 percent) and depth of analysis (3.58, or 56 percent).

Overall, the EDs and advisors who responded (50) expressed higher satisfaction with evaluations than Bank staff. EDs and advisors rated relevance of IEG evaluations to their work highest (mean of 4.9, or 88 percent of respondents rating it 4 or higher), followed by ease of understanding (4.8, or 90 percent) and

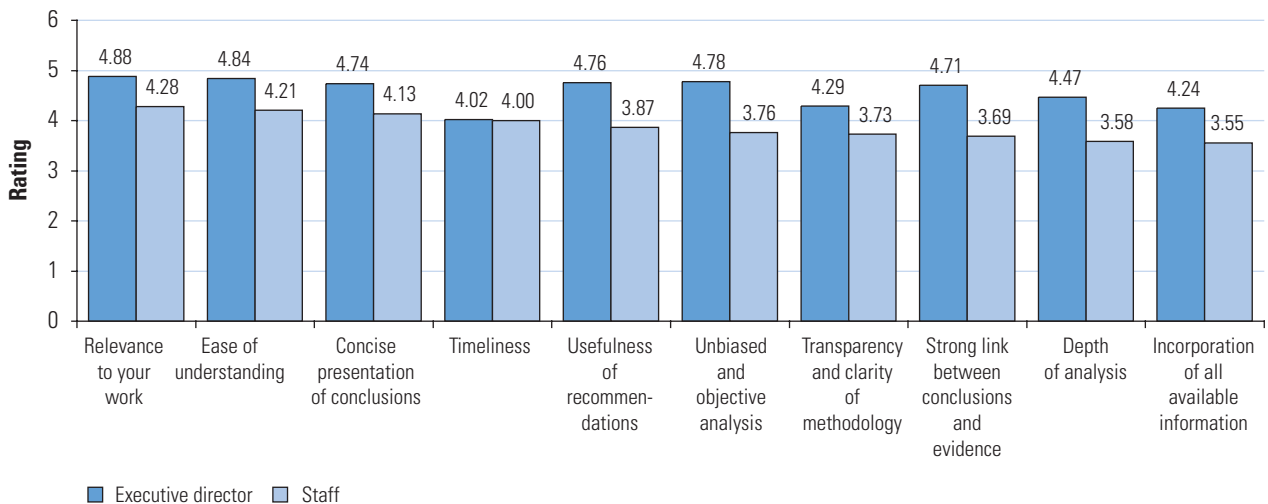
Figure E.3: Readership and Awareness of IEG Evaluations



Source: IEG data.

unbiased and objective analysis (4.8, or 90 percent). They were least satisfied with incorporation of all available information (4.2, or 82 percent) and timeliness (4.0, or 69 percent).

Figure E.4: Bank Staff and Executive Director Satisfaction with Quality of IEG Evaluations



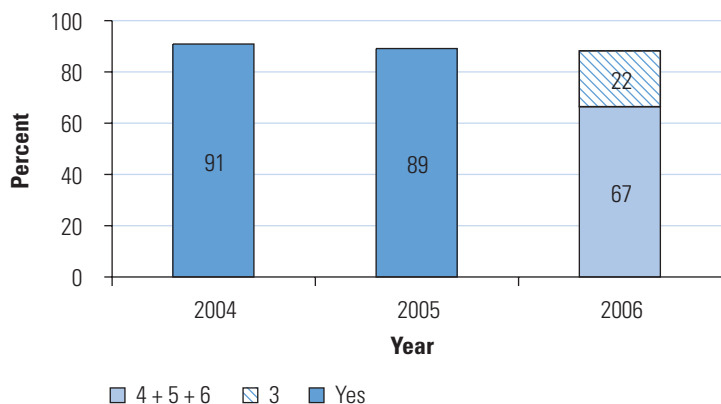
Source: IEG data.

Note: Mean on a six-point scale, N = 292 for staff and N = 50 for executive directors.

As noted, although these results cannot be directly compared with previous years, the 2006 survey shows timeliness ratings of IEG evaluations that are comparable with previous years (figure E.5). 89% of Bank staff rated their satisfaction 3 or higher, and 67 percent rated it 4 or higher, compared with 89 percent who considered evaluations to be timely last year. The responses are similar for EDs and advisors, of whom 89 percent rated their satisfaction 3 or higher and 69 percent rated it 4 or higher, compared with 80 percent who considered IEG’s evaluations to be timely last year.

Comparing those attributes that used a four-point scale in previous years shows that relevance of IEG evaluations to Bank staff’s work was rated highest this year, with 76 percent rating it 4 or higher, compared with 67 percent rating it excellent or good last year and 73 percent in 2004. Link between conclusions and evidence (2006, 61 percent; 2005, 58 percent; 2004, 65 percent) and concise presentation of conclusions (2006, 72 percent; 2005, 73 percent; 2004, 75 percent) fare similarly well. Satisfaction with depth of analysis, however, was rated 56 percent this year, compared with 71 percent in 2005 and 80 percent in 2004 (figure E.6).

Figure E.5: Bank Staff’s Satisfaction with Timeliness



Source: IEG data.

Influence of IEG evaluations

This year, Bank staff were asked to rate on a six-point scale—from 1 = not at all to 6 = a great deal—the influence of an IEG evaluation on their understanding of the subject area, the Bank’s country assistance and sector strategy, lending and nonlending services, and how outcomes are linked to outputs (results framework) at the project, sector, and country levels. Moreover, they were asked to rate how much the evaluation helped them gain a view of essential lessons and best practices. This year 268 Bank staff responded to this question.

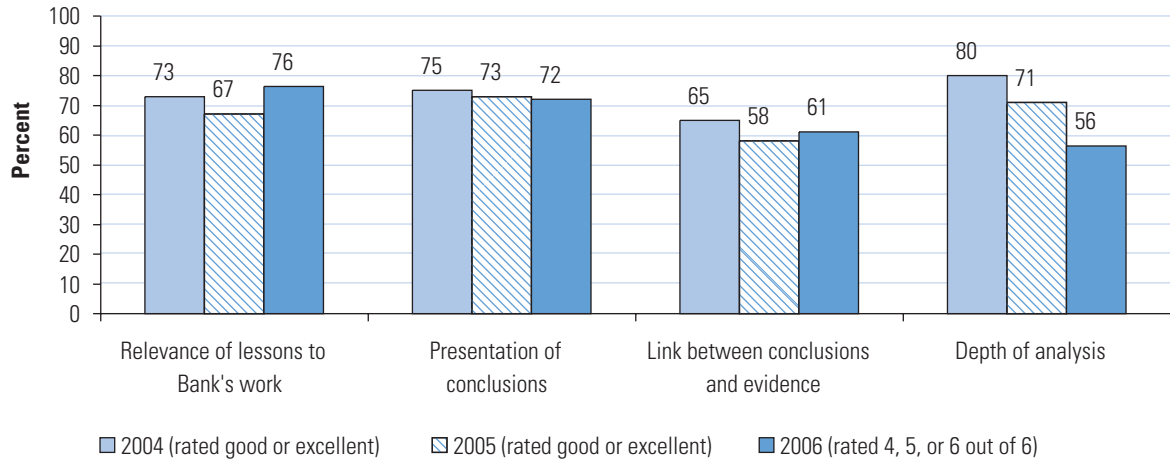
EDs and advisors were asked to use the same six-point scale to rate how much IEG evaluations influenced their understanding of the subject area; helped them better understand Bank activities at country, sector, and project levels; and helped them gain a view of good practices and essential lessons. Forty-nine EDs and their advisors responded to this question.

This year more than three-quarters (79 percent) of Bank staff respondents rated the influence of evaluations on their understanding of the subject area 3 or higher, and 28 percent rated it 5 or 6 (figure E.7). This result is similar to previous years, when a four-point scale of not at all, little, somewhat, and a great deal was used. In 2005, 78 percent rated IEG evaluations as having either somewhat or a great deal of influence on their view of the subject area. The rating in 2004 was 73 percent.

Overall, as in previous years, EDs and advisors perceive themselves to be strongly influenced by IEG products. Among EDs and advisors, 90 percent of respondents rated the influence of IEG evaluations on their understanding of a subject area 4 or higher, with 69 percent rating it 5 or 6. Last year, 95 percent of respondents rated IEG influence 3 or 4 out of 4. IEG was considered helpful to gaining a clear view of lessons by 94 percent (mean of 4.7) and to gaining a clear view of good practices by 86 percent (mean of 4.4).

Among Bank staff respondents, IEG’s recognized strengths at compiling lessons and good

Figure E.6: Quality of Studies over Time



Source: IEG data.

practices do not translate into an equally strong rating for informing staff about the make-up of results chains (figure E.8). Staff respondents ranked IEG evaluations highest on helping them to gain a clear view of lessons (mean of 3.8, 65 percent ranking it 4 or higher) and of good practices (3.51, or 52 percent). Influence on how outcomes are linked to outputs is ranked lowest (3.22, or 44 percent average over project, sector, and country level). Forty-four percent and 46 percent of staff respondents rated IEG's influence on the Bank's sector and country strategies, respectively, at 4 or higher, compared with 47 percent rating it somewhat or a great deal in 2005.

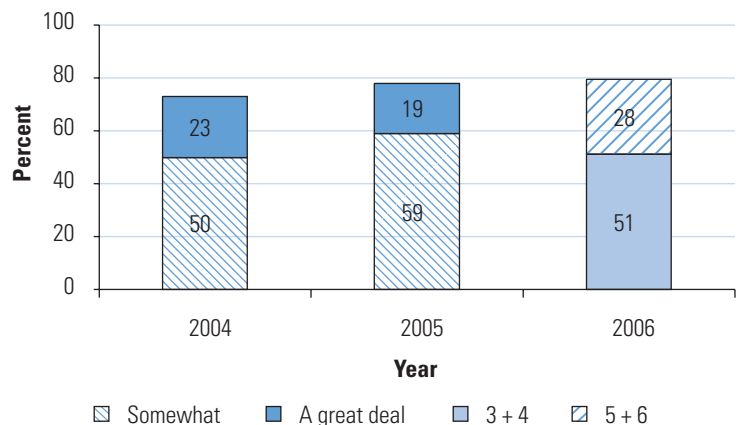
Use of Evaluations

This year, the client survey also probed the use of evaluations (figure E.9). This section was added to obtain a better understanding of evaluation use. EDs and advisors were asked how they use evaluations to assess the Bank's work. Staff were asked about different modalities for use, such as giving advice, commenting, designing new products and strategies, and modifying ongoing operations and strategies. Two hundred forty-one Bank staff and 53 EDs responded to this question. Overall, EDs and advisors use evaluations primarily to fulfill their

oversight function. Bank staff report using evaluations mostly for commenting, making a case, and providing advice, less for modifying ongoing operations or designing new Bank products.

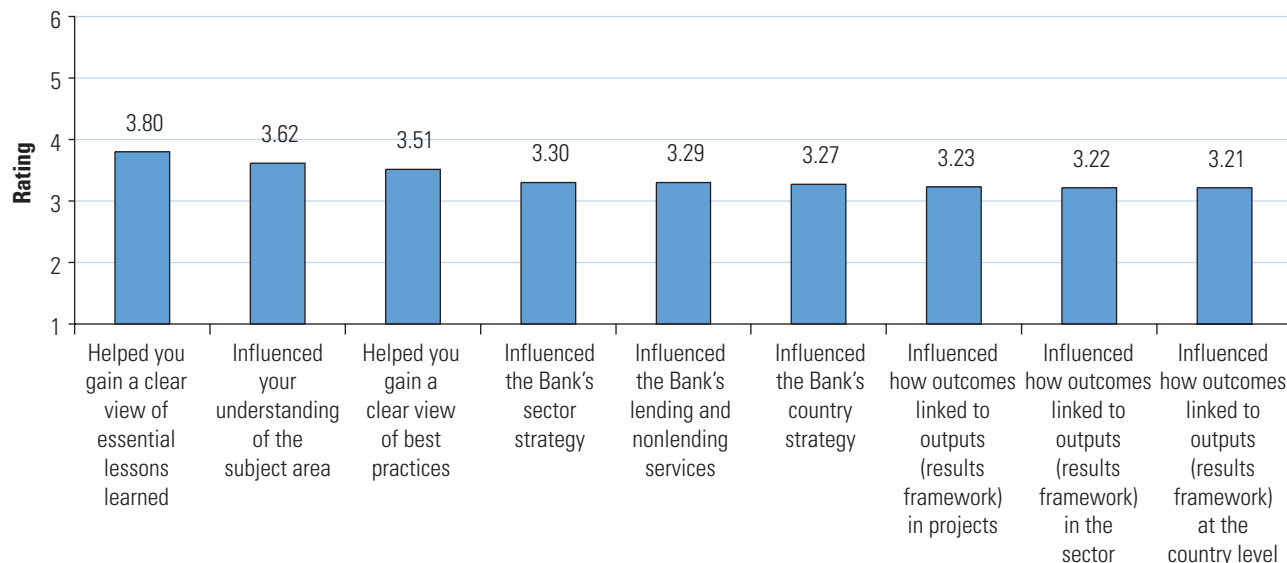
Bank staff use evaluations mostly for commenting on the work of others, making a case for a particular course of action, and providing advice to clients, and less for modifying strategies or

Figure E.7: IEG's Influence on Bank Staff's Understanding of Subject Area



Source: IEG data.

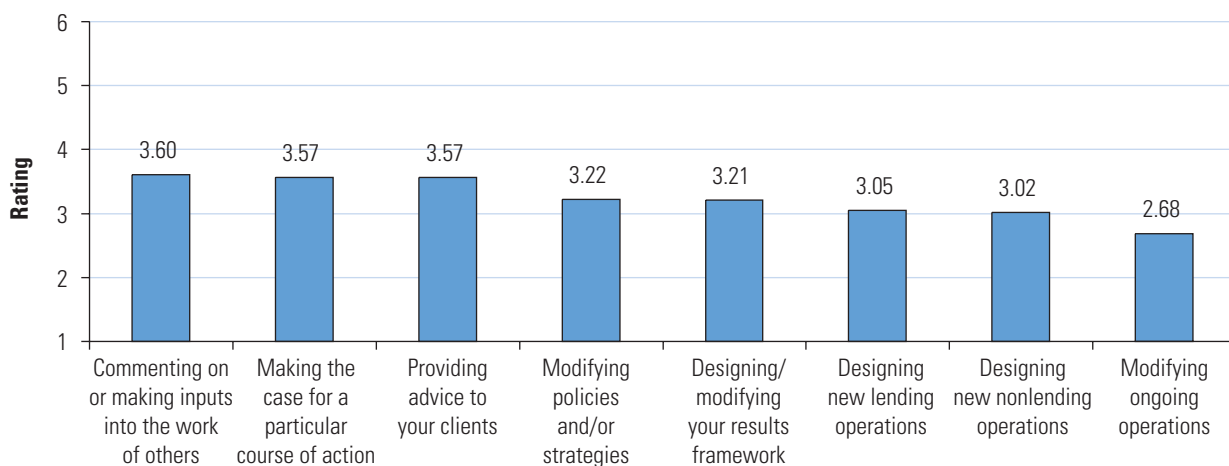
Figure E.8: Bank Staff's Rating of IEG's Influence



Source: IEG data.

Note: Mean on a six-point scale, N = 268.

Figure E.9: Bank Staff's Use of IEG Evaluations



Source: IEG data.

Note: Mean on a six-point scale, N = 241.

operations, or designing new projects or programs. About one in four (27 percent) respondents acknowledged that they use evaluations to modify ongoing operations, assigning a rating of 4 or higher. EDs and advisors use IEG

evaluations heavily to assess Bank policies and procedures; 90 percent of respondents rate it 4 or higher (mean of 4.7), followed by use to assess country strategies (mean of 4.8, 88 percent rating it 4 or higher) and sector strategies (4.6,

or 81 percent). EDs and advisors use evaluations to a lesser degree when assessing projects (3.7, or 57 percent).

Recommendations from respondents

When asked how to improve the quality of evaluations, Bank staff respondents selected “make findings more operational” as their top recommendation. However, this year just 57 percent of respondents chose this recommendation, compared with 72 percent last year (figure E.10). Nearly half of non-ED respondents also suggested that obtaining more evidence and broadening consultation with Bank staff could help improve evaluation quality.

EDs and advisors were asked to select from the same set of recommendations for IEG evaluations. Sixty-four percent of respondents said they would make the findings more operational, down from 87 percent last year. However, more respondents than last year suggested increasing external consultation (2006, 56 percent; 2005, 46 percent) to obtain more evidence (2006, 38 percent; 2005, 29 percent) and to broaden consultation with Bank staff (2006, 38 percent;

2005, 21 percent). Thus, EDs and Bank staff have a similar view on where and how IEG needs to improve.

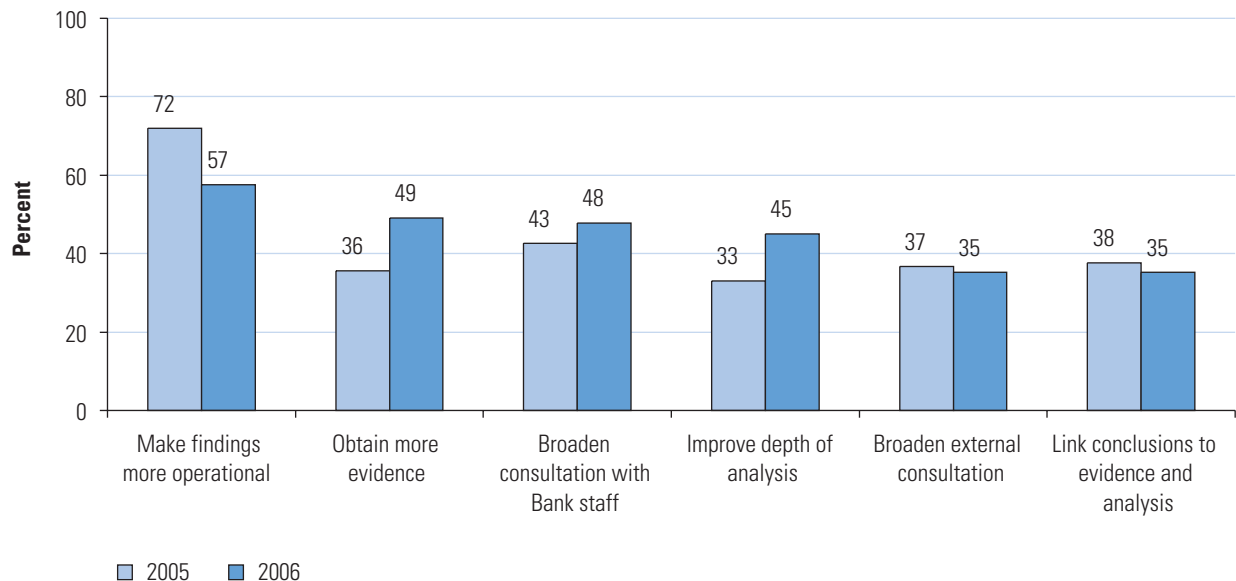
Both EDs and Bank staff were asked for their views on various recommendations for improving IEG’s outreach (figure E.11). Respondents generally called for more summaries of IEG findings and suggested better notification about new documents and events. In addition, Bank staff would like to have more access to evaluation data, while EDs and their staff want IEG content to be more accessible online.

Results of Survey of External Clients

This year’s external survey queried 2,759 external clients from governments, international organizations, bilateral donor organizations, nongovernmental organizations (NGOs), academia, and the general public. They were asked about the following:

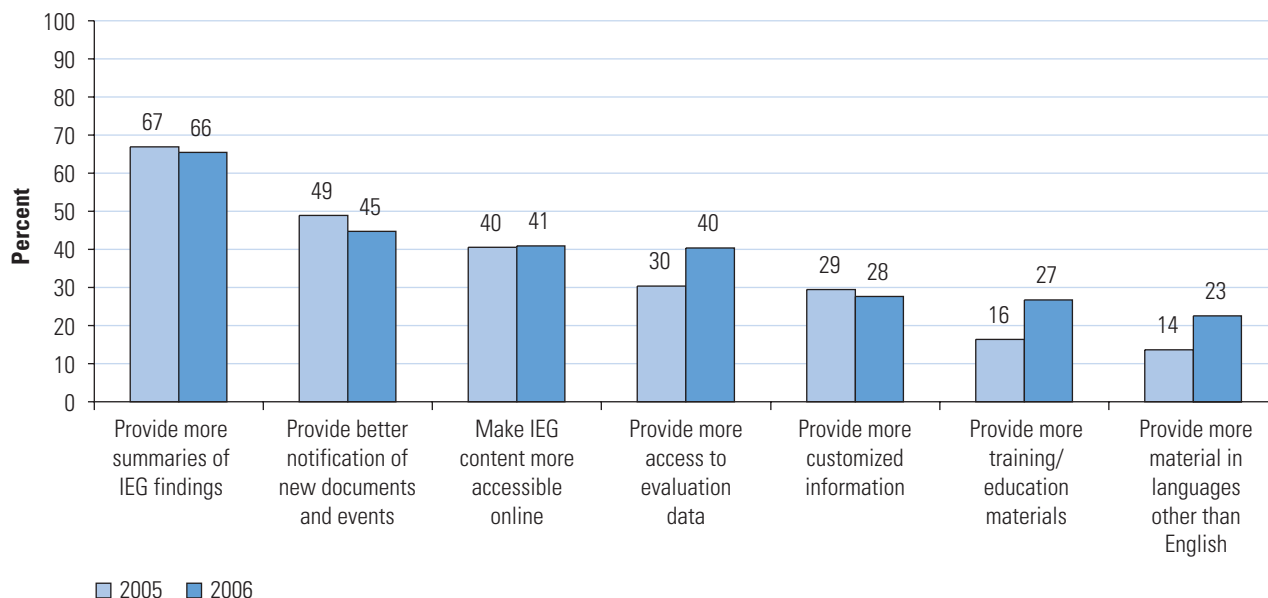
- Two corporate evaluations
 - The 2004 ARDE
 - *Capacity Building in Africa: An OED Evaluation of World Bank Support*

Figure E.10: Bank Staff Recommendations to Improve IEG’s Evaluation Quality



Source: IEG data.

Figure E.11: Bank Staff Recommendations to Improve IEG’s Outreach



Source: IEG data.

- Three sector and thematic evaluations:
 - *Committing to Results: Improving the Effectiveness of HIV/AIDS Assistance*
 - *Pension Reforms and the Development of Pension Systems: An Evaluation of World Bank Assistance*
 - *The Effectiveness of World Bank Support for Community-Based and -Driven Development (CBD-CDD)*
- One synthesis paper: *Improving the World Bank’s Development Effectiveness: What Does Evaluation Show?*

Surveys were sent to individuals who had either received an IEG report through an e-mail announcement or as a hard copy, had participated in a seminar, had been consulted during the evaluation, or had placed a request concerning a report with the IEG Help Desk. For the CBD-CDD study, Pension Reforms, and the synthesis paper, surveys were only sent to those e-mail recipients who had acknowledged receipt of the report by opening the report’s Web page. Thus, as for the internal survey, the sample was targeted, not random. The following is a summary of survey responses.

Response rates

The surveys were sent electronically to the reported e-mail addresses of 2,759 external clients during a two-week period in April 2006. Valid responses were received from 322 people, a response rate of 12 percent. Response rates were highest where surveys were sent exclusively to those who had acknowledged receipt of a report. Multiple reminder notifications that directly targeted nonrespondents were used. As for the Bank staff survey, results from the external survey are not generalizable to the overall surveyed population, but are indicative for the sizeable number of respondents.

Asked to identify the type of their organization, about one quarter (27 percent) of respondents were from international organizations, 15 percent from governments, and 9 percent from governmental donor organizations. Fifteen percent were working with NGOs and 20 percent with academia/research organizations.

Readership and awareness

More than half (55 percent) of all respondents read, participated in a discussion, or attended an

event on the evaluation for which they were surveyed (figure E.12). Overall awareness (including those who were aware of the findings but had not read the evaluation or participated in an event) reached 76 percent on average.

The majority of respondents (69 percent) had heard about the evaluation via e-mail, 28 percent had visited the evaluation Web site, and 15 percent had received a hard copy. About seven percent of respondents stated that they had been consulted during the evaluation.

The majority of respondents had shared information or the findings from the evaluation with colleagues. Fifty percent shared it with two to four colleagues and 18 percent shared it with more than 10 colleagues, expanding the indirect reach of evaluations.

Quality of IEG evaluations

External survey audiences were asked to rate their satisfaction with different attributes of quality using a six-point scale. Respondents rated their satisfaction with the quality of IEG evaluations very favorably (figure E.13). Ease of understanding received the highest rating, with a mean of 4.75 and 89 percent of respondents rating it 4 or higher, followed by concise presentation of conclusions (4.73, or 89 percent). Eighty-five percent of respondents rated their satisfaction with the evaluation’s relevance to their work at 4 or higher (mean = 4.66). Although still high, satisfaction was lowest with depth of analysis (4.32, or 80 percent) and incorporation of all available information (4.15, or 75 percent).

Comparing these ratings with last year’s results, this year’s respondents rate their satisfaction with IEG’s quality more favorably overall.

A strong majority of all respondents rated their satisfaction with each of the aspects positively:

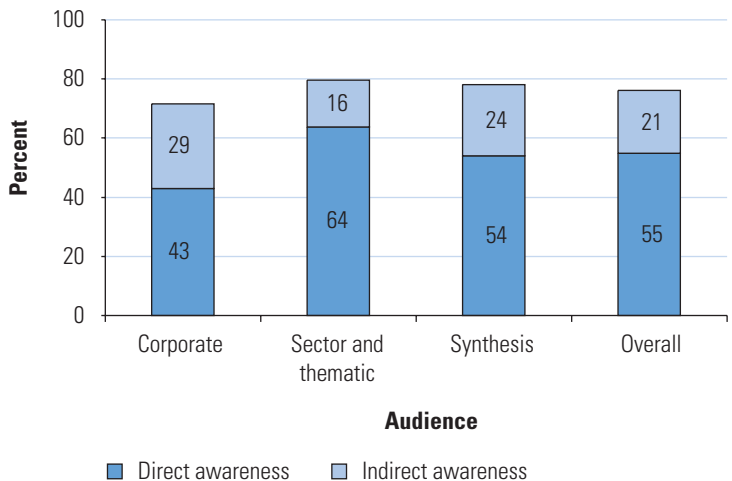
- Ease of understanding (mean of 4.75, 89 percent rated 4 or higher)
- Concise presentation of conclusions (4.73, or 89 percent)

- Relevance to their work (4.66, or 85 percent)
- Usefulness of recommendations (4.51, or 84 percent)
- Transparency and clarity of the methodology (4.49, or 86 percent)
- Timeliness (4.45, or 78 percent)
- Unbiased and objective analysis (4.34, or 78 percent)
- Strong link between conclusions and evidence (4.34, or 79 percent)
- Depth of analysis (4.32, or 80 percent)
- Incorporation of all available information (4.15, or 75 percent).

Influence of IEG Evaluations

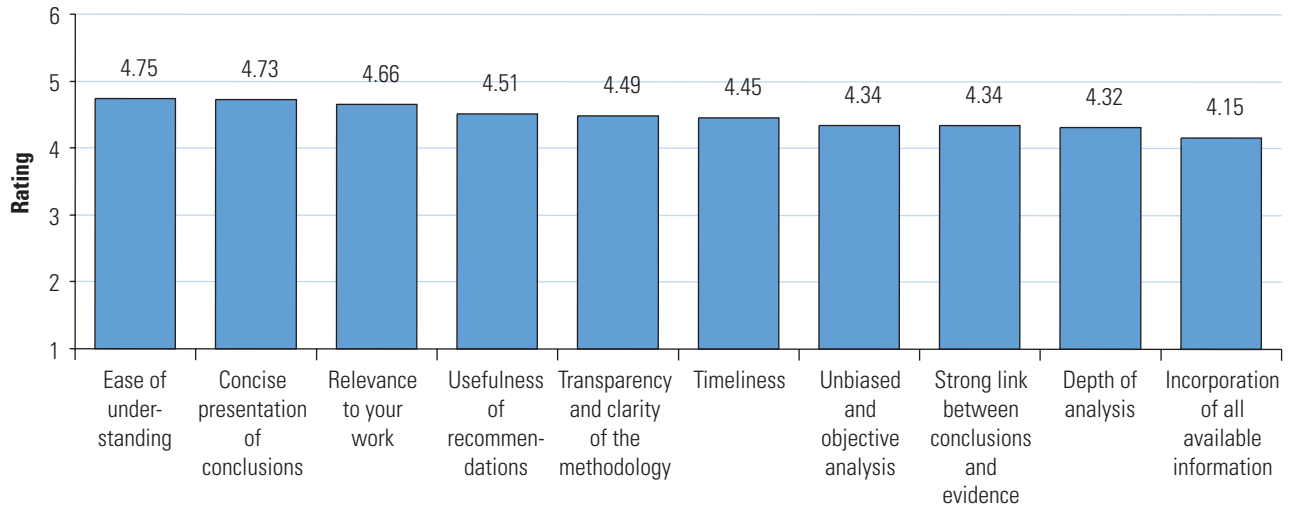
External respondents were asked to rate on a six-point scale the extent to which the evaluation influenced their understanding of the subject area, helped them better understand the Bank’s activities within a sector, and conveyed essential lessons learned, best practices, and other relevant information to help them make decision on strategies and programs (figure E.15). Respondents rated all dimensions highly. Influence on their understanding of the subject area was rated 4 or higher by 81 percent, compared with 68 percent who rated it as good or excellent last year.

Figure E.12: Awareness among External Audiences



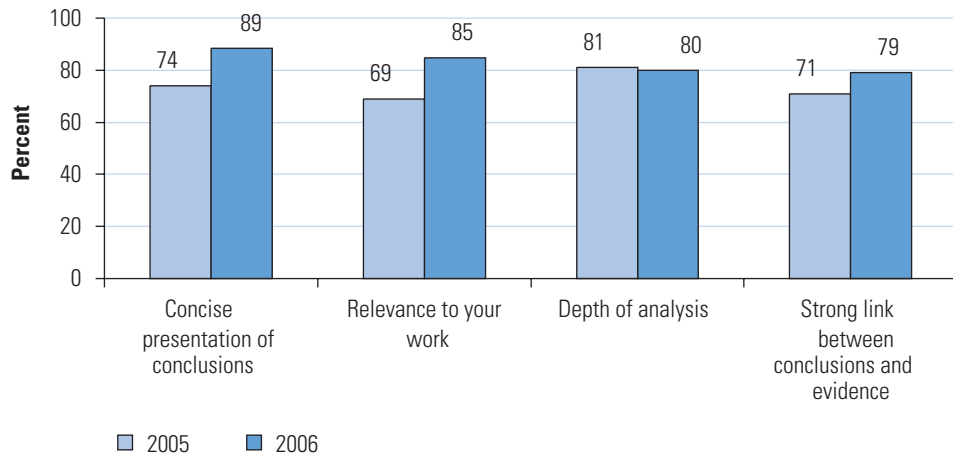
Source: IEG data.

Figure E.13: External Client Satisfaction with Quality of IEG Evaluations



Source: IEG data.
 Note: Mean on a six-point scale. N = 161.

Figure E.14: External Client Satisfaction with Evaluation Quality 2005–06



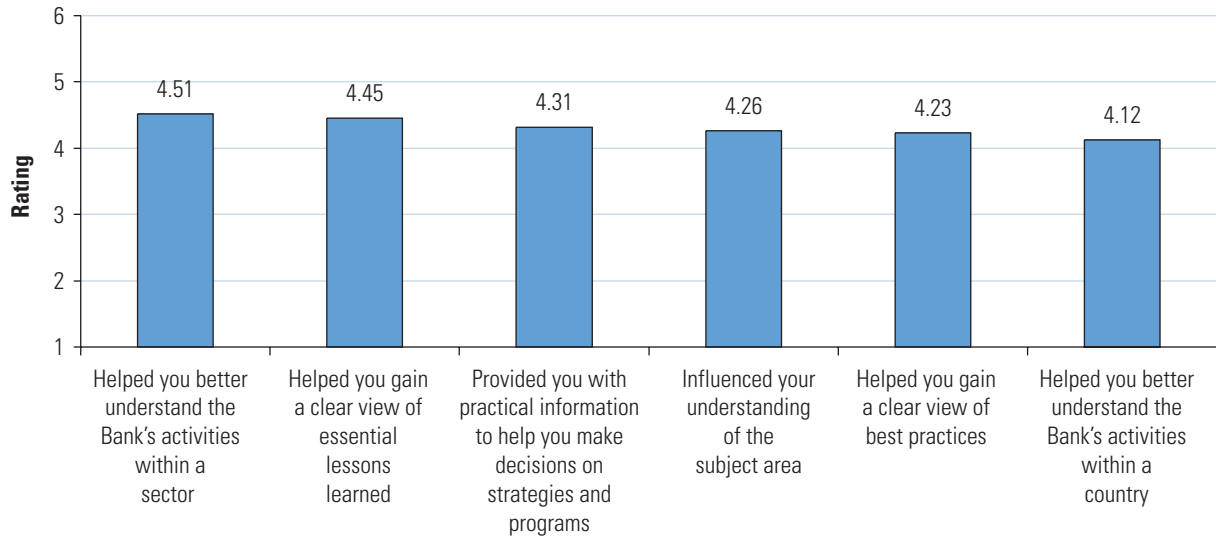
Source: IEG data.

Use

Respondents were asked to rate—using a six-point scale—their use of the evaluation for research, making the case for a particular course of action, refocusing ongoing strategies or programs and education. A majority of respondents assigned the highest “use” ratings from evaluations to research (mean of 4.0, 69 percent

rating it 4 or higher) and advocacy, that is, making the case for a particular course of action (3.95, 70 percent). Sixty-two percent of respondents relied on evaluations to refocus ongoing strategies. Use of evaluations for educational purposes was ranked fourth, with a mean of 3.65 and 57 percent of respondents rating it 4 or higher (figure E.16).

Figure E.15: Influence of IEG Evaluations on External Clients



Source: IEG data.
 Note: Mean on a six-point scale. N = 154.

Recommendations from respondents

When asked about improvements, the most often-selected recommendation was to broaden external consultation (58 percent), up from 31 percent the previous year, followed by making findings more operational (52 percent; 2005, 61 percent) and improving depth of analysis (44 percent; 2005, 50 percent). Thirty-nine percent of respondents recommended that IEG obtain more evidence, up from 32 percent last year (figure E.17).

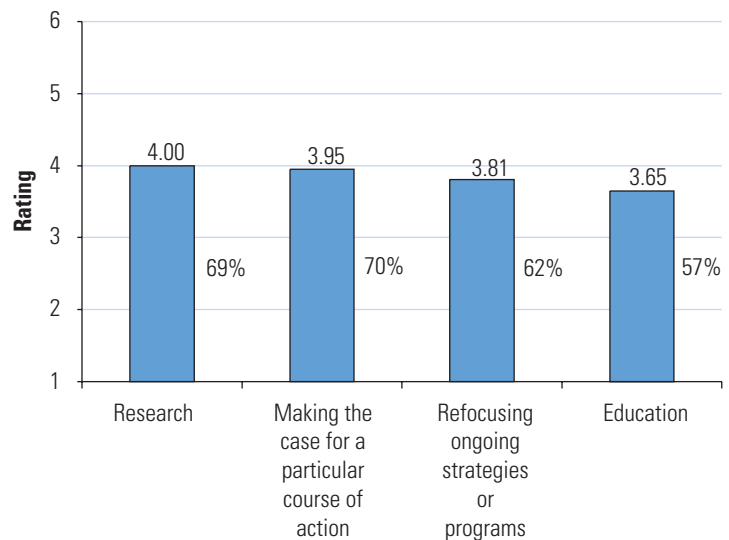
When asked what recommendations they would make to improve IEG's outreach, 58 percent of external survey respondents suggested providing more summaries of IEG findings. About half of the respondents would like to have more access to evaluation data, more training and education materials, better notification of new documents, and more online accessibility (figure E.18). Overall, there is little difference from last year's recommendations.

Implications for IEG

The survey results indicate that IEG provides good services to its main client, the Board.

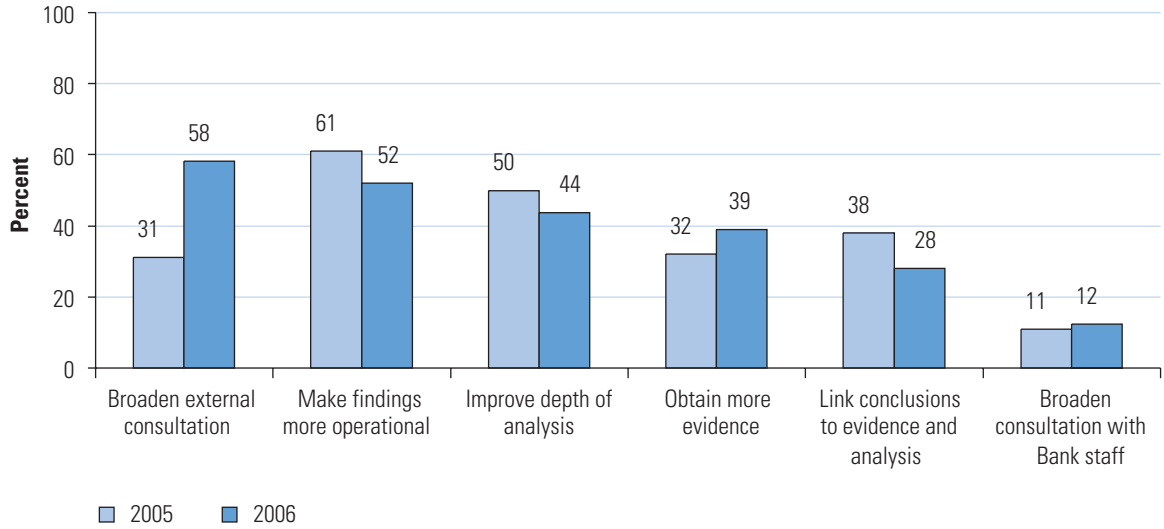
Board members are largely satisfied with the quality of IEG evaluations and use them extensively to assess the Bank's activities at the sector and country level.

Figure E.16: External Client Use of IEG Evaluations



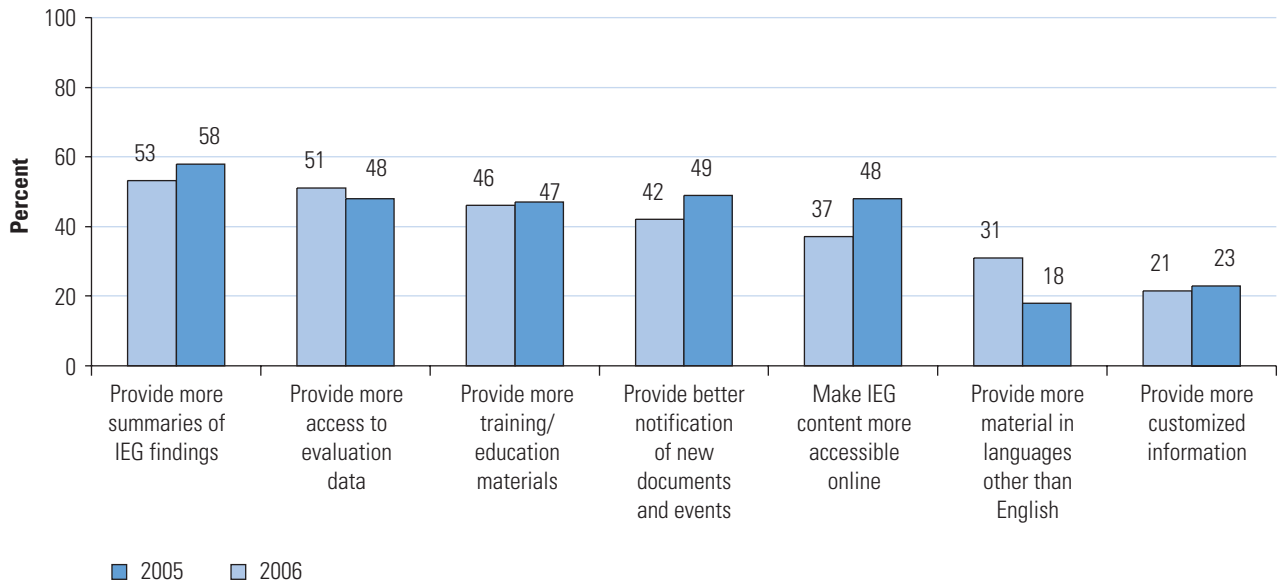
Source: IEG data.
 Note: Percentages show how many respondents rated 4 or higher.

Figure E.17: External Clients' Recommendations to Improve IEG Evaluations



Source: IEG data.

Figure E.18: External Clients' Recommendations to Improve IEG's Outreach



Source: IEG data.

External audiences, which consist mainly of other multilaterals, governments and bilateral donors, NGOs, and academia, were also highly satisfied with the quality of IEG evaluations and considered them influential. Thus, IEG is successful in turning its evaluation knowledge into a valuable global public good. However, external audiences would like IEG to broaden its external consultation during evaluations.

IEG products help deepen staff understanding of a given subject area and lessons, but staff use evaluations less frequently in their day-to-day work of designing and implementing projects and programs. Thus, IEG provides high-level knowledge that is useful for assessing programs, giving advice to clients, and making comments,

but it needs to focus on influencing ongoing and future operations.

Across all audiences, the survey results suggest that improving its timeliness to provide information when decisions need to be made; deepening its analyses through more consultation and attention to the context of evaluated operations; and focusing on the operational value of its findings will enhance IEG's influence on operations and subsequently on development outcomes.

As IEG looks for novel ways to contribute to the Bank's results agenda, one possible entry point may be to help identify results chains and operational knowledge "nuggets" that can be readily applied in ongoing operations.

APPENDIX F: OVERVIEW: CHANGES IN IEG'S DISSEMINATION METHODS

Disclosure of evaluations on the Internet		
Before	Now	Why?
Post electronic document on Web site without formal sign-off process	Systematic disclosure clearance process requiring TTL sign-off (e-mail) copying relevant IEG group managers	(1) To comply with disclosure requirements (2) To inform TTLs and managers of postings, to minimize error in timing or contents of posted items
	Notify OPCS prior to disclosure of major evaluations	To notify OPCS of major outreach campaigns that may draw attention from external audiences
E-mail marketing		
Before	Now	Why?
None	Conduct targeted internal and external audience research and analysis	To ensure core audiences are reached by e-mail and/or hard-copy distribution, including those identified by previous AROEs as neglected, such as country office and operational staff
None	Send e-mail announcements for important sector, thematic, and corporate evaluations to selected internal and external audiences (ranging from 1,000 to >10,000)	To significantly increase awareness within the Bank and externally of IEG's role and its major evaluation findings
None	Include in e-mail announcement a tracking code to identify audience response patterns	To monitor who is reached by, and has taken an interest in, the evaluation. Monitoring mechanisms allow IEG to tap into virtual networks, for example, by identifying individuals who forwarded the announcement to colleagues.
Web promotions		
Before	Now	Why?
None	Google AdWords	To increase visibility of IEG and its products on Google
None	Cross links from key internal and external websites	(1) To increase ranking of IEG and its products on Google search results (2) To provide links to IEG products from relevant pages, such as CAEs on the Bank's country office home pages
None	Section on World Bank's home page dedicated to featuring IEG content	To increase the visibility and reach of IEG content by placing links to new products on the Bank's high-traffic home page

(Continues on the following page.)

Media campaigns

Before	Now	Why?
"Today" stories and kiosk announcements	Increased coverage of IEG events (IEG conferences, Good Practice Awards, launch events)	Announce evaluation products or events
None	Definition of key messages	(1) Identify key messages for internal and external audiences (2) Express messages so they are concise, engaging, relevant to the current policy agenda inside the Bank and outside, consistent across dissemination channels, and grounded in evaluation evidence
None	Media strategy	To maximize impact with targeted journalists in major media organizations covering evaluation topic area
None	Press kit: press release, media advisory, fact sheet	To provide effective briefing materials for accurate reporting and increased coverage
None	Press conference	To generate interest from media to increase coverage
None	Media monitoring following campaign launch	To monitor media coverage of evaluation, to measure impact of campaign

Web site enhancements

Before	Now	Why?
Simple one- or two-page Web sites for products	Content and feature rich Web sites for promoting completed and ongoing evaluations	To make user experience more engaging, to increase repeat visits and referrals
None	Careful attention to Web site messages	(1) To ensure consistency across IEG's message delivery products, such as the press release, executive summary, and so forth (2) To ensure balanced messages that highlight both the positive and negative and are constructive in tone
Basic Web statistics software	Enhanced Web statistics tracking software including campaign tracking	To monitor the impact of outreach campaigns and measure impacts of different outreach instruments
Contact IEG Help Desk	Request page for ordering hard copies and providing feedback	To provide internal and external audiences with a quick mechanism for ordering hard copies

Multilingual outreach

Before	Now	Why?
Translation of print publications, but limited translation of Web content into other languages	Increase in volume of content translated for Web-only publication, or for Web site-only text	(1) To increase the cost effectiveness of IEG's translation budget (2) To reach non-English language audiences through electronic media, as recommended by the Bank's translation framework
None	Partnerships with EXT team to announce IEG content in non-English languages on Web sites and newsletters in relevant languages	To extend dissemination of IEG content electronically in multiple languages to reach non-English-speaking audiences
None	IEG key Web pages in multiple languages	To provide visitors to the Bank's external home page in Spanish, French, and Arabic with IEG Web pages in these languages

IEG help desk

Before	Now	Why?
Ad hoc requests on wide range of topics	Significant increase in volume of requests for hard copies and questions related to launched products	To integrate the Help Desk into promotional campaigns for products

Note: EXT = External Affairs, Communications and United Nations Affairs Vice Presidency.

APPENDIX G: ACCELERATING THE RESULTS AGENDA

In its most recent progress report on the Bank’s results agenda (World Bank 2006b), OPCS laid out a set of program actions to be undertaken over the next two years. The salient features of the action program are summarized in the following table. The AROE found the plan to be

comprehensive and to reflect elements critical to furthering the Bank’s results agenda. However, the plan could benefit from more specificity on how it would be implemented, especially its cost and funding implications. The AROE’s assessments are noted in table G.1.

Table G.1: OPCS Action Plan and AROE Assessments

Report section III—Program of action	AROE assessment
Pillar I: Strengthening country capacity to manage for results	
Country Statistical Capacity Building: Continue to support countries improve their statistical capacity through DGF in line with Marrakech Action Plan.	
Sectoral Concerns and Management Capacity: Enhance government capacity to interpret statistics and use them for policy implementation at the sectoral and/or cross-sectoral levels.	Critical element going forward. “Develop tools for country/sectoral teams to assess the incentive/demand for results based planning” was identified as a Next Step in the Results Framework (annex B in the report). Some Regions (for example, Latin America and the Caribbean) are more advanced in working with countries on M&E support and evaluation capacity development (ECD). The Bank could go beyond analysis, reporting, and tool development to providing funding for direct support. This would give M&E and its use more prominence and send a signal to staff that management considers this important.
Pillar II: Strengthening results in the Bank	
Regional action program:	
Describe how the Region will organize itself to manage for results.	
Provide just-in-time training for front-line staff.	Critical element going forward. (See comments on OPCS’s knowledge and learning strategy, below.)
Develop a process for identifying regional projects/ programs that should receive support for the design and implementation of impact evaluations in coordination with the Development Impact Evaluation (DIME) Initiative.	

(Continues on the following page.)

Table G.1: OPCS Action Plan and AROE Assessments (continued)

Report section III—Program of action	AROE assessment
Review, comment, and advise on the results orientation and quality of all operational products.	Critical element going forward. Regions already strengthening review/support functions in vice presidents' offices.
Review and comment on the preparation of Regional strategy and performance contracts to ensure both clarity of focus on results and consistency with regional and country-level internal budgets and work programs.	
Facilitate sector/country discussions of work program alignment	Critical element going forward. Regions (with support from OPCS) should provide support to CDs and country teams to refine the results framework, identify common CAS and intermediate outcomes, and align Bank operations (across sectors where relevant) with CAS outcomes.
Engage in preparation and review of Country Assistance Strategy Completion Reports (CASCRs) and other elements of the Regional Monitoring and Learning System.	
Organize the Region's contribution to the proposed Operational Performance and Results Report.	
Network Action Program:	
Describe how the network will organize itself to manage for results.	
Provide advice on the development of performance indicators and implications for design of institutional arrangements and the specification of data requirements for monitoring sectoral performance at the country level.	Critical element going forward. AROE focus groups stated that country teams and task teams are looking to networks for guidance good indicators to use. In addition, AROE focus groups raised difficulties teams were facing in coming up with results chains when developing results frameworks—that is, how to conceptually link higher-level outcomes with specific Bank operations. They were looking to networks for support.
Encourage development of methodologies and approaches to enable comparative assessments and benchmarking of country-level performance in sectors/themes covered by the network.	Critical element going forward. Some networks are selectively engaging.
Identify priorities for selection of sectoral projects/programs that should receive support for the design and implementation of impact evaluations in coordination with the DIME initiative.	
Provide just-in-time training for front-line staff.	Critical element going forward (see comments on OPCS's knowledge and learning strategy, below)
Engage in preparation and review of CASCRs and other elements of the Regional Monitoring and Learning System.	
Organize network's contribution to the proposed Operational Performance and Results Report.	

Report section III—Program of action	AROE assessment
Results Steering Group: Create a Results Steering Group to replace current network of volunteer results focal points.	
<p>Knowledge and Learning: OPCRX will lead the design and implementation of a results knowledge and learning strategy. Strategy will aim to:</p> <ul style="list-style-type: none"> – Clarify and promote the managing for results approach among managers at all levels – Support staff’s efforts to adopt and apply results approaches and knowledge and to identify factors that influence the demand for use of information on results in countries’ decision making – Facilitate the application of a “results lens,” including the provision of training to TTLs preparing projects and identification of the incentive environment at the country/sectoral level for managing for results. – Collaborate with sector and networks to heighten learning about indicators, statistical tools for managing for results – Improve outreach to a globally dispersed staff – Ensure the quality and consistency of results knowledge and learning by further dissemination of good practices. 	Critical element going forward. It includes the key areas identified in AROE. Focus is on traditional learning. Annex F provides details of learning priorities and proposes two new initiatives—a flagship Managing for Results training module and an Institutionalization of Managing for Results self-assessment tool. Bank needs to go beyond passive support (it did not work in MEIP) but be more active in supporting operational activities, say with funding for enhancing government capacity to use M&E information for policy making as per Pillar 1. The Bank should also pursue opportunities to incorporate IEG products into learning strategy.
<p>Results Monitoring and Learning System: To be developed.</p> <p>Tier 1: Country Indicators—IDA 14 + growth, macroeconomic management, social protection, and trade</p> <p>Tier 2: Bank indicators—CASCR ratings, QAG/IEG ratings on project quality</p> <p>Global indicators—global programs</p>	OPCS should consolidate corporate-level monitoring around a common set of key indicators. OPCS should work with SFRM to use SPCs to link corporate-level reporting with resource allocation and regional reporting with the CAS results frameworks and country reporting.
Results Reporting: Operational Performance and Results Report—apex report building on ARPP, CAS Retrospective, and SSIU.	
<i>Pillar III: Strengthening global capacity</i>	
Collaborative Activities: Continuing activities following from the Paris Declaration, the OEDC/DAC Joint Venture for Managing for Results, and the MDB Working Group on Managing for Results	
Finalize Sourcebook on Managing for Results	
COMPAS	
New “Mutual Learning Initiative”	
Develop communities of practice	
Third Roundtable on Managing for Development Results	

Note: OPCRX = Results Secretariat; MEIP = Metropolitan Environmental Improvement Program; SFRM = strategy, finance, and risk management; SPCs = strategy and performance contracts.

APPENDIX H: WORLD BANK MANAGEMENT ACTION RECORD

IEG influences the Bank through recommendations to management as part of sector, thematic, and corporate evaluations, as well as Country Assistance Evaluations. Management is accountable to the Board for follow-up. One of the intermediate outcomes for IEG is the extent to which management incorporates IEG recommendations and findings in policy advice, program design, and project design.

The Management Action Record (MAR) allows IEG to track recommendations from sector, thematic, and corporate evaluations and to monitor the degree to which management has adopted them. The MAR tracks two indicators: the level of adoption and the implementation status of individual recommendations. The MAR presents management's ratings on these two indicators and IEG's validation of the same on an annual basis.

The rating categories for the level of adoption of recommendations are as follows:

- 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
- Negligible—no evidence or plan for adoption, or plans and actions for adoption are in a very preliminary stage.

The categories for the status of recommendations are the following:

- Active and remain actionable by management

- Complete and archived
- Obsolete or overtaken by events and archived
- Difference of Opinion between management and IEG.

The MAR contains information on all IEG recommendations, subsequent management actions, and IEG's assessments. Validation of management actions responds to CODE's interest in a transparent documentation of all recommendations, their implementation status, and substantive differences between management and IEG as to their level of adoption.

The 2005 AROE included recommendations to improve the MAR. A pilot three-year expiration criterion has been introduced for MAR recommendations from fiscal 2006 onward. Recommendations issued in the previous three calendar years (calendar years 2002–04) are being retained in the MAR. Recommendations from earlier years (calendar year 1998–2001) are submitted to CODE/the Board and archived, subject to review by CODE/the Board.

The 2006 Management Action Record

The 2006 MAR tracks management actions on 79 recommendations. These include 21 new recommendations from the six IEG studies (excluding CAEs) presented to the Board in calendar year 2005, and 58 recommendations carried forward from calendar year 2002 to 2004.

For the 2006 MAR, IEG rated 66 percent of its recommendations as having been adopted by management at high or substantial levels (figure H.1). In 2005, 75 percent of recommendations were noted by IEG as having been adopted at a high or substantial level, compared with 62

percent in 2004. The ratings for 2006 reflect a trend of high or substantial adoption by management of about two-thirds of IEG’s recommendations. This year, for six percent of recommendations, implementation was rated by IEG as negligible, which is slightly higher than in previous years. There is a difference of opinion between management and IEG on two recommendations; these have therefore not been rated and are not included in the percentages on adoption.

IEG and management agreed on the rating on level of adoption for 65 percent of recommendations (50 of 77). This represents a decline from 70 percent in 2005 and 78 percent in 2004. It should be noted that the introduction of the three-year expiration rule for recommendations reduces the comparability of these numbers over time, because the 2006 MAR contains proportionally more recent evaluations than the MARs from prior years. Where there was disagreement in the level of adoption, IEG’s ratings were lower than management’s ratings for all except one of the recommendations. Table H.1 shows the distribution of adoption ratings by management and IEG.

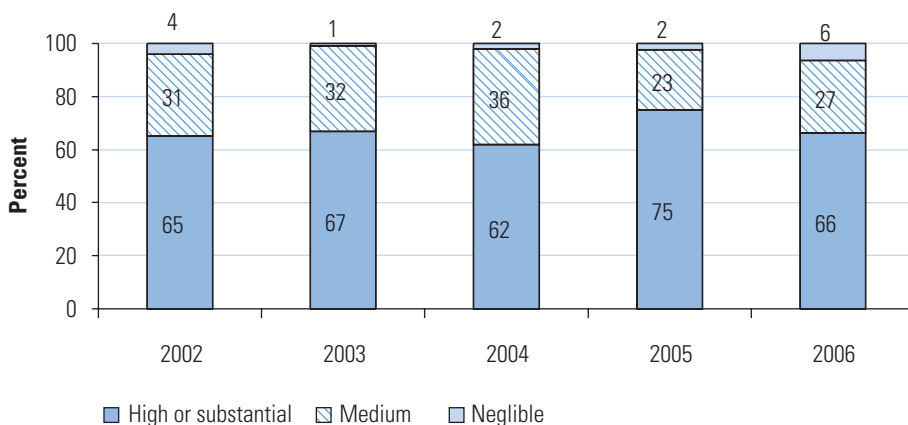
Fifty-eight recommendations have been carried over from the 2005 MAR to the 2006 MAR. IEG upgraded its rating of adoption for about one-

fourth (14) of them and downgraded it for five recommendations. Management downgraded its rating for one recommendation and upgraded eight, five of which IEG concurred with. Management considered 14 recommendations as completed. IEG agreed in seven cases.

IEG and management agreed on the level of adoption for about three-quarters (43 out of 58) of these recommendations. IEG rated 79 percent of the recommendations with “high” or “substantial” for the level of adoption. Examples for positive follow-up action to IEG recommendations from calendar year 2002–04 are the reassessment of the Bank’s approach to infrastructure in response to IEG’s evaluation of private sector development in the power sector; reforms of the PRSP Joint Staff Assessment following IEG’s PRSP study; and changes in the HIPC debt sustainability analysis in response to IEG’s HIPC evaluation.

Of the 21 new recommendations entered into the MAR during calendar year 2005, IEG concurs with management’s adoption ratings on 40 percent (8) of the recommendations. There is one difference of opinion. Management rates its adoption high or substantial for 85 percent (17 of 20) of the recommendations and medium for 15 percent. IEG considers the adoption of 29

Figure H.1: Level of Adoption of IEG Recommendations Similar to Previous Years



Source: IEG data.

Table H.1: Management and IEG Ratings of Recommendation Adoption Levels

Management ratings	IEG ratings				Difference of opinion	Sum of management ratings
	High	Substantial	Medium	Negligible		
High	17	8	7	2		34
Substantial		25	6	2		33
Medium		1	8	1		10
Negligible						
Difference of opinion					2	2
Sum of IEG ratings	17	34	21	5	2	79

Source: IEG data.

Note: Highlighted fields show agreement between management and IEG; fields to the right of the diagonal represent higher rating by management than by IEG, fields to the left of the diagonal show higher rating by IEG than by management. The sixth column shows the sum of ratings in each category by management; the sixth row shows the sum of ratings by IEG.

percent of the recommendations “high” or “substantial,” 43 percent “medium,” and 24 percent “negligible.”

For MAR recommendations from calendar year 1998–2001 that are being archived, among 7 IEG reports, 10 recommendations are incomplete, and there is a difference of opinion on one recommendation.

Selected Examples of Follow-up Action

The 2006 MAR contains a number of recommendations that have prompted responsive actions by management and improvements of Bank operations. MAR recommendations that have had an impact include IEG’s review of lines of credit (IEG 2006e), the global programs review (IEG 2004a), and the evaluation of the World Bank Group’s experience with extractive industries (IEG-WB, -IFC, -MIGA 2005).

IEG’s 2004 evaluation of lines of credit (LOCs) offers a good example of corrective action by management following IEG recommendations. IEG found poor results of Bank-funded LOCs and widespread lack of compliance with operational policies (OPs) governing LOCs. It recommended that Bank management update OP 8.30 to cover all LOCs across all sectors and Regions. CGAP’s self-evaluation of LOCs for compliance with OP 8.30 confirmed IEG’s findings. To address IEG’s recommendations,

management subjected all identified LOCs to a quality assurance system to foster compliance with OP 8.30. Comments from a staff survey indicate that the evaluation has affected projects; for example, one comment was: *“evaluation of Bank LOCs has provided guidance on our recent design of an LOC in the area of micro and small lending. Lessons learned in that report have proved very enlightening.”*

IEG’s evaluation of the World Bank’s approach to global programs found that the Bank’s selectivity and oversight of its global program portfolio had been weak, because the existing selectivity criteria were broad and difficult to apply and because of poorly defined expectations with respect to the roles, responsibilities, and accountabilities of Bank staff who oversee individual programs.

In response to IEG’s recommendation that the Bank improve its selectivity and oversight of global programs, management has put in place new business processes that recognize GPPs as a separate Bank product line and has initiated work on a monitoring framework and indicators to assist global program task managers in improving oversight. QAG has also initiated quality-at-entry reviews of new DGF-supported programs to help strengthen the adoption of these new business processes. Management

rated the level of adoption of all four recommendations from this report high; IEG rated the adoption as substantial for two recommendations and high for the remaining two.

IEG's evaluation of the World Bank Group's experience with extractive industries (EI) found that, while the Bank's EI projects have produced positive economic and financial results, the Bank could do more to enhance the EI sector's contribution to sustainable development and poverty reduction. Among the issues highlighted were ensuring greater compliance with environmental and social safeguards, promoting improved governance and management of EI revenues, and developing an integrated sector strategy that went beyond returns to investors.

Following IEG's recommendations, management has adopted a clearer strategy for the sector and is mainstreaming new guidelines for better addressing EI sector issues in strategy documents (CAS), particularly for low-income countries under stress. Guidelines are being prepared for improved upstream screening of projects for safeguards requirements. The EI Transparency Initiative is being expanded, and the distribution of project benefits among key stakeholders is being done as standard practice. A World Bank Group community development facility has been established for increasing local community participation in EI projects. The level of adoption for all recommendations from the EI evaluation was rated high by both management and IEG in the 2006 MAR.

APPENDIX I: MANAGEMENT RESPONSE

Introduction

The *2006 Annual Report on Operations Evaluation* (AROE) updates the actions taken since the 2004 and 2005 AROEs to strengthen the results focus in monitoring and evaluation. It explores first the implications for managing for results in Bank operations and second the extent to which the Bank's M&E systems provide staff with the information they need to better manage for results. The report also evaluates the products and services of IEG. The AROE also makes recommendations for further progress on implementing the results architecture and on ways IEG might encourage more timely and widespread use of IEG findings by operational staff.

Emerging results architecture

Management welcomes this AROE, the third in as many years to focus on the Bank's progress in articulating a results architecture and procedures for sharpening its focus on results. This year's report confirms the relevance of the emerging results architecture. That architecture is grounded in the country business model that envisions a results management system focused on the articulation of an RBCAS as the overarching framework for prioritizing the Bank's program of lending and knowledge support to client countries. Key supporting elements are the CAS mid-term progress report and self-evaluation through the preparation of a CASCR. Last, the CASCR ratings of program achievements and Bank performance are validated by IEG.

The AROE's analysis confirms management's own assessment of recent progress (World Bank 2006b) and contributes more in-depth analyses, particularly of RBCASs and feedback from staff

focus groups on the opportunities and challenges of strengthening the results focus of operational activities. The analysis is a useful complement to management's work to deepen its understanding of the key factors that will determine the success of its results initiative, notably how best to encourage staff efforts and improve the quality and utility of the key instruments in the results architecture.

AROE Analysis

The AROE catalogues the measures taken to strengthen the Bank's results orientation in M&E at the country, sector, and product levels, updating the overview in the 2003 and 2004 AROEs and focusing particularly on the current status of the M&E tools. Like management's recent update on the results agenda (World Bank 2006b), it notes progress but also areas for improvement.

Challenges to developing results-based monitoring

The AROE's compilation of the challenges to developing results-based monitoring systems, derived from focus group discussions within the Bank, yields a familiar listing of key issues: (a) accelerating work to help staff overcome gaps in their understanding of the practical dimensions of constructing useful results frameworks, (b) ensuring client capacity and buy-in, and (c) helping country and task teams integrate the demands of the results agenda into already overloaded work programs.

Overcoming internal challenges

Work to overcome many of these conceptual constraints to more active implementation (moving from seeing M&E as primarily a tool of

relevance to project completion to viewing it as a tool to facilitate responsive management during implementation) is under way. Notably, management is developing a Bank-wide learning program that is supported by the Knowledge and Learning Board.

A critical element of creating a results culture in the Bank is to align the incentive and rewards structure with the results agenda. To improve incentives, management has made portfolio quality oversight a required performance evaluation item for sector managers. Staff have also been recognized Bank-wide in a variety of ways for their contributions to results, notably through the President's Awards for Excellence, IEG's Good Practice Awards, and the Results Showcase, which brought competitively selected good practice cases to senior management attention.

In addition, management has moved to increase the focus on results in its learning and accountability processes around operational products (RBCASs, operational product reviews during preparation and implementation, and the QAG process). More needs to be done, and management is committed to continuing to focus on incentives and expanding programs to communicate to staff the importance of this agenda. Two key elements going forward are management attention to the IDA 14 results measurement system and the results monitoring and learning system under preparation (see the results framework for the agenda on managing for results in the recent update) (World Bank 2006b, p. 33).

The global element in the focus on results

The AROE reviews progress internally in improving the results focus and externally in supporting countries in strengthening their focus on results. Management notes, however, that its work on the results agenda includes a third element, at the "global" level. The Bank has put considerable effort into harmonizing its approach with that of other development partners and working with them to align with country results efforts, a topic that this year's

AROE analysis does not explore. Management's recent update, discussed at the Board's Committee on Development Effectiveness on March 8, 2006, documents the progress on this front (World Bank 2006b, p. 24).

Challenges at the country level

The AROE rightly cites these challenges. As noted in the recent update, in countries preparing PRSs, the Bank has continued to encourage countries to improve their PRS monitoring systems through ESW, lending for public sector management, training programs, and knowledge sharing. The Bank has given attention to helping partner countries gather national and sub-national data for monitoring progress toward their PRS goals. Analytic work and policy advice focus on aligning allocations with the PRS priorities and on achieving operational efficiency, service delivery, and outcomes.

However, more needs to be done to stimulate greater demand for results within and across the line agencies that are most directly responsible for the delivery of public goods and services, as well as from citizens. One implication is that what is needed is a shift in incentives—from donor-driven results reporting to building the countries' capacity to provide credible evidence of effectiveness to their own citizens. Increasingly, management seeks to signal to staff that they are accountable for helping their country partners build systems and procedures that enhance accountability and transparency. One element in this process is the Development Impact Evaluation Initiative. This strategic scaling up of impact evaluations will provide countries with additional information that they can use in evidence-based decision making.

Self-evaluation tools and the bank's results agenda

Management agrees with the AROE analysis that there is a need to bring a results focus to the methods and tools that are used to improve performance during project implementation. The Rapid Results Approach cited by the AROE seems to hold promise, and management is assessing this experience. Management

envisions this as the first of a series of collaborative learning activities across Regions, networks, and central service units designed to develop tools and examples that task and country teams can use in “real time” as they struggle to meet—and to help countries meet—the challenge of managing for results.

AROE Recommendations

The AROE 2006 proposes two core recommendations for management consideration:

- Building on progress in advancing the results agenda, agree on a three-year action program with a corresponding budget to fund implementation of the next stage, differentiating new allocations and redeployed resources. Critical action items would be as follows:
 - Support country directors and country teams in their efforts to refine and use results frameworks at country and sectoral levels to manage country programs. Similar efforts need to be undertaken to strengthen the results frameworks of thematic and global programs.
 - Assess the effectiveness of (self-) evaluation approaches during project and country program implementation and provide guidelines to staff on their use.
 - Strengthen incentives and accelerate a results-oriented training and communica-

tions program for management and staff to encourage use of M&E information.

- Identify and support in-depth learning opportunities to develop and use results-based approaches with operational teams, particularly in challenging country cases and in complex multisectoral settings.
- To enhance countries’ capacity and demand to manage for results by strengthening the collection and use of performance information, provide budgetary support to task teams and technical advice to countries that intend to institutionalize M&E systems.

Management views

Management broadly agrees with the thrust of these recommendations, with one major exception. A results focus is the responsibility of country teams and task teams and needs to be included in the formulation of RBCASs and in task preparation and task supervision budgets. To move forward on agreed priorities, a Results Steering Group has been formed to facilitate communication and coordination among Regional and network results teams, which are now charged with responsibility for leadership in identifying the priorities for implementing the action plan proposed in the recent progress report. A response to the AROE recommendations is included in the Summary of Management Response to IEG Recommendations (p. xv).

ENDNOTES

Chapter 2

1. The figure was adapted from an actual country example in the East Asia and Pacific Region. It is based on a middle-income country with significant Bank program. Smaller country programs may have fewer projects. In such cases, a results chain may help to identify the key gaps in the country program.

2. Performance indicators, taken from the Project Appraisal Document, had been included in an annex to the PSR and were often ignored and not updated. The ISR brought these indicators into the main document, giving them greater visibility.

3. The 2005 AROE found that only 18 percent of the 25 CASs reviewed were considered well targeted and 13 percent adequately baselined. However, the RBCAS had not been mainstreamed when these CASs were prepared, so a direct comparison between 2005 and 2006 CAS reviews is not appropriate.

4. Milestones are “outputs, actions, or outcomes expected to be realized during CAS implementation and which can serve as progress makers toward CAS outcomes.”

5. QAG went on to observe that problem of defining indicators was further compounded by the absence of practical intermediate indicators and failure to regularly update the key performance indicators and to use them to shed light on obstacles to outcomes through the implementation period (World Bank 2005b).

6. The Bank’s safeguard and fiduciary policies had led to a doubling of the Regional procurement and financial management staff between fiscal 1998 and 2003. Between fiscal 2000 and 2003, headquarters expenditures on compliance with environmental and social safeguards, including costs during project preparation, increased by 52 percent from an estimated \$21 million to about \$32 million. No update on these costs is currently available.

7. A program chain logic assessment would review: (a) the deployment and sequencing of the ac-

tivities, resources, or policy initiatives to achieve the project or program outcome (the desired change in beneficiaries) and (b) the plausibility of achieving that desired outcome, based on similar prior efforts and on the research literature. The purpose is to avoid failure from a weak design that would have little or no chance of success in achieving its intended outcomes.

8. Impact evaluation establishes a counterfactual to make a comparison of the situation with and without the project or program.

9. OPCS announced the creation of impact evaluation as a new AAA product line, effective July 14, in an effort to respond to an increasing demand for impact evaluation both within the country and the Bank. This would also allow the Bank to track the amount of Bank resources being used for such activities (World Bank 2005d).

10. IEG’s evaluation of extractive industries found that ex post cost-benefit analysis was done in 80 percent of reviewed investment projects, but hardly at all in technical assistance or sectoral adjustment loans. Similarly, IEG’s recent evaluation of agricultural water management found that “project economic analysis seems to have been allowed to slide over the past 20 years.” Also, IEG found, in these and other evaluations, that many cost-benefit analyses are flawed or too simplistic.

Chapter 3

1. IEG’s independence has been validated according to the standards of OECD/DAC and the U.S. Government Accountability Office (among others) (IEG 2003).

2. The three key criteria for assessing the results orientation of M&E are the results framework, design of M&E, and implementation of M&E (IEG 2004b, box 4).

3. A good example for a CASCR and CASCR Review is the Philippines CASCR from April 2005 and IEG’s re-

view from May 2005. This CASCR won IEG's 2006 Good Practice Award for a "candid and thoughtful assessment of the achievements and deficiencies in implementing the Country Program" (quoted from <http://www.worldbank.org/ieg/awards>).

4. A desk review showed that all 38 calendar year 2005 PPARs used the word *outcome* an average of 16 times, but only 53 percent (20) of these PPARs used the word *output* more than once.

5. The award categories are Projects, ICRs, Country Programs, Country Assistance Strategy Completion Reports (CASCRs), Monitoring and Evaluation, and Initiatives with Demonstrated Impact/Results.

6. During the 12-month period April 2005–March 2006, ECD materials accounted for 36 percent of total recorded downloads of IEG documents.

7. In previous years, these were yes/no questions and four-point scales. This year, respondents were asked to rate all questions in the three sections on a six-point scale. For further details on the conversion, see appendix D.

8. For example, one respondent wrote, "It was not really useful or relevant to our work. We would have liked the evaluation team to think on [sic] the context in which we operate (rather than on its Washington stakeholders), so as to contribute ideas that can be useful for our teams at the trenches."

9. In the first four months of 2006, IEG was mentioned more than 20 times in leading English language mass media, which is almost triple the number of mentions of IEG in each year from 1996 to 2004.

10. Level of adoption ratings are high–fully adopted; substantial–largely adopted but not fully incorporated into policy, strategy, or operations as yet; medium–adopted in some operational and policy work but not to a significant degree in key areas; and negligible–no evidence or plan for adoption, or plans and actions for adoption are in a very preliminary stage.

11. The status of recommendations is rated as active and remain actionable by management; complete and archived; obsolete or overtaken by events and archived; and difference of opinion between management and IEG.

12. Due to the new 3-year expiration criterion, comparability between the numbers of rated recommendations in the 2006 MAR to MARs from prior years is limited.

13. Two recommendations constitute a difference of opinion and did not receive an adoption rating.

These are excluded from the sample when comparing levels of adoption.

Chapter 4

1. The Bank has been focusing on improving the outcomes of its projects since the Wapenhans Task Force Report in the early 1990s. In response to the report, the Bank introduced project logframes, which linked project components with outputs, project development objectives and country (CAS) objectives, and sector key performance indicators, which could be used to monitor performance. QAG's review of quality of supervision for fiscal 1999 found that monitoring was satisfactory in 58 percent of the projects reviewed and that task teams do not know how to improve monitoring and evaluation or do not have the resources to do so, despite technical support from OPCS (or its equivalent at that time). The Bank responded with an effort to strengthen M&E at the project level. A working group, organized in 1999 to come up with an action plan, concluded that the weaknesses in M&E were due to (a) limited incentives to systematically focus on results, (b) unclear roles and accountabilities in the Bank and between the Bank and borrowers, and (c) inadequate capacity to carry out M&E.

Appendix A

1. For investment loans/credits, Template and Guidelines for the Project Appraisal Document, from Operations Policy and Country Services (OPCS) is used. For development policy loans, OPCS's *Good Practice Note for Development Policy Lending: Results in Development Policy Lending* is used (World Bank 2005a).

2. A new OP/BP 8.60, Development Policy Lending, was approved in August 2004 to replace the previous OD 8.60, Adjustment Lending, and nine operational memoranda. A series of good practice notes, including one on results, has been issued to provide guidance to staff on specific aspects when designing and implementing a development policy loan.

3. Previously the performance indicators were included in an annex, which was not automatically attached to the report when it was viewed electronically. The annex tended to be neglected and data often were not updated regularly.

4. In its fiscal 2005 Annual Report on Portfolio Performance, the Quality Assurance Group (QAG) noted

a slowdown in analytical and advisory activities (AAA) implementation and attributed this to a lack of management attention (after the concept stage) and difficulties in using the system. QAG observed that “some of those leading the AAA work in the Regions view the Bank-wide task reporting system as virtually unusable and pay little attention to it” and that the users “argue that creation of a separate code for each task combined with the lack of user-friendliness in inputting information and updating the system leads to inaccuracy in reporting.”

5. About 75 percent of the CPPRs reviewed were attempting to incorporate elements of a results approach, although only 16 percent could be considered results-based country performance reviews.

6. The IDA 14 Results Measurement System includes 14 country outcome indicators covering areas that are consistent with the MDGs, are priorities in Poverty Reduction Strategies, and reflect the International Development Association’s (IDA) activities: income/poverty, health, gender, basic education, water supply, infrastructure, private sector development, governance, and economic development.

7. The Annual Report on Portfolio Performance reports on the percentage of project that have meet their PDOs, assesses the percentage of ongoing projects which may be at risk of not achieving their objectives, and recommends actions the Bank needs to take to remedy problems identified through the analysis.

8. These are percent of IDA projects with (a) satisfactory outcome ratings, (b) satisfactory quality at entry ratings, (c) first ISRs with satisfactory baseline data on expected outcomes, and (d) Implementation Completion Reports with satisfactory data on project outcomes.

9. The COMPAS master matrix of categories and indicators is included as annex 4.1 in the *Global Monitoring Report* (World Bank 2006a).

10. Revised formally through approval by the Bank authority that approved the original loans, credits, or grants.

11. Previously, the ACS was required for activities costing more than \$50,000. In 2004, the Bank simplified the tracking of AAA but required ACS for all activities, irrespective of cost, because of the increase in the number of tasks costing less than \$50,000 and the Bank’s emphasis on tracking and results and the impacts of operational activities.

12. The three GPRs were on Partnership for Child Development, the Trust Fund for Statistical Capacity Building, and the Global Corporate Governance Forum.

Appendix C

1. In addition, IEG reviews the CASCR.

2. The section on implementation states: “The Bank collaborates with the government and external partners in implementing CAS activities. The country team monitors and evaluates progress toward achieving CAS objectives, including outcomes identified in the results framework. The CAS M&E system uses national M&E systems, where possible, and draws on the M&E arrangements built into Bank operations.”

3. Capability maturity models were first developed at Carnegie-Mellon University in the mid-1980s as a way of comparing management practices of private sector companies providing services to the U.S. Department of Defense. Maturity models have since been developed and employed across a wide variety of management practices. A maturity model identifies key characteristics that an evaluator expects to find as the process in question becomes more fully developed and integrated into related management activities. The maturity model was chosen for this analysis because it does not require a particular process to be employed; it requires only that the process chosen produces the desired characteristics.

4. The CASs covered the following countries: Albania, Bangladesh, Bhutan, Burkina Faso, Cambodia, Djibouti, the Dominican Republic, Egypt, El Salvador, Gabon, Georgia, Guatemala, Jamaica, Lao People’s Democratic Republic, Morocco, Nigeria, Philippines, Poland, São Tome and Principe, Sierra Leone, Tajikistan, Timor-Leste, Uganda, Uruguay, and Uzbekistan.

Appendix E

1. Before 2005, individuals did not receive a unique survey login and could enter the survey multiple times. They might have had multiple records within the database. For the 2005 and 2006 surveys, individuals were limited to a single record but could re-enter that record until they had completed the survey to their satisfaction.

2. This new approach to obtaining IEG client views is based on the Massachusetts Institute of Technology Center for Technology, Policy, and Industrial Development approach to managing information quality, which compares consumers’ ratings

of importance of attributes of information quality with their satisfaction with these attributes. Assuming that all attributes of quality listed in the survey are important, clients rated their satisfaction with these attributes. In previous years, clients rated evaluations on the attributes themselves. The switch to a six-point scale was made for two reasons: to achieve coherence with the six-point scale of satisfaction (highly unsatisfactory-highly satisfactory) that is used in IEG evaluations and has also recently been adopted by QAG; and to provide room for clearer distinction between the different attributes.

BIBLIOGRAPHY

- Abadzi, Helen. 2006. *Efficient Learning for the Poor: Insights from the Frontier of Cognitive Neuroscience. Directions in Development*. Washington, DC: World Bank.
- Behn, Robert D. 2003. "Why Measure Performance? Different Purposes Require Different Measures." *Public Administration Review* 63 (5).
- Bennett, Charles, and Herb Hill. n.d. "Performance Management Systems: The Importance of Defining Their Purpose." Richmond, Virginia: Virginia Department of Planning and Budget. <http://www.dpb.state.va.us/VAResults/Material/PolicyService.pdf>.
- Binnendijk, A. 2001. "Results-Based Management in the Development of Co-operation Agencies." Paris: Organisation for Economic Co-operation and Development–Development Assistance Committee.
- Harvard University, Kennedy School of Government, Public Sector Performance Management, Executive Session. 2001. *Get Results Through Performance Management: An Open Memorandum to Government Executives*. Cambridge, MA.
- Hatry, Harry. 2003. "Results Matter: Suggestions for a Developing Country's Early Outcome Measurement Effort." In *Handbook on Public Sector Performance Review*, Volume II, *Measuring Government Performance in the Delivery of Public Services*, ed. A. Shah. Washington, DC: World Bank.
- IEG (Independent Evaluation Group-World Bank). Various years. IEG Good Practice Award Web site. <http://www.worldbank.org/ieg/awards>.
- . 2006a. *Assessing World Bank Support for Trade, 1987–2004*. Washington, DC: World Bank.
- . 2006b. *Engaging with Fragile States: An IEG Review of World Bank Support to Low-Income Countries Under Stress*. Washington, DC: World Bank.
- . 2006c. *From Schooling Access to Learning Outcomes: An Unfinished Agenda. An Evaluation of World Bank Support to Primary Education*. Washington, DC: World Bank.
- . 2006d. *Hazards of Nature, Risks to Development: An IEG Evaluation of World Bank Assistance for Natural Disasters*. Washington, DC: World Bank.
- . 2006e. *World Bank Lending for Lines of Credit: An IEG Evaluation*. Washington, DC: World Bank.
- . 2005a. *2005 Annual Report on Operations Evaluation*. Washington, DC: World Bank.
- . 2005b. *Capacity Building in Africa: An OED Evaluation of World Bank Support*. Washington, DC: World Bank.
- . 2005c. *Committing to Results: Improving the Effectiveness of HIV/AIDS Assistance: An OED Evaluation of the World Bank's Assistance for HIV/AIDS Control*. Washington, DC: World Bank.
- . 2005d. *Country Assessment Evaluation Retrospective: An OED Self-Evaluation*. Washington, DC: World Bank.
- . 2005e. *The Effectiveness of World Bank Support for Community-Based and -Driven Development*. Washington, DC: World Bank.
- . 2005f. "Mandate of the Director-General, Operations Evaluation." <http://www.worldbank.org/ieg/about.html>.
- . 2004a. *Addressing the Challenges of Globalization: An Independent Evaluation of the World Bank's Approach to Global Programs*. IEG Study Series. Washington, DC: World Bank.

- . 2004b. *2004 Annual Report on Operations Evaluation*. Washington, DC: World Bank.
- . 2004c. *Evaluation Capacity Development: OED Self-Evaluation*. Washington, DC: World Bank.
- . 2004d. *The Poverty Reduction Strategy Initiative: An Independent Evaluation of the World Bank's Support Through 2000*. Washington, DC: World Bank.
- . 2003. "Independence of OED." *OED Reach*. February 23. [http://lnweb18.worldbank.org/oed/oeddoclib.nsf/b57456d58aba40e585256ad400736404/50b9b24456b788be85256ce0006ee9f8/\\$FILE/Independence_Reach.pdf](http://lnweb18.worldbank.org/oed/oeddoclib.nsf/b57456d58aba40e585256ad400736404/50b9b24456b788be85256ce0006ee9f8/$FILE/Independence_Reach.pdf).
- . 2002. *The World Bank's Approach to Global Programs: An Independent Evaluation. Phase 1*. Washington, DC: World Bank.
- IEG-World Bank, -IFC, and -MIGA (Independent Evaluation Group-World Bank, -International Finance Corporation, and -Multilateral Investment Guarantee Agency). 2005. *Extractive Industries and Sustainable Development—An Evaluation of World Bank Group Experience*. Washington, DC: World Bank.
- . 2003. *Power for Development: A Review of the World Bank Group's Experience with Private Participation in the Electricity Sector*. Washington, DC: World Bank.
- IEG/Thematic Group for Poverty Analysis, Monitoring, and Impact Evaluation. 2006a. "Conducting Quality Impact Evaluations under Budget, Time and Data Constraints." Washington, DC: World Bank.
- . 2006b. "Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management." ECD Working Paper Series, No. 15. Washington, DC.
- Kusek, Jody, and Ray Rist. 2004. *Ten Steps to a Results-based Monitoring and Evaluation System*. Washington, DC: World Bank.
- Mackay, Keith. 2006. "Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management." IEG Evaluation Capacity Development Working Paper Series, No. 15. Washington, DC: World Bank.
- Rist, Ray, and N. Stame, eds. 2006. *From Studies to Streams: Managing Evaluative Systems. Comparative Policy Evaluation, Vol 12*. New Brunswick, NJ: Transaction Publishers.
- World Bank. n.d. *Template and Guidelines for the Project Appraisal Document*. Washington, DC.
- . 2006a. *2006 Global Monitoring Report*. Washington, DC.
- . 2006b. *Accelerating the Results Agenda: Progress and Next Steps*. Washington, DC.
- . 2006c. *FY05 Annual Report on Portfolio Performance*. Quality Assurance Group. Washington, DC.
- . 2006d. *Quality at Entry in FY04–05 (QEA7)*. Quality Assurance Group. Washington, DC.
- . 2005a. *Good Practice Note for Development Policy Lending: Results in Development Policy Lending*. Washington, DC.
- . 2005b. *Quality of Supervision in FY03–04 (QSA6)*. Quality Assurance Group. Washington, DC.
- . 2005c. "Reporting in the ISR on Progress Toward Outcomes for Investment Lending Operations." Operations Policy and Country Services—Advisory Services. Washington, DC.
- . 2005d. "Updating the Bank's Operational Policy on Monitoring and Evaluation." Consultation Paper. Washington, DC.
- . 2004. "Implementation of the Agenda on Managing for Results: Progress Report." Operations Policy and Country Services. Washington, DC.
- . 2003a. *Country Assistance Strategies: Retrospective and Future Directions*. Operations Policy and Country Services. Washington, DC.
- . 2003b. *Quality of Supervision in FY01/02 (QSA5)*. Quality Assurance Group. Washington, DC.
- . 2000. *Report of the Working Group on Improving Quality of Monitoring and Evaluation in Bank Financed Projects*. Washington, DC.
- The World Bank, Portfolio Management Task Force. 1992. *Effective Implementation: Key to Development Impact: Report of the World Bank's Portfolio Management Task Force*

- (also known as *Wapenbans Report*). Washington, DC.
- Zaltman, Ariel. 2006. "Experience with Institutionalizing Monitoring and Evaluation Systems in Five Latin American Countries: Argentina, Chile, Columbia, Costa Rica, and Uruguay." IEG ECD Working Paper Series No. 16. Washington, DC.

IEG PUBLICATIONS

Study Series

2004 Annual Review of Development Effectiveness: The Bank's Contributions to Poverty Reduction
Addressing the Challenges of Globalization: An Independent Evaluation of the World Bank's Approach to Global Programs
Agricultural Extension: The Kenya Experience
Assisting Russia's Transition: An Unprecedented Challenge
Bangladesh: Progress Through Partnership
Brazil: Forging a Strategic Partnership for Results—An OED Evaluation of World Bank Assistance
Bridging Troubled Waters: Assessing the World Bank Water Resources Strategy
Capacity Building in Africa: An OED Evaluation of World Bank Support
The CIGAR at 31: An Independent Meta-Evaluation of the Consultative Group on International Agricultural Research
Country Assistance Evaluation Retrospective: OED Self-Evaluation
Debt Relief for the Poorest: An OED Review of the HIPC Initiative
Developing Towns and Cities: Lessons from Brazil and the Philippines
The Drive to Partnership: Aid Coordination and the World Bank
Economies in Transition: An OED Evaluation of World Bank Assistance
The Effectiveness of World Bank Support for Community-Based and –Driven Development: An OED Evaluation
Evaluating a Decade of World Bank Gender Policy: 1990–99
Evaluation of World Bank Assistance to Pacific Member Countries, 1992–2002
Financial Sector Reform: A Review of World Bank Assistance
Financing the Global Benefits of Forests: The Bank's GEF Portfolio and the 1991 Forest Strategy and Its Implementation
Fiscal Management in Adjustment Lending
IDA's Partnership for Poverty Reduction
Improving the Lives of the Poor Through Investment in Cities
India: The Dairy Revolution
Information Infrastructure: The World Bank Group's Experience
Investing in Health: Development Effectiveness in the Health, Nutrition, and Population Sector
Jordan: Supporting Stable Development in a Challenging Region
Lesotho: Development in a Challenging Environment
Mainstreaming Gender in World Bank Lending: An Update
Maintaining Momentum to 2015? An Impact Evaluation of Interventions to Improve Maternal and Child Health and Nutrition Outcomes in Bangladesh
The Next Ascent: An Evaluation of the Aga Khan Rural Support Program, Pakistan
Nongovernmental Organizations in World Bank–Supported Projects: A Review
Poland Country Assistance Review: Partnership in a Transition Economy
Poverty Reduction in the 1990s: An Evaluation of Strategy and Performance
The Poverty Reduction Strategy Initiative: An Independent Evaluation of the World Bank's Support Through 2003
Power for Development: A Review of the World Bank Group's Experience with Private Participation in the Electricity Sector
Promoting Environmental Sustainability in Development
Putting Social Development to Work for the Poor: An OED Review of World Bank Activities
Reforming Agriculture: The World Bank Goes to Market
Sharing Knowledge: Innovations and Remaining Challenges
Social Funds: Assessing Effectiveness
Tunisia: Understanding Successful Socioeconomic Development
Uganda: Policy, Participation, People
The World Bank's Experience with Post-Conflict Reconstruction
The World Bank's Forest Strategy: Striking the Right Balance
Zambia Country Assistance Review: Turning an Economy Around

Evaluation Country Case Series

Bosnia and Herzegovina: Post-Conflict Reconstruction
Brazil: Forests in the Balance: Challenges of Conservation with Development
Cameroon: Forest Sector Development in a Difficult Political Economy
China: From Afforestation to Poverty Alleviation and Natural Forest Management
Costa Rica: Forest Strategy and the Evolution of Land Use
El Salvador: Post-Conflict Reconstruction
India: Alleviating Poverty through Forest Development
Indonesia: The Challenges of World Bank Involvement in Forests
The Poverty Reduction Strategy Initiative: Findings from 10 Country Case Studies of World Bank and IMF Support
Uganda: Post-Conflict Reconstruction

Proceedings

Global Public Policies and Programs: Implications for Financing and Evaluation
Lessons of Fiscal Adjustment
Lesson from Urban Transport
Evaluating the Gender Impact of World Bank Assistance
Evaluation and Development: The Institutional Dimension (Transaction Publishers)
Evaluation and Poverty Reduction
Monitoring & Evaluation Capacity Development in Africa
Public Sector Performance—The Critical Role of Evaluation

IEG

INDEPENDENT EVALUATION GROUP



THE WORLD BANK



ISBN 0-8213-6845-1