

IEG

INDEPENDENT EVALUATION GROUP

WORLD BANK



2005 Annual Report on Operations Evaluation



THE WORLD BANK

INDEPENDENT EVALUATION GROUP

ENHANCING DEVELOPMENT EFFECTIVENESS THROUGH EXCELLENCE AND INDEPENDENCE IN EVALUATION

The Independent Evaluation Group (IEG) reports directly to the Bank's Board of Executive Directors. IEG assesses what works, and what does not; how a borrower plans to run and maintain a project; and the lasting contribution of the Bank to a country's overall development. The goals of evaluation are to learn from experience, to provide an objective basis for assessing the results of the Bank's work, and to provide accountability in the achievement of its objectives. It also improves Bank work by identifying and disseminating the lessons learned from experience and by framing recommendations drawn from evaluation findings.



2005 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
Manufactured in the United States of America

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

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

Rights and Permissions

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

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

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

Cover photo: Curt Carnemark.

ISBN-10: 0-8213-6521-5
ISBN-13: 978-0-8213-6521-2
e-ISBN: 0-8213-6522-3
DOI: 10.1596/978-0-8213-6521-2

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

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



Printed on Recycled Paper



Contents

v	Acknowledgments
vii	Foreword
xi	Acronyms and Abbreviations
3	1 Introduction
	3 Relationship of the 2005 AROE to Previous AROEs
	3 Importance of a Country Focus
	4 Overview of the 2005 AROE
7	2 Measuring and Managing Performance at the Country Level
	8 Strategic Planning
	11 Performance Measurement
	16 Performance Management
	18 Conclusions
23	3 Update on Monitoring and Evaluation of Bank Operations
	23 Monitoring of Bank Operations
	28 Evaluation
	31 Conclusions
35	4 Improving IEG's Effectiveness
	35 IEG's Response to 2004 AROE Recommendations
	37 IEG Client Survey Results
	38 Recommendation for IEG
39	5 Conclusions and Recommendations
	39 Recommendations
41	Appendixes
	43 A: Review of CAS Results Matrices
	57 B: Country Director Survey Results on the Use of Monitoring and Evaluation Systems and Information at the Country Level
	63 C: Monitoring and Evaluation Framework in the Bank
	65 D: IEG Client Surveys
	75 E: Management Response
	81 F: Chairman's Summary: Committee on Development Effectiveness (CODE)

85 **Endnotes**

87 **References**

Boxes

- 7 2.1 Elements of Results-Based Management
- 9 2.2 What Is a Results-Based CAS?
- 12 2.3 Harmonization Around Results
- 13 2.4 Observations from IEG's Review of the Poverty Reduction Strategy Initiative
- 14 2.5 IEG Lessons on Harmonization from Evaluation
- 15 2.6 Developing the Capacity to Manage by Results at the Country Level
- 17 2.7 How Performance Information Is Used
- 19 2.8 Incorporating Results in Portfolio Reviews
- 32 3.1 Lessons on Results-Based Management in the U.S. Federal Government

Figures

- 10 2.1 Distribution of Performance Measures by Evaluation Level
- 17 2.2 Important Uses of Information



Acknowledgments

This report was prepared by a team comprised of Janardan Prasad (J.P.) Singh (Task Manager), Lauren Kelly, Howard Smith, and Yoshine Uchimura. The peer reviewers were Harry Hatry, Hartwig Schafer, and Hasan A. Tuluy, who provided invaluable advice and guidance in preparing the report.

IEG is grateful to the country directors who participated in the survey conducted for the report and who shared their time and experience. Our thanks also go to Nick Gompper (Caliber Associates), Vivian R. Jackson, and Ray Rist and for their support in survey design and analysis.

Many IEG staff, including Soniya Carvalho, Laurie Effron, Victoria Elliott, Nils Fostevedt, Patrick Grasso, Fareed Hasan, John Johnson,

Keith Mackay, and Kyle Peters, provided extensive support and contributions to all phases of this report. Basil G. Kalvasky and F. Stephen O'Brien provided feedback on the final draft. We gratefully acknowledge their assistance.

Others who have provided valuable input and assistance include Princess Moore-Lewis. Lydia Ndebele and Juicy Qureishi-Huq provided administrative support. Alex H. McKenzie, Julius Gwyer, Maria J. Mar, 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, Independent Evaluation, World Bank Group: *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*



Foreword

The World Bank launched its Managing for Results initiative in 2002 to enhance the effectiveness of its operations. The focus on results is a welcome development and reflects an interest within the development community and the Bank to better demonstrate the effectiveness of its assistance, and needs to be encouraged.

The *2005 Annual Report on Operations Evaluation* (AROE) focuses on the country because it is the main unit of account for monitoring, managing, and evaluating performance. It examines Country Assistance Strategies (CASs) for how well they link country goals, CAS outcomes, and Bank programming. It also examines how Bank information from M&E systems is actually used to manage the performance of Bank country programs. Finally, the report takes stock of the measures taken since the 2003 and 2004 AROEs to strengthen the results focus in monitoring and evaluation.

Measuring and Managing Performance at the Country Level

The Results-Based Country Assistance Strategies provide the framework to better link country goals with Bank operations. They represent a strengthening of the traditional CASs, which set broad goals but did not clearly articulate the outcomes the Bank or the country wanted to achieve at the end of the

CAS period. A review of CASs conducted for the 2005 AROE found that most attempted to link country development goals with CAS outcomes and Bank interventions. These linkages were better developed in education, health, and the environment than in other disciplines. These particular CASs included performance measures to track progress toward achieving goals and outcomes. However, many of these measures lacked baselines, specific targets, or both, limiting their effectiveness for monitoring, management, and evaluation.

A greater results focus leads to shared accountability for development results between the government and the Bank. A survey of country directors conducted for the 2005 AROE found that only one-third have developed a common system with the government to develop information and jointly monitor progress toward CAS objectives. The reason often noted for the absence of joint systems is a lack of country capacity or a lack of interest in performance measurement, a finding confirmed

by other IEG studies. This indicates that more progress is needed in this area.

The Bank is supporting countries in developing performance measurement through Bankwide efforts such as the Statistical Capacity and the Evaluation Capacity Development Programs and through country-specific activities as part of its country programs. It is important that capacity development be conducted systematically, in collaboration with development partners, to ensure that key constraints are identified and addressed. The review of the CASs conducted for the 2005 AROE found that the CASs include references to monitoring and evaluation, but that monitoring and evaluation are not being approached in a systematic or strategic manner. The Bank should assess the country's performance measurement capacity and identify actions needed to strengthen it as part of CAS preparation to ensure consistency with priority country needs.

To achieve results, information on performance has to be useful to both government and Bank country managers. Country directors surveyed for this AROE report that improving country-level management to achieve greater results is the main use of information at the country level. Within countries there is often a perception that monitoring and evaluation information is collected to meet external reporting requirements.

The Country Portfolio Performance Reviews (CPPRs) could be made into a tool for the government and the Bank to jointly manage performance and used to develop an understanding of the use of performance information. The 2005 AROE found several instances where the Bank used CPPRs to increase awareness of results and identified possible uses of performance information to improve operational outcomes. The CPPR could help clarify and strengthen the links between CAS outcomes and individual Bank operations, help identify possible areas of synergy, and better align the results of individual Bank operations with overall country results.

Closing the Gaps to Increase a Results Focus in Monitoring and Evaluation Systems

The Bank is making progress in establishing guidelines to strengthen the results focus of its monitoring and evaluation at the country, sector, and instrument levels. At the project level, monitoring and reporting on results has been strengthened with the revised Project Appraisal Document (PAD), which now outlines who is responsible for producing the required information for monitoring implementation progress. The Project Supervision Report has been replaced by the Implementation Status and Results Report (ISR), which gives more prominence to results. Outcome tracking systems have been established for the Bank's economic and sector work and technical assistance and for trust fund-supported activities, but these systems are not yet being widely used.

At the sector level, the initial Sector Strategy Implementation Update (SSIU) replaced individual sector strategy updates and documented progress in all sectors. Remaining challenges include developing results frameworks for each sector and theme, integrating them into the Bank's operational work, and reconciling tensions between country priorities and global sector priorities. This reaffirms the 2004 AROE, which identified sectoral fragmentation and unclear accountabilities in the Bank's matrix structure as a major constraint to a stronger results orientation in the Bank.

At the corporate level, the Millennium Development Goals (MDGs) and the IDA14 performance measures provide an overall framework for monitoring results Bankwide. The Regions select from the MDGs, IDA14 country outcome indicators, and the sectoral indicators developed under the Implementation Follow-Up exercise when determining key performance indicators for final Regional outcomes in their Strategy and Performance Contracts. Also, the Bank is designing a Results Reporting System to capture and disseminate information for corporate learning and accountability.

IEG's Role in Increasing the Results Focus in Monitoring and Evaluation

IEG has strengthened its attention to results-oriented monitoring and evaluation in its evaluations at the country, sector, and project levels, and this work will need to continue and deepen. IEG is also working to improve awareness of its products. The results of the annual IEG client survey for 2005 show that 45 percent of Bank staff who responded to the survey were aware of the evaluation for which they were surveyed, a marked increase over the awareness levels found in previous AROE surveys (2004: 27 percent). There is an ongoing need for IEG to pursue improvements in the communication and dissemination of its evaluation products within the Bank and among its external clients and partners. The new Outreach and Communication Strategy launched by IEG will be tracked by key performance indicators. Lessons learned during the pilot phase will inform the full implementation of the strategy.

Recommendations

For Management

- Incorporate in the CAS a diagnosis of the country's performance measurement capacity and recommendations for action to strengthen that capacity. The diagnosis and actions should be customized to country capacity and focused on critical information needed for the government to implement its program.
- Expand use of the Results-Based CAS as a tool to manage country programs and strengthen the link between the Results-Based CAS and the country program by providing country teams with guidance designed to increase the results orientation of CPPRs.

For 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 communication strategy.



Nils Fostvedt

Acting Director-General, Operations Evaluation

ACRONYMS AND ABBREVIATIONS

AAA	Analytical and advisory activities
ACS	Activity Completion Summary
AIS	Activity Initiation Summary
AUS	Activity Update Summary
AROE	Annual Report on Operations Evaluation
ARPP	Annual Review of Portfolio Performance
BP	Business Process (World Bank)
CAE	Country Assistance Evaluation
CAS	Country Assistance Strategy
CCR	Country Assistance Strategy Completion Report
CFP	Concessional Finance and Global Partnerships
CODE	Committee on Development Effectiveness
CPPR	Country Portfolio Performance Review
DEC	Development Economics/Chief Economist
DECDG	Development Economics/Chief Economist Development Data Group
DFID	Department for International Development (U.K.)
DGF	Development Grant Facility
DIME	Development Impact Evaluation Initiative
DPL	Development Policy Loan
ECD	Evaluation capacity development
ED	Executive director
ESSD	Environmentally and Socially Sustainable Development (World Bank network)
ESW	Economic and sector work
GCR	Grant Completion Report
GPP	Global programs and partnerships
ICRs	Implementation Completion Reports
ICM	Implementation Completion Memorandum
IDA	International Development Association
IDF	Institutional Development Fund
IEG	Independent Evaluation Group
IFU-7	Implementation Follow-Up 7
IPDET	International Program for Development Evaluation Training
ISR	Implementation Status and Results Report
KPI	Key performance indicators
MAR	Management Action Record
MDGs	Millennium Development Goals
M&E	Monitoring and evaluation
OP	Operational Process (World Bank)
OPCS	Operational Policy and Country Services
PAD	Project Appraisal Document
PPAR	Project Performance Assessment Review

PRSC	Poverty Reduction Support Credit
PRSP	Poverty Reduction Strategy Paper
PSR	Project Status Report
QAG	Quality Assurance Group
RBCAS	Results-Based Country Assistance Strategy
RMS	Results measurement system
RRS	Results reporting system
SAP	Systems, Applications, and Products in Data Processing
SPC	Strategy and Performance Contract
SSIU	Sector Strategy Implementation Update
SSP	Sector Strategy Papers
TA	Technical assistance
TF	Trust funds
TFO	Trust Fund Operations
VPU	Vice presidential unit

Chapter 1: Evaluation Highlights

- The 2005 AROE focuses on the country and examines the use of information by Bank managers to improve results and enhance Bank effectiveness at the country level.
- It examines three elements that are critical to a strong results focus: strategic planning, performance measurement, and performance management.
- It reviews the actions taken by the Bank to strengthen the results orientation in M&E at the country, sector, and product levels.
- It evaluates IEG's effectiveness.



Introduction

The *Annual Report on Operations Evaluation* (AROE) fulfills IEG's mandate to assess the progress, status, and prospects for monitoring and evaluating the development effectiveness of World Bank activities. The objective of the 2005 AROE is to assess the use and usefulness of monitoring and evaluation (M&E) information for Bank country program management.

Relationship of the 2005 AROE to Previous AROEs

The 2003 and 2004 AROEs reviewed the Bank's Managing for Results initiative and the results focus in M&E. The 2003 AROE analyzed M&E at the instrument, country, sector, and global levels and commented on its overall effectiveness. The 2004 AROE examined how results-oriented the Bank's M&E systems are, the extent to which they contribute to managing for results in the Bank, and the organizational changes and incentives needed to move toward managing for results.

Importance of a Country Focus

The 2005 AROE examines the use of information by Bank managers to improve results and enhance Bank effectiveness at the country level. The country has supplanted the project as the focus of development over the past few decades. In the Bank, the country has become the principal unit of account for performance M&E. The Millennium Development Goals (MDGs) have highlighted the importance of country-level

results. The Bank's research on aid effectiveness and the Monterrey consensus have emphasized the importance of developing country ownership of sound policies and governance.

Country-level management also is important for achieving results. An IEG review of Country Assistance Evaluations (CAEs) found that one-third of evaluated country assistance programs were unsuccessful (IEG 2004c). While project outcomes may have been rated satisfactory, they did not always add up to a successful country program. The review concluded that a more comprehensive view of the outcome of the Bank's assistance programs is possible from the country level. The Quality Assurance Group has similarly found that "the quality, and hence the impact, of the country AAA [analytical and advisory activities] programs reviewed are often less than the sum of their parts" (World Bank 2005c).

The Bank's results initiative agenda was launched in September 2002 in the Development Committee's paper "Better Measuring, Monitoring, and Managing for Results." The country program cycle was identified as the

focal point for formulating, monitoring, and achieving greater development effectiveness. In contrast to a project-by-project approach, analysis at the country level captures the broader context of a country and the political economy of reform. Barriers to development can be better understood, and interventions designed accordingly. Moreover, the Bank's effectiveness will then be judged by the results the countries achieve with Bank support.

Overview of the 2005 AROE

The 2005 AROE has three parts. First, it examines three elements critical to a strong results focus at the country level: strategic planning, performance measurement, and performance management. It reviews the steps the Bank has taken in these three areas. To determine current practices, a review was conducted of Country Assistance Strategies (CASs) submitted to the Board in fiscal years 2004 and 2005. This was supplemented with a survey of country directors. Second, the AROE catalogues the measures taken to strengthen the results orientation in M&E at the country,

sector, and product levels within the Bank, updating the overview in the 2003 and 2004 AROEs. Third, it evaluates IEG's effectiveness. It concludes with a set of recommendations for Bank management and IEG.

Chapter 2 assesses strategic planning, performance measurement, and performance management at the country level and identifies areas that need strengthening. The analysis done for the assessment used a review of the CASs (Appendix A) and a survey of country directors (Appendix B). Chapter 3 reports on changes made since the 2003 AROE to enhance M&E and identifies areas where it needs further strengthening. Chapter 4 examines IEG effectiveness and reports on its response to recommendations made in the 2004 AROE. It also presents highlights of the results of the annual IEG client survey (Appendix D). Chapter 5 summarizes conclusions and presents recommendations for Bank management and IEG. In addition to the sources already mentioned, this report draws on reports from Operations Policy and Country Services (OPCS), the Quality Assurance Group (QAG), and on IEG evaluations.

Chapter 2: Evaluation Highlights

- Results-Based CASs allow a better articulation of the contribution Bank operations make to the achievement of country goals.
- While performance measures are included to track progress, improvement is needed in developing baselines and targeted performance measures.
- The Bank will need to deepen its work with clients to strengthen their monitoring and evaluation systems.
- As the Bank increases its results orientation, the focus of performance management will shift from the Bank to the country and toward shared accountability for results.
- Successful performance management requires creating demand for performance information in countries.



Measuring and Managing Performance at the Country Level

Three elements are essential for effective results-based management at the country level: strategic planning, performance measurement, and performance management (see box 2.1).

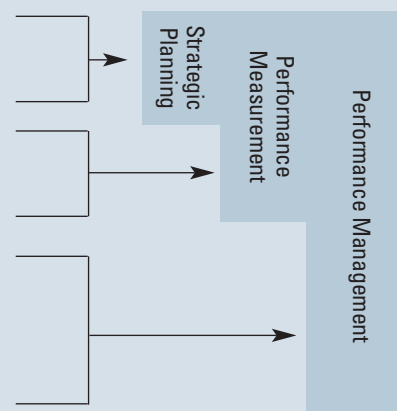
Strategic planning is needed to align the Bank's products and services with desired country outcomes to maximize their effectiveness. The strategy needs to be supported by a system to monitor progress and evaluate effectiveness in achieving country outcomes. Finally, the information obtained from M&E needs to be useful to Bank managers for making operational decisions.

To see how these three elements fit together in Bank operations, two reviews were conducted as part of this AROE. In the first review, the results matrices of 24 CASs presented to the Board in fiscal years 2004 and 2005 were analyzed to assess their usefulness as a management tool and their impact

Results matrices of 24 CASs were analyzed to assess their usefulness as a management tool.

Box 2.1: Elements of Results-Based Management

- Identify clear and measurable objectives
- Select indicators that will be used to measure progress toward each objective
- Set explicit targets for each indicator, used to judge performance
- Develop performance monitoring systems to regularly collect data on actual results
- Review, analyze, and report actual results vis-à-vis the targets
- Integrate evaluations to complement performance information not readily available from performance monitoring systems
- Use performance information for internal management accountability, learning, and decision making processes, and also for external performance reporting to stakeholders and partners



Source: Binnendijk 2001.

on enabling the CAS to become more results-based. The review examined the strength of the linkages between country development goals, CAS or intermediate outcomes, and Bank programming, which together constitute the results framework. The performance measures were examined to determine whether they were clearly defined, baselined, and targeted and whether they provided information that would be useful in managing the country program. The method and results are presented in Appendix A.

For the second review, country directors who had presented CASs to the Board in fiscal years 2004 and 2005 were surveyed to learn how they use information from M&E systems.¹ The survey responses indicate the degree to which

Country directors who had presented CASs to the Board were surveyed to learn how they use information from Bank M&E systems.

the Bank's current M&E systems enable them to effectively manage their country operations. The complete results of the survey and the method are presented in Appendix B.

Strategic Planning

The Bank is making progress in strengthening country-level strategic planning, yet improvement is needed in developing baselines and targeted performance measures.

A management review of Bank CASs (World Bank, OCPS 2005a) concluded that it was difficult to use the CAS to assess Bank effectiveness. The CAS policy and program matrix that underpin the self-evaluation framework set broad goals, but did not clearly articulate the outcomes that the Bank or country wanted to achieve by the end of the CAS period. Nor did it clearly differentiate between long-term country goals and goals to which the Bank program is designed to contribute. It also lacked baseline data and targets to monitor progress and assess whether the intended results were achieved.

The Bank introduced the Results-Based CAS (described in box 2.2) to sharpen the strategic focus of the CAS and strengthen it as a planning, management, and evaluation tool.

The results-based approach has been taken up by CASs other than the seven pilots. Of the 24 CASs reviewed for this report, 13 had adopted a results-based approach.

The Results-Based CAS allows a better articulation of the contribution Bank operations make to the achievement of country goals. The Results Matrix requires the Bank to distinguish and then link country development goals, CAS and intermediate outcomes, and Bank operations. Of the 24 CASs reviewed for the AROE, 19 attempted to establish such causal linkages.² The review concluded that the Results Matrix makes it easier to see the contribution of Bank operations to the CAS and to country goals.

Countries that had Results-Based CAS pilots or that had gone through a Poverty Reduction Strategy Paper (PRSP) exercise were able to formulate improved linkages.

The understanding of results chains and use of the matrix could be strengthened, however. For example, some of the weaker CAS matrices simply listed goals, measures, and programs under pillar headings (the pillars: strengthening country capacity, enhancing the results focus in operations, and fostering a global partnership to manage for results). This approach makes it very difficult to understand the assumed interactions between actions and results and to assess the contributions of particular programs to results.

Lack of baselines and targeted performance measures limits the effectiveness of the Results Matrix as a monitoring, management, and evaluation tool. A qualitative assessment of the performance measures for country, CAS, and Bank intermediate outcomes was done to assess how well they were defined, whether baselines were provided, and whether they included targets. The country, CAS, and Bank intermediate performance measures for each CAS were sorted into four categories, with a higher-numbered category indicating that the measures were better defined, baselined, or targeted. For example, performance measures in one CAS may be rated Level 3 for definition if more than 75 percent are well defined, Level 2 for baselines if more than half but less than 75

Box 2.2: What Is a Results-Based CAS?

A **Results-Based Country Assistance Strategy** is oriented toward achieving realistic outcomes and includes a results-oriented M&E system. Its central innovation is the design of a results framework that specifies expected linkages between the Bank's interventions and long-term development goals. A greater focus on results, in turn, affects the CAS's vision, diagnosis, and programming. The features of the new CAS framework are:

- More explicitly defined outcomes that establish clear links between the two ends of the results chain: development goals of the country and Bank-supported interventions
- A stronger alignment of the Bank's program with the country's vision for economic and social development and poverty reduction, and development goals and priorities that derive from this vision

- An M&E framework for learning and self-evaluation that uses existing country's systems (to the extent possible) to steer implementation toward outcomes.

Results are selected during CAS design, tracked during implementation, and evaluated at the end of the CAS period. The building blocks of the design include three main components: (a) a **Results Matrix** that summarizes the results framework, providing the links upward from CAS outcomes to the broad development goals of the CAS, and downward to the proposed Bank programming; (b) a **CAS Progress Report** for mid-point stocktaking and, as necessary, course correction; and (c) self-evaluation of the country assistance program in a **CAS Completion Report**. IEG conducts a review of the CAS Completion Report.

Source: World Bank, OCPS 2005a.

percent have baselines, and Level 1 for targeting if less than half have clear targets.³ The methods and results of the review are described in Appendix A. Figure 2.1 presents the combined results of the evaluation for definition, baselining, and targeting. It shows that the performance measures were generally adequately defined, but they often lacked baselines and targets.⁴ The review concluded that 54 percent of the country, CAS, and intermediate measures combined were considered to be well defined (at Level 3), while only 18 percent were considered well targeted, and 13 percent adequately baselined. At the same time, 24 percent of the measures were rated poor (Level 0) for targeting and 28 percent for baselines, but only 15 percent of the measures were considered to be poorly defined.

Results-based approaches and performance measures are better developed in some areas than in others, which may explain some of the disparity. Performance measures generally were better developed at the country level, where more established measures (such as the MDGs) are available. Countries with a PRSP had an advantage because PRSPs included results monitoring at the country level. The review also noted that results chains and performance measures were better developed in education,

health, and the environment than elsewhere.

That is, the causal linkages between country outcomes, Bank outcomes, and Bank activities were stronger in these sectors, so the CASs were able to present a more compelling, results-oriented narrative through the Results Matrix.

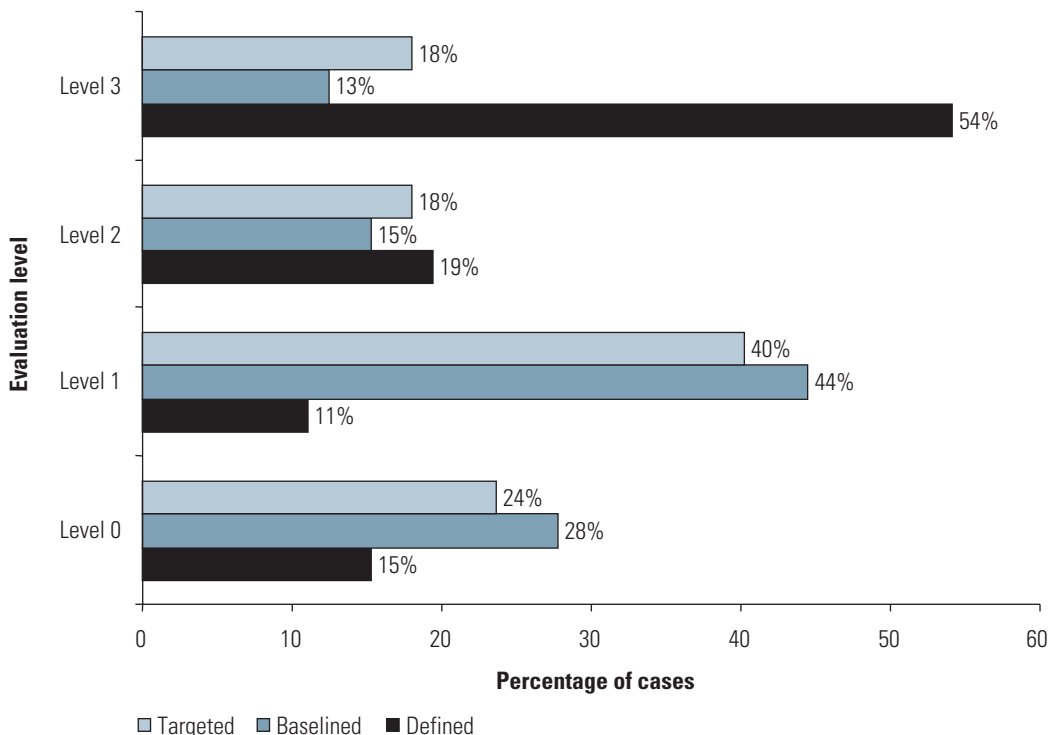
Other studies have also found a stronger results focus in education and health. In reviewing projects under implementation, QAG (World Bank 2005d, Annex 9) noted that projects in health and education, along with those aimed at poverty reduction, took a more systemic approach to monitoring and evaluation, often using broad national outcome indicators and relying on the monitoring of project performance indicators under another project or a government-managed evaluation program.

IEG's CAE Retrospective (IEG 2005a) also found that Bank country programs were more successful in education and health than in

Lack of baselines and targeted performance measures limit the effectiveness of the Results Matrix as a monitoring, management, and evaluation tool.

Results chains and performance measures were better developed in education, health, and the environment.

Figure 2.1: Distribution of Performance Measures by Evaluation Level (percent)



other sectors. A possible explanation is that these sectors have had a stronger focus on outcomes. For example, educators have been concerned about the impact of their programs on student learning, and health professionals are focused on the improvement in the health of patients.

The large number of measures may have exacerbated the problem of baselining and targeting. The average number of country performance measures was 22; CAS outcome measures, 39; and intermediate outcome

The Bank needs to be selective in determining which measures to monitor, especially in view of data constraints in countries.

measures, 40. One CAS had 80 outcome measures and 97 intermediate measures. The extent of Bank involvement and the size of the program would drive the number of measures needed to manage the program. It is not possible to attribute the lack of baselines and targets directly to large numbers because

many CASs, even those with smaller numbers of measures, lacked adequate baselines and targets. However, larger numbers of measures would increase the challenge for measurement and determination of baselines.

While the number of measures should reflect the scope of the Bank’s country program, the Bank needs to be selective in determining which measures to monitor, especially in view of the data constraints within the country. A large set of measures may be needed for evaluation at the end of the CAS period, but a smaller set could be used to monitor and manage the country program. Some CASs summarize the key performance measures, establishing a priority among them. A more systematic approach will be to establish the information necessary to manage the country program, and ensure that the measures include baselines and targets. Some of these measures can be qualitative, such as policy or administrative milestones.

Country ownership of monitoring data is important. The country's capacity to collect and analyze the information should be considered in establishing the performance measures. Availability of data, however, should not be a main reason for selecting measures. It may be necessary to include indicators for which data collection systems will need to be developed during the CAS period. The link to country capacity is discussed in more detail in the following section.

Performance Measurement

To manage for results, the Bank will need to deepen its work with clients to strengthen their monitoring and evaluation systems and take a systematic approach to performance management at the country level.

The Bank has systems to help managers manage their country operations and needs to continue to make the information readily available. The Bank's internal information systems allow Bank managers and staff to monitor budget use, product deliveries, and the quality of the portfolio at the country level. In a survey of selected country directors carried out for this report (Annex B), 67 percent of respondents indicated that they were able to obtain the information needed from the Bank's information systems. Many (75 percent) packaged the information to better meet their needs by consolidating data from different Bank systems or by including outcome information obtained from country sources. Several country directors expressed a need for more systematic assembly of country information for program management. The Bank is working to improve management of country-level information; one proposed solution is a Country Director's Portal.⁵

A stronger focus on results will require the Bank to work closely with its clients to collect data regularly on the outcomes of Bank-supported operations and to adjust the way it monitors country program implementation. At the project level, the emphasis on outcomes requires the implementing agency to collect information on the impact of its services and products on benefi-

ciaries. These data collection efforts could become more elaborate, time consuming, and expensive in the future. For example, an agency may rely on its operating units to provide information on outputs such as kilometers of roads built or training programs conducted, but may need to conduct surveys to collect information on beneficiary impact. Similarly, the move toward monitoring higher-level country outcomes shifts the focus of data collection from units within agencies to agency-wide or national systems. While strengthening the results focus and project-level monitoring and evaluation systems continues to be important, more attention would need to be given to strengthening performance measurement at the sector or agency level and at the national level.

The Bank will need to work closely with country and development partners to coordinate information needs and performance measurement at the country level. Country data, such as the MDGs, will be used by the government, other donors and development partners, and the Bank to monitor progress toward achieving country strategic goals. To be effective, performance measurement systems need to be country-owned and used, focused on essential information, and harmonized between the country and other stakeholders, including other donors, multilateral institutions, and the Bank.

A multi-partner evaluation of the Comprehensive Development Framework (CDF Evaluation Management Group 2003) noted that "development performance should be evaluated through measurable, on-the-ground results" and recommended that all donors and others "strengthen and use country-led M&E systems, view country development outcomes

A stronger focus on results will require the Bank to work closely with its clients to collect data regularly on the outcomes of Bank-supported operations.

The Bank will need to work closely with country and development partners to coordinate information needs and performance measures at the country level.

as a joint product, to which donors contribute under country leadership,” and see results monitoring as “a shared responsibility in which donors and governments use the same information flows for their individual purposes.”

The development community supports greater collaboration and country ownership in performance measurement. Following the Second International Roundtable on Managing for Results in Marrakech in 2004, the leaders of the development agencies stressed the need to “align cooperation programs with desired country results, define the expected contribution of our support programs to country outcomes, and rely on and strengthen countries’ monitoring and evaluation systems to track progress and assess outcomes” (African Development Bank and others 2004). The Paris Declaration on Aid Effectiveness issued in March 2005 reaffirmed the commitment of the countries, donors, and multilateral institutions to country ownership and coordination around country-led information systems (box 2.3).

The survey of country directors found that 87 percent of those responding regularly

discuss progress toward CAS objectives, along with operational matters such as portfolio status and readiness of projects under preparation, with their government counterparts. These discussions are conducted as part of regular consultations and program and portfolio reviews. In low-income countries, PRSP and Poverty Reduction Support Credit reviews are important vehicles for these discussions.

But most survey respondents do not yet have a common system or process to develop and share results information with the government. Fewer than one-third (31 percent) of the responding country directors indicated that they have a system or process in common with the government to jointly develop and share information for monitoring progress toward meeting CAS objectives. Most existing systems and processes were centered on the PRSP.

Several country directors indicated that they were currently developing a joint system with the government through an Institutional Development Fund (IDF) grant or other means. Lack of government capacity or interest was often cited as the reason for not developing a joint monitoring system. Lack of interest

Box 2.3: Harmonization Around Results

The ministers from developing and developed countries and the heads of multilateral and bilateral development institutions met in Paris in February-March 2005 to discuss measures to improve aid effectiveness. One of the commitments from the meeting is to “managing for results.” Specifically:

- **Partner countries** commit to:
 - a. Strengthen linkages between national development strategies and annual and multi-annual budget processes.
 - b. Endeavor to establish results-oriented reporting and assessment frameworks that monitor progress against key dimensions of the national and sector development strategies, and that these frameworks should track a manageable number of indicators for which data are cost-effectively available.
 - c. Link country programming and resources to results
- **Partner countries and donors** jointly commit to:
 - d. Work with partner countries to rely, as far as possible, on partner countries’ results-oriented reporting and monitoring frameworks.
 - e. Harmonize their monitoring and reporting requirements, and—until they can rely more extensively on partner countries’ statistical, monitoring, and evaluation systems—partner with countries to the maximum extent possible on joint formats for periodic reporting.
 - f. Work together in a participatory approach to strengthen country capacities and the demand for results-based management.

Source: Paris Declaration on Aid Effectiveness, March 2, 2005.

was more prominent in middle-income countries where there is no PRSP equivalent to focus the country, donors, and the Bank on results.⁶ Several country directors mentioned the time and effort required to develop joint systems with the government.⁷

IEG reviews have also identified lack of capacity for data collection and limited country demand for performance information as constraints to developing country performance measurement systems. The multi-partner evaluation of the Comprehensive Development Framework (CDF Management Group 2003) noted that most government systems were ill-suited to the often “elaborate and complex monitoring approaches” donors required of them.

An IEG review of the Poverty Reduction Strategy Initiative (IEG 2004a) noted that results orientation was one of the weaker areas of the process, although “results orientation with a focus on outcomes that benefit the poor” was one of the five underlying principles of the PRSPs. Establishment of monitoring systems was constrained by lack of capacity for M&E in the countries. A results orientation was impeded in some cases by a decision to track a large number of indicators for which too little data was available. Efforts to build monitoring capacity and to monitor results were seen by governments in many countries as a requirement imposed from outside. The study noted that countries need to select and monitor relevant indicators that

can support domestic decision making (box 2.4).

A similar observation was made in an IEG assessment of evaluation capacity development efforts (IEG 2004b).

Following in-depth reviews in two countries with extensive Bank involvement, the report observed that weak demand could result in weak systems and that greater attention should have been paid to sector ministries. These ministries are responsible for service delivery and could use the information and become the focus of M&E.

IEG has also cautioned that harmonization may take time and effort. An overview of IEG’s work reconfirmed the importance of harmonization in reducing transaction costs and enhancing development effectiveness. IEG noted that successful harmonization hinges on country capacity and government commitment, especially where changing “the existing way of doing business” could conflict with different donors’ and agencies’ own interests. It also cautioned that harmonization can have significant up-front costs and prove frustratingly slow relative to expectations. The lessons are summarized in box 2.5.

Most country directors surveyed do not have a system or process to develop and share results information with the government.

Lack of both capacity for data collection and demand for performance information are constraints to developing country performance measurement systems.

Box 2.4: Observations from IEG’s Review of the Poverty Reduction Strategy Initiative

Added attention to monitoring, but results focus remains weak.

The Poverty Reduction Strategy (PRS) process has begun to orient stakeholders toward a results focus, although the development of country-specific indicators and monitoring systems to track them is still at a preliminary stage in most countries. Available information is generally not linked to decision making.

Some countries have adopted PRSP targets that are unrealistic in light of their initial conditions, resources, and recent expe-

rience. In many cases, monitoring systems are designed to meet donor data requirements—for example, through the progress reports—even when they exceed the country’s needs or capacity.

Since the majority of low-income countries will probably not be able to meet most of the Millennium Development Goals, PRSPs must be more effective in enabling countries to select and monitor relevant indicators that can support domestic decision making on a sustained basis.

Source: Extract from IEG, 2004a, p. xv.

The Bank is working with countries to strengthen their M&E systems.

The Bank is working with countries to strengthen their M&E systems. The Development Economics Department's Development Data Group (DECDG) manages the global Trust Fund for Statistical Capacity Building, which helps countries to improve the collection, processing, analysis, dissemination, and use of statistics, especially to support poverty reduction. DECDG has designed a new lending program, the Statistical Capacity Program, to support the development of more effective and efficient statistical systems in developing countries. The program is under way in Burkina Faso and Ukraine, and preparations have started in four additional countries.⁸ DECDG will take the lead in improving statistical capacity in International Development Association (IDA) countries under IDA14.

In addition, IEG has been engaged in evaluation capacity development (ECD), helping countries to develop their M&E systems through technical assistance and training (IEG 2004b). IEG has provided ECD support to 34 countries since 1999 and has engaged intensively in 2 countries, Uganda and Egypt, and to a lesser extent in 15 others, through diagnostics, organization of workshops and other training, and support of local evaluation associations. The Uganda experience is described in box 2.6.

IEG has supported another 17 countries, primarily through knowledge sharing (provision of materials, identification of experts, and the like) and has prepared and widely disseminated case studies, manuals, and other resource materials on ECD. Finally, over 630 participants from 62 countries have attended the International Program for Development Evaluation Training (IPDET), a four-week course developed by IEG and Carleton University. The IEG self-evaluation of ECD concluded that "IEG has been highly active in ECD, and the ECD strategy has contributed to a number of ECD results, outputs and outcomes, over the past five years" (IEG 2004b, p. ii).

Third, the Bank is supporting efforts to strengthen M&E at the country, sector, and project levels through use of Institutional Development Fund grants, coordination with other donors, and AAA. These efforts are led by the country departments. Examples of some recent innovative approaches where the Bank is focusing on country capacity development at a national level are described in box 2.6.

Treatment of performance measurement in CASs is relatively weak and inconsistent. Most of the 24 CASs reviewed for this report included references to M&E, but many did not provide much detail. Only a few discussed data availability and contained proposals to strengthen monitoring and evaluation of the

Box 2.5: IEG Lessons on Harmonization from Evaluation

IEG's evaluations have demonstrated that proliferation of uncoordinated procedures and requirements across aid agencies is imposing substantial costs on both donors and recipients and limits the overall effectiveness of aid. The evaluations have so far found only modest progress in the harmonization of aid processes and programs. Nevertheless, the evaluations support the following lessons about harmonization:

- Harmonization can be improved through mechanisms such as Consultative Groups, Sector-Wide Approaches, and

Poverty Reduction Strategies.

- Where successful, harmonization has required capacity for and commitment to reform on the part of both donor and recipient countries.
- Harmonization can have significant up-front costs, as has been the case with the Poverty Reduction Strategy Initiative.
- Harmonization among donors may not always align donor strategies with country development goals.
- Changing donor procedural requirements has proved to be slow and difficult.

Source: IEG documentation.

Box 2.6: Developing the Capacity to Manage by Results at the Country Level

- **Evaluation Capacity Development in Uganda:** At the request of the Bank's country director and Ugandan authorities, IEG started providing evaluation capacity development (ECD) support to Uganda in 2000. The work included a government-wide diagnosis of M&E and a study of M&E in the education, health, and water sectors and workshops to disseminate M&E practices and experiences.

Thirty Ugandan officials were given scholarships to attend the International Program for Development Evaluation Training (IPDET). IEG also provided support to the fledgling Uganda Evaluation Association. IEG's work was linked closely to the reform agenda under the Poverty Reduction Support Credit (PRSC).

The government decided in 2004 to implement a National Integrated Monitoring and Evaluation system based in substantial part on IEG input. IEG partnership agreements with Norway and the Netherlands helped fund this work and enabled IEG to scale up its support.

An IEG review of the experience concluded that, despite insufficient attention paid to bureaucratic realities and to sector ministries, on balance, the impact on evaluation capacity in Uganda had been positive and significant. It also commented that the most important impact was that "the lip service given in 2000 to [M&E] has become four years later genuine interest, and awareness of the importance of results and of outward accountability has percolated through Uganda's public sector and much of society at large."

- **Brazil Multi-Sectoral Program to Support Evaluation (BRAVA):** M&E is challenging in Brazil because the country has neither a strong mandate for evaluations nor methodological frameworks for program assessments. The Bank-supported BRAVA

Program is intended to help the government develop a policy framework for comprehensive program evaluation and conduct a series of comprehensive evaluations to pilot the framework, most of them in Bank-supported programs in social development, health, education, agrarian development, housing and urban development, and micro-credit.

The policy framework would establish incentives for line ministries to comprehensively evaluate their programs, put forward a methodological framework and guidelines for conducting comprehensive evaluations, and specify the institutional roles needed to implement a results-based evaluation system within the federal government. The program is expected to promote a culture of evaluation and complement the Bank's activities to strengthen results orientation in the government through the Results-Based CAS and M&E in individual operations.

- **Training of Trainers in Morocco:** Discussions are under way in Morocco for a program for training of trainers in results-based project management and M&E, with support from an Institutional Development Fund grant. This follows a results-based Country Portfolio Performance Review that developed a common understanding of each operation's relationship to country goals, results matrix, performance indicators, and M&E requirements.

The training will be done at a local institute that is expected to develop expertise and continue such training beyond the period of the grant. The expected outcome of the program is a network of resource persons with strong knowledge of results-based management, M&E, and fiduciary procedures (especially as they relate to improving project outcomes) who will be given the responsibility to respond to training needs in the country in the same areas of focus.

Source: IEG 2004b.

CAS during implementation.⁹ While the Bank is taking measures to strengthen performance measurement at the country level, it is not clear whether this is being done by identifying priorities, identifying gaps, and factoring appropriate support into the Bank country program.

Preparation of the Results-Based CAS should take a systemic approach to performance measurement at the country level by taking stock of the M&E information needs; identifying capacity development needs; and building in measures, where appropriate, to develop

country capacity. Bank CAS guidelines as of March 2004 mention that "as appropriate, the CAS should discuss efforts to develop monitoring and evaluation capacity in borrowing countries as well as efforts at the country level to harmonize monitoring and evaluation approaches and requirements among different partners"¹⁰ (World Bank, OCPS 2004b,

Preparation of Results-Based CASs should take a systemic approach to country-level performance measurement.

p. 7). These guidelines are being revised to mainstream the experience from the pilot Results-Based CASs and propagate good practices. As part of the revision, country teams could be required to undertake a diagnosis or Readiness Assessment to identify areas where the Bank and the government could work together to ensure that performance measurement and management are addressed in a timely manner.

Use of the information for operational performance enhancement and achievement of agreed outcomes should be the driving force behind development of performance measurement systems. The sources of information, the data to be collected, responsibility for data collection and analysis, and capacity development needs and proposed support could be summarized in a Monitoring and Evaluation Action Plan. A good practice note should be prepared based on ongoing country efforts to encourage and assist the country teams with the Readiness Assessments and action planning.

Performance Management

As the Bank increases its results orientation, the focus of performance management will shift from the Bank to the country and toward shared accountability for results. Establishing the use of, and creating demand for, performance information will be critical for successful performance management.

Performance information is meaningless if it is not used. An OECD/DAC report identified two categories of use for performance information: (a) management improvement (managing for results) and (b) performance reporting (accountability for results). These concepts are

described in box 2.7. The former is driven by internal organizational concerns, the latter by external stakeholders.

Within the Bank, management improvement is seen as the main use of information at the country level, but the

need to report on performance at the country level is expected to increase. Within the country, there is often a perception that monitoring and evaluation is driven by external concerns (see box 2.4). However, there is a growing need to make performance information part of overall decision making. This will occur only if there is demand for performance information within the government.

Many country directors who responded to the survey carried out for this report see learning and management improvement as the key uses of information.¹¹ They identified learning, business planning, and modifying operations as the most important uses of information (figure 2.2). Accountability was considered less important. Several commented on the difficulty of determining the extent of the Bank's contribution to an outcome as a reason for giving accountability a lower rating.

Bankwide reporting on results may increase the pressure on country directors, sector directors, and managers to focus on performance reporting for accountability for results. The Strategy and Performance Contracts that are being introduced in fiscal 2006 will provide the Regions with a list of indicators from which to select key performance indicators for final outcomes.¹² While the contracts will be at the Regional level, the actual results are achieved in the countries. Some links will need to be established between the Regional indicators and the countries, most likely through the Results-Based CASs. The experience with PRSPs has highlighted some of the challenges: balancing country-specific information needs and Bankwide monitoring and reporting, ensuring that the information will be useful for country decision makers, and keeping the level of information to be collected manageable and consistent with country capacity.

As the Bank increases its results orientation, the focus on performance management will shift from the Bank to the country and toward shared accountability for results. The Bank has a responsibility to deliver loans and credits and to provide analysis and advice through studies and technical assistance, as agreed with the government. The government, in turn, will

Performance information needs to be part of overall decision making.

This will occur only if there is demand for performance information within the government.

Box 2.7: How Performance Information Is Used

An OECD/DAC review of results-based management in multi-lateral and bilateral aid agencies distinguished two uses of performance information:

- **Management improvement (managing for results):** Performance information: (a) promotes continuous management *learning* about what results are being achieved and why and (b) facilitates operational *decision making* (implementation, strategy and policy formulation, resource allocation, and the like).
- **Performance reporting (accountability for results):** Performance information is used to *report* performance and fulfill *accountability* to external stakeholders and internal management.

Source: Binnendijk 2001.

The OECD/DAC review noted a potential for conflict between the two uses for performance information. As the focus moves up the results chain (from inputs, to outputs, to outcomes), problems of aggregation and attribution emerge. That is, it becomes more difficult to “add up” the results from lower levels (projects to country) and attribute changes to Bank operations (clearly demonstrate that the change was due to a Bank intervention).

The review suggested resolving this conflict by moving from a “command and control” view of accountability, holding managers accountable for achieving results, to shared accountability, where managers are held accountable for “working with partners to achieve higher-order results, for learning from failures, and for continually using performance information in their decision making.”

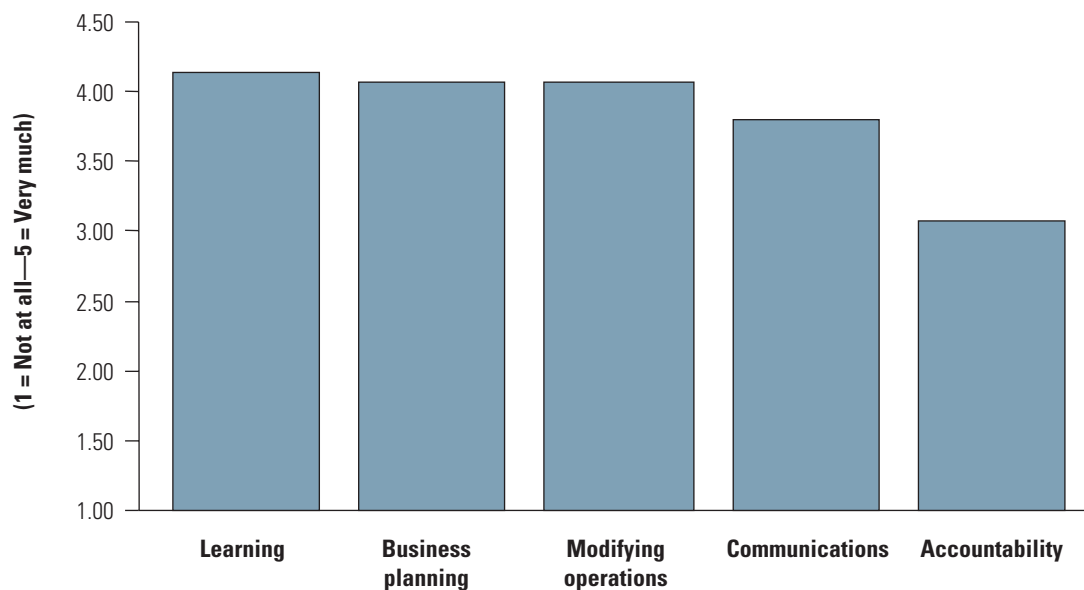
need to use these inputs from the Bank effectively and provide infrastructure, goods, facilities, and services to their beneficiaries. It is essential that the Bank generate demand for performance information and continue to foster the interest, political will, and capacity to manage for results within the government.

The Country Portfolio Performance Review (CPPR) can be made into a tool for the government and the Bank to jointly manage perform-

ance. Traditional CPPRs have been Bank-led, focused solely on the ongoing project portfolio, addressed primarily implementation problems, and were often a review of problem projects. The CPPR can be made more effective

As the Bank increases its results orientation, the focus on performance management will shift from the Bank to the country and toward shared accountability for results.

Figure 2.2: Important Uses of Information



by providing the country with a leadership role, expanding beyond ongoing projects to include AAA, and focusing the discussion around progress toward agreed results (country-level outcomes and CAS-level outcomes), rather than solely around implementation issues. Project administration issues can then be discussed in the context of overall progress toward agreed results and foster joint accountability.

The CPPR can be developed into a vehicle to conduct regular reviews of the overall Bank program.

An expanded CPPR would help clarify and strengthen the links between CAS outcomes and the outcomes of individual Bank instruments (such as projects and AAA), help identify possible areas of synergy across operations, and begin to address the disconnect between overall country and individual instrument results. It would complement the CAS Progress Reports and Completion Reports by generating information needed for country program management and by incorporating results into the discussion of the program. Finally, the anticipated increase in lending comes with the risk that the quality of operations may drop. A regular review of program performance and its implications for overall country results could ameliorate some of this concern.

Some countries are piloting this new approach (see box 2.8). For example, the Government of Malawi and the Bank agreed to take a results-oriented approach to Malawi's CPPR. While some participants initially focused on implementation issues, the review built an understanding and awareness of the importance of results. In the Middle East and North Africa Region, the Bank also used the CPPRs to raise awareness of the importance of results.

The CPPR is an established tool, and several country directors who responded to the survey

An M&E readiness assessment and action plan should be made part of the CAS preparation process.

reported that they use it as an occasion to review progress toward CAS objectives with the government. The CPPR can be developed into a vehicle to conduct regu-

lar reviews of the overall Bank program, rather than projects under implementation, and to assess program effectiveness in achieving CAS outcomes, and thereby generate demand for performance information.

Conclusions

The move to a results approach shifts the focus on performance measurement and management from within the Bank to the country. The shift to outcomes will require the Bank to work with the government to put in place systems to collect and analyze the necessary information. This should not become a burden for the government. The Bank will need to: (a) use government systems and help build capacity to develop and maintain such systems; (b) help to generate demand for, and use of, the information; (c) simplify and emphasize essential information; and (d) harmonize information systems across Bank operations, as well as with other development agencies.

The Bank should take a more systematic approach to performance measurement during CAS preparation. The Results-Based CASs (RBCASs) provide the strategic framework for managing for results at the country level, although the lack of baselines and targets for performance indicators makes them less effective as a monitoring, management, and evaluation tool. While the RBCAS is new and is only now being mainstreamed, the Bank needs to focus on strengthening baselines and targets. More attention needs to be given to monitoring performance during CAS implementation. The Bank could work with the government to:

- a. Determine the information needed to monitor progress toward country goals and manage the Bank country program.
- b. Set targets and establish baselines.
- c. Identify the sources of information.
- d. Establish responsibility for collection and analysis of the data.
- e. Determine areas where additional capacity building and assistance may be required and addressed through the CAS.

Box 2.8: Incorporating Results in Portfolio Reviews**The Malawi Country Program Review**

Preparation of the Country Program Review started in the fall of 2004 as a partnership with government. Both partners wanted to take a results-oriented approach, with the focus on outcomes and how the program was (or was not) positioned to support Malawi in achieving its development goals. Implementation issues (such as procurement or M&E) would be looked at to see how they might be supporting or hindering achievement of the goals and priorities.

The review was designed to initiate: (a) a continued dialogue on results; (b) a collaborative process to design the next CAS; and (c) preparation of the CAS Completion Report. Both parties agreed that an interactive, workshop-style format would be appropriate for an open exchange of perspectives between the government and the Bank and between technical and more senior levels of government.

During the week-long event, the level and depth of discussions around outcomes and results evolved. Initially some participants focused largely on implementation bottlenecks. As the focus on results caught on, the participants engaged in strategic discussion on outcomes, although some found the change in mindset challenging.

The event proved an excellent forum for the government to express some concerns over how to translate the president's vision into operational programs—while balancing the growth strategy with social policies. Senior and technical government officials and Bank staff exchanged views and offered solutions on (a) the alignment of Bank support to the government's development priorities; (b) identification of key institutional implementation issues;

Source: World Bank documentation.

and (c) next steps. Steering groups were set up to continue work on the strategy and implementation issues.

Results-Focused CPPRs in the Middle East and North Africa Region

In the Bank's Middle East and North Africa Region, Country Portfolio Performance Reviews (CPPRs) in Algeria, Morocco, and Tunisia focused on enhancing results on the ground by using a participatory approach in evaluating and strengthening the countries' M&E. The Region moved away from traditional CPPRs that focused on inputs and outputs and made outcomes its essential focus.

The discussions with the government started with development objectives and included results chains, performance indicators (measurable, realistic, relevant, precise, with baseline and targets), and capacity and risk (readiness of systems in place, who does what). The Region helped its counterparts achieve a better understanding of results-based management and a common understanding of each operation's results matrix and performance indicators, along with the extent to which objectives are related to the overall country's goals, on which CASs are based.

The CPPRs went well beyond diagnosis and reached agreements on action plans for sound M&E systems. To ensure sustainability of the efforts, the Region is working with the governments of Algeria and Morocco on building permanent in-country capacity in results-based management and in helping the countries link national strategies to budget process and public expenditure management through the use of Institutional Development Fund grants (box 2.6).

An M&E readiness assessment and action plan should be made part of the CAS preparation process.

The government and the Bank will be jointly accountable for results. Bank managers are becoming increasingly focused on managing for results, yet external pressure for the Bank to report on its results would most likely lead to increasing emphasis on accountability for results. While Bank managers will continue to be accountable for managing resources and delivering quality products to the country, the government and the Bank will be jointly accountable for outcomes. Performance information is essential for effective results-based manage-

ment, but the information has to be useful for both Bank and government managers for learning and decision making, as well as for reporting.

The Bank should expand the CPPR to include AAA and place greater emphasis on project and AAA outcomes and their linkages to CAS and country development goals, in addition to implementation issues. This would allow the Bank and the government to

The Bank should expand the CPPR to include AAA and place greater emphasis on project and AAA outcomes and their linkages to CAS and country development goals.

focus discussion on the performance of the Bank's overall program and contribution to achieving country development and Bank CAS objectives. An expanded results-focused CPPR would become the

tool for the government and the Bank to jointly manage performance and would operationalize the links between the CAS outcomes and project or sector-level outputs and outcomes.

Chapter 3: Evaluation Highlights

- The Bank continues to integrate a results focus in monitoring and evaluation systems.
- Country-level monitoring for results is gaining momentum.
- Monitoring of sector-level results is not as far along, but progress is being made.
- New lending guidelines and templates have been issued that require more systematic collection of data for monitoring of lending operations.
- Monitoring of analytic and advisory activity remains a challenge.
- Corporate-level monitoring is gaining ground.
- Country-level evaluation is becoming more developed and progress in sector-level evaluation needs to be deepened.
- Trust fund and global program evaluation are not fully effective.



Update on Monitoring and Evaluation of Bank Operations

This chapter updates the 2003 and 2004 AROE on the status of the Bank's work in revising and introducing guidelines and templates for incorporating a results focus in the Bank's M&E of its products at the country, sector, and instrument levels. It summarizes the key observations and recommendations from the previous reports and highlights steps taken to continue to integrate a results focus in M&E systems in the Bank.

Monitoring of Bank Operations

Country-Level Monitoring for Results Is Gaining Momentum

The CAS presents the Bank's proposed program of lending and nonlending activities to support the country's development vision and includes an M&E framework to gauge the results of the Bank's performance. The 2003 AROE noted that most CASs had not specified clear outcomes for the Bank, linked Bank interventions to expected outcomes, or defined performance indicators.

The Results-Based CASs address some of the concerns raised in the 2003 AROE. These CASs include results frameworks that link country outcomes, CAS outcomes, and Bank interventions and results matrices that summarize these results frameworks with performance indicators for M&E. A stocktaking report by OPCS of the seven Results-Based CAS pilots suggests that these CASs might be changing the content of conversations within country teams, and

between country teams and clients (World Bank, OPCS 2005a) to include a greater focus on results. As discussed in greater detail in Chapter 2, performance indicators still need to be better baselined and targeted. The Results-Based CAS can be further enhanced by taking a systematic approach to identifying, prioritizing, and ensuring collection of country information as part of the CAS preparation process.

The 2003 AROE noted that the CPPR was to be replaced by the Country Portfolio Review. The intent of the new Country Portfolio Review was to facilitate a more dynamic and integrative process to link the CAS with Bank projects, as well as selected AAA and grants and guarantees. The Bank did not formally revise its policies on CPPRs to incorporate the new approach because it was looking to reduce the number of formal policies as part of the simplification agenda.

With decentralization, three-quarters of country directors are in the field, and country managers or portfolio coordinators with portfo-

lio responsibility are in many country offices, and country portfolio discussions tend to be more continuous. As discussed in Chapter 2, several country directors are already exploring new CPPR approaches to incorporate AAA and are making the results link explicit. It would be possible and desirable to build on the experience to date, enhance the results focus of the CPPR, develop demand for performance information, and use the CPPR to jointly manage the Bank program with the government.

Sector-Level Results Monitoring Is Not as Far Along, but Progress Is Being Made

The 2003 AROE noted that the framework for monitoring outcomes of the Bank's sector strategies and programs had been evolving, but had not advanced as far as it had for country programs and specific instruments. Sector Strategy Papers lacked performance measures for M&E that included baselines and targets over a specified period.

Since the 2003 AROE, the Implementation Follow-Up (IFU) exercise has resulted in seven sectors with results frameworks and outcome indicators for annual progress monitoring.¹ This pilot approach will now be expanded to other sectors. While most of these frameworks need further development, they provide a starting point for Bank staff as they develop country results chains that link country development goals with specific Bank operations. A key challenge will be adapting a generic framework to country circumstances.

The Sector Strategy Implementation Update (SSIU) of February 2005 is a new report that replaces individual Sector Strategy Paper updates. It documents progress in all the sectoral (and thematic) strategies. It anticipates that the Results-Based CASs will improve the Bank's internal capacity to collect and report on sector-level results and will highlight intermediate and final outcome indicators.

The initial SSIU exercise highlighted two challenges: The first is developing and improving results frameworks for each sector and theme and integrating these frameworks into the Bank's operational work. The second challenge is strengthening the capacity and

demand to manage for results in client countries. For the next SSIU, it would be useful to report on progress in the collection of data for sectoral indicators and progress in strengthening the linkages between the indicators and outcomes.

Also, tensions between country priorities and global sector strategies need to be reconciled. Global sector objectives that are reflected in country-level sector strategies need to bear a relationship to country-level interests and feasibility.

Lending Guidelines and Templates to Enhance Monitoring Are Too New for Their Impact To Be Determined

The 2003 AROE noted that the framework for monitoring investment projects is well established, and the 2004 AROE pointed out that M&E has received the greatest attention in guidelines at this level. The AROEs pointed out two areas of weakness: establishing outcome indicators and monitoring and reporting on implementation progress. Since then, the Bank has taken several steps to strengthen the M&E focus of its lending instruments by requiring more systematic data collection and management for monitoring its lending operations and tracking implementation.

For investment lending, the Project Appraisal Document (PAD), which develops the rationale for the proposed investment loans, was revised in 2004 to increase the focus on results-oriented project design and monitoring. Instead of a logframe, the PAD now includes an annex on the M&E of outcome results and a results framework that includes the project development objectives and indicators, to measure progress toward meeting them. The indicators are accompanied by baseline and target values. The PAD format requires that arrangements for monitoring and data collection be noted, along with how local capacity for data collection will be supported where needed. Task teams now need to be more precise in indicating how and where the performance information will be collected and who will be responsible for preparing the data.

A desk review of 28 randomly selected PADs submitted to the Board in the first three quarters of fiscal year 2005 found that monitoring frameworks are being improved with the new format. Of the 28, 24 used the new format, while 4 used the logframe. The majority (21 of the 24) included targets for most of their indicators, but baselines continue to be a concern. One-third (8 of the 24) had inadequate baseline data. Often the baseline was “to be determined.” At the same time, one of the PADs using the logframe included detailed performance indicators broken down by year, with specific amounts. While the change in the PAD addresses the concerns raised in the previous AROEs, the key will be ensuring that the frameworks are actually used during project implementation to improve monitoring and implementation.

A new policy on development policy lending was issued in August 2004 that emphasizes the results framework in development policy lending and tightens the links between program objectives, reform measures, and outcomes. Information on good practice was issued in June 2004 to provide staff with guidance in preparing development policy operations. An internal good practice note on results measurement and monitoring in development policy was issued in June 2005. The program document highlights the country program being supported, the objectives of the operation, and the specific results expected from the policy program being supported by development policy lending. Measurable indicators are required that will serve as the basis for monitoring progress during implementation, as well as for evaluating the program on completion.

A desk review of 10 program documents submitted to the Board during the first three quarters of fiscal 2005 found that the policy matrices included the expected results of the loan credit. Development policy lending seeks to change practices within governments, and the impact may be difficult to observe, and therefore challenging to monitor and quantify. More experience with the new policy will be needed to understand the effect of its emphasis

on results frameworks and whether they have been strengthened.

With regard to tracking implementation progress, the 2003 AROE noted that the Project Status Report (PSR) was not fully effective in reporting on implementation performance during supervision. Few line managers used the PSR as a management tool because it did not report adequately on project outcomes. In 2005, a QAG report pointed out that the data underpinning portfolio monitoring indicators continued to be hampered by the absence of candor and accuracy in project performance reporting (World Bank 2005d).

To provide better-quality reporting on the status of project implementation and outcomes, the PSR was replaced in fiscal 2005 by an Implementation Status and Results Report (ISR). This tool strengthens and streamlines task team reporting to managers on the implementation performance of projects and their prospective outcomes. Sector managers are required to approve or disapprove updated ISRs. The ISR was introduced in January 2005, and after June 30, 2005, an ISR was requested for every active project.

The success of the ISR will depend on the degree to which it addresses the challenges encountered with the PSR, which included weak incentives for its use as a management tool. To encourage more focus on, and realism in, project supervision, portfolio oversight will be included in the results agreements of country and sector managers.

Monitoring AAA Remains a Challenge

The 2003 AROE noted that the monitoring framework for nonlending operations is relatively weak and that measurable outcomes and performance indicators are rarely set. It recommended that management strengthen the outcome orientation and evaluation of the Bank’s nonlending services at the aggregate and task level. The 2004 AROE reiterated the difficulty in applying results-based M&E and the Bank’s accountability for results to the majority of its AAA products.

An OPCS review of progress on AAA in 2004 led to a simple framework to improve the

accuracy of tracking the results of economic and sector work (ESW) and technical assistance (TA) activities and to link development objectives more closely to actual development results. Beginning July 1, 2004, task team leaders were required to define at the beginning of the ESW and TA activities the desired interim outcomes and to identify monitorable indicators of success.

The task-level tracking is a positive step for monitoring. It provides an important, albeit partial, perspective on monitoring and impact. The effectiveness of ESW and TA tasks is likely to be strengthened when the tasks are coordinated to have impact in aggregate and are strategically and operationally linked with complementary activities. While the impact of an individual task may be quite limited, the combination of strategically structured tasks could be significant. Designing ESW and TA accordingly and identifying outcomes from strategically focused work can come from strengthened country management oversight of the overall ESW and TA program. It would also help to address the challenge of managing AAA in a matrix environment and allow for better coordination across sectors, themes, and management units.

The 2003 AROE noted QAG's pilot program in the use of a country-level approach to assessing AAA quality. In the pilot, which began in 2002, QAG chose a sample of all ESW tasks during the CAS period in three countries and reviewed the quality of individual tasks and their contribution to the CAS objectives and country program. QAG has mainstreamed the pilot and has conducted 36 country AAA reviews, comprising 289 individual tasks. The country AAA assessments have found that while task-by-task assessments resulted in quality

ratings of 90 percent or more satisfactory, the strategic relevance of the entire program has been judged satisfactory in about 86 percent of countries, and the quality of follow-up activities and program

coherence was rated at 78 percent satisfactory. AAA has greater impact when the tasks have strategic relevance and are integrated, owned, and supported by the country and are characterized by dialogue and dissemination.

AAA can be better incorporated into the management of the Bank's country programs. The *Annual Report on Portfolio Performance* for fiscal 2004 recommended that Regions consider including AAA in the regular CPPRs (World Bank 2005c). It also suggested that AAA reporting be balanced away from the individual task and toward the country level. A similar theme is noted by OPCS regarding AAA, with the suggestion that the CAS Completion Report examine the contribution of AAA to the success of the overall country program, a recommendation that has been implemented. These steps would help to strengthen country management oversight of the AAA program.

Trust Funds. The 2003 AROE noted inadequacies in SAP (Systems, Applications, and Products in Data Processing – the Bank's budget and cost monitoring software) reporting on the use of trust funds and that VPUs did not have the ability to properly monitor the use of trust funds and their impact on Bank activities and country outcomes. The report recommended monitoring and reporting on the use of performance indicators for trust funds that are linked with the country activity they support.

The Bank has made a number of changes to improve the monitoring and reporting on trust funds. A new grant-monitoring tool requires reporting on progress in outputs, implementation, procurement goals, and disbursement targets. The Initiating Brief for Trust Funds Form was revised in February 2004. Ongoing work is needed. An internal review of the trust fund portfolio for 2004 concluded that little progress has been made in developing standard performance indicators to measure the results and development impact associated with a trust fund-supported program.

The 2003 AROE noted that OP/BP 14.40 was under revision. The content of the revisions is set out in the July 2004 Board presentation on trust funds, and the focus of the revisions is

AAA has greater impact when the tasks have strategic relevance and are integrated, owned, and supported by the country.

noted as primarily addressing nationality restrictions on procurement (World Bank 2004c). Strengthening M&E for trust-funded activities remains an unfinished agenda.

Corporate-Level Monitoring Is Gaining Ground

Millennium Development Goals. Since the Millennium Declaration was adopted five years ago, the international community has been engaged in ongoing stocktaking of MDG implementation and how progress toward achieving the MDGs can be accelerated. The urgency of this task is growing as the 2015 deadline for achieving the MDGs approaches. The Bank is pursuing an agenda that includes sharpening its emphasis on development results, as noted in the *Global Monitoring Report 2005* (World Bank 2005e). Two areas of priority for the Bank and development partners are highlighted in the report: managing for results and measuring for results. Key themes emphasize building capacity in countries to help them manage for results through the use of independent and self-evaluations.

IDA14. The Results-Based CAS is one element of the IDA14 results measurement framework, along with efforts to build results measurement capacity in client countries. The IDA14 results measurement framework has two tiers. The first tier focuses on monitoring country outcomes. It includes 14 country outcome indicators that are consistent with the MDGs, PRSP priorities, and IDA's activities in IDA countries. In addition, all new IDA operations in health, education, water supply, and rural road transport will, beginning in fiscal 2006, report on and strengthen country capacity to enable monitoring of key sector outcome indicators included in the IDA14 results measurement framework.

The second tier of the IDA14 results measurement framework monitors IDA's contribution to country outcomes. At the country level, the introduction of Results-Based CASs in IDA countries will be monitored and targeted to ensure progress in incorporating country-level results frameworks. At the

project level, four indicators will be tracked.² Finally, output data from completed projects in four sectors (health, education, water supply, and rural road transport) will be obtained from project completion reports and aggregated.

The success of the IDA14 results measurement framework will depend on the quality of the data for measuring progress on key indicators. An earlier review of IDA indicators found that only a small number have a well-established source of data (World Bank 2004d).³ At the other end of the spectrum, there are no trend data for other indicators. In between, data for a sizable number of indicators come from surveys that do not provide a consistent supply of data. In the short term, progress is needed to strengthen the data sources necessary to track progress on the IDA14 outcome indicators. For the long term, country capacity building, discussed in Chapter 2, will be required.

Strategy and Performance Contracts. Strategy and Performance Contracts (SPCs) are part of the Corporate Budget Reform and are intended to support a new type of dialogue between vice presidential units (VPUs) and senior management, and between senior management and the Board. The intent is to strengthen accountability for choices on strategy and resources and focus on key performance indicators. The SPCs build on the previous VPU Unit Compacts and introduce key performance indicators that will be used to monitor implementation. This approach to allocating administrative resources across Bank units was introduced in fiscal 2005.

The SPC raises the question as to its link with the results framework, where the country is the focus. In the SPCs, the Regions and the networks are the focal points. It will be important to explain operationally how performance measurement

IDA14 established a results-based framework that will be implemented in part through the Results-Based CAS.

The corporate budgeting and resource allocation process can provide incentives to advance the results agenda.

in the CASs will be aggregated at the Regional and network levels. For IDA countries, the IDA14 results framework builds a bridge between individual countries and an aggregate of countries. Guidance will be needed on how this process will take place for other types of countries. A balance also will be needed between standardizing performance measurement and accounting for variable country circumstances. The budgeting and resource allocation process will provide incentives that can advance the results agenda. As the budget reform process proceeds, it will be important for these to be aligned to further deepen the operationalization of the results framework.

Results Reporting System. OPCS is preparing a concept paper on a Results Reporting System (RRS) that is intended to reinforce the Bank's focus on helping its clients manage for results; build on existing systems; and harmonize, simplify, and streamline many different forms of reporting on results. This is a first step. RRS is expected to answer such questions as whether countries, sectors, and global programs are achieving outcomes and whether the Bank is contributing to these outcomes. The RRS will draw out linkages with the IDA14 results measurement system, the reformulation of the SPC, and other recent efforts to sharpen the results focus in other key business processes. The simplification process that the RRS may engender could be an important counterbalance to a proliferation of multiple reporting systems.

Evaluation

The Bank evaluates its operations through self-evaluation by the units responsible for operational programs and activities and by QAG and OPCS. Independent evaluation is carried out by IEG. The 2003 AROE recommended that management issue an OP/BP that sets out the mandate, framework, and roles and responsibilities for self- and independent evaluation. This OP/BP remains to be issued.

Country-Level Evaluation Is Developing

Country Assistance Evaluations provide the Board with an independent assessment of the

achievements and limitations of the Bank's country program, usually covering up to three CAS periods, and they offer lessons learned for the preparation of the upcoming CAS. Since the 2003 AROE, IEG has conducted a self-evaluation of CAEs (IEG 2005a). It noted that the CAEs yield a more complete picture of the outcome of the Bank's assistance programs than do individual project evaluations.

In one-third of the country programs evaluated, aggregate outcomes from projects were satisfactory, but the overall country assistance program was unsuccessful. The CAE Retrospective also concluded that CAEs place too much emphasis on Bank instruments and formal compliance issues and tended to neglect the Bank's development impact and the results achieved. A more results-based approach is needed that links the Bank's objectives, instruments, and the outcomes achieved. The recommended approach for future CAEs is that they outline key objectives and results the Bank has been trying to achieve based on the CASs, assess the relevance and efficacy of the instruments used to achieve each objective, and evaluate the Bank's contribution to the objective. In this way, CAEs will evaluate Bank impact at the country level.

The CAS Completion Report (CCR) assesses outcomes achieved relative to expectations in the four-year CAS time horizon. It addresses whether the previous CAS included an M&E system and performance indicators and whether information from the system was properly factored into the discussion on the achievement of specific CAS outcomes.

IEG conducts a review of the CCR. IEG's review is an intermediate product, with less in-depth coverage than a CAE. IEG draws evaluative judgments to independently validate major CCR findings. IEG's criteria for exemplary CCRs include a strong focus on M&E. IEG delivered 11 CCR reviews in the last 2 months of fiscal 2005. The timing of IEG's delivery depends on the CAS schedule.

The Sector Level Is Getting More Attention, but Progress Needs to Be Deepened

The 2003 AROE reported that there is no

framework for self-evaluation of sector outcomes. In the past, IEG has provided independent evaluations of Bank experience in numerous sectors, timed to serve as inputs to a new or updated sector strategy. IEG and management will need to develop a formal framework and timeline for evaluating the outcomes of sector strategies.

Efforts Are Under Way to Strengthen Lending Evaluation

The 2003 AROE noted that a constraint to evaluation of lending operations is that the design of project completion reporting has not kept pace with the evolution of lending instruments. It was recommended that the special features of Development Policy Loans, Adaptable Program Loans, and Learning and Innovation Loans be reflected in the guidelines for their Implementation Completion Reports (ICRs).

Since the 2003 AROE, OPCS and IEG have provided guidance on simplified ICRs for programmatic development policy lending (DPL). This was in response to concerns raised by operational managers and staff who questioned the requirement to prepare an ICR within six months of closing programmatic adjustment loans, since individual loans are linked, by design, to produce results at the end of the series. In consultation with operational staff, OPCS and IEG agreed on guidelines for simplified ICRs that apply to each operation in a programmatic series, except the last. The simplified ICRs meet the requirements in Operations Policy/Bank Procedure 13.55 on Implementation Completion Reporting. They also reduce possibly duplicative work and meet reasonable information and self-evaluation requirements. A standard ICR that encompasses the entire programmatic series is required after the last operation.

In January 2005 new joint OPCS/IEG guidelines were issued for evaluating projects with formally revised objectives. They apply to both ICRs and IEG assessments and seek to harmonize portfolio monitoring and rating methodologies. The guidelines require that outcome ratings be weighted in proportion to the

share of actual loan or credit disbursements made in the periods before and after approval of the revision. These guidelines contribute to strengthening the self-evaluations and independent evaluations of Bank projects.

Greater focus is being placed on impact evaluations. The Development Impact Evaluation Initiative (DIME), a pilot initiative led by the Development Economics Vice Presidency (DEC), is intended to address gaps in the Bank's conduct of project impact evaluations and to strengthen the development community's global learning. DIME is supporting systematic evaluation of programs in a selected number of strategic priority themes in the Human Development, Infrastructure, and Environmentally and Socially Sustainable Development (ESSD) Networks.

As an example, impact evaluations of land reform programs in Brazil, Colombia, India, Mexico, and South Africa will be undertaken to determine the effect of the programs on the productivity of land use and on the welfare of affected populations. In addition, DIME will support evaluation studies focused on specific countrywide economic policies, such as the poverty impact of reforms on infrastructure policy and trade.

This initiative may prove useful for supporting and strengthening M&E in the Bank and can be expected to provide information that could be valuable to borrowers and others in the development community who want to know what works in different country contexts. DIME also has the potential to improve the identification of monitoring indicators of intermediate outcomes.

IEG has completed two impact evaluations in partnership with DFID: the Ghana Basic Education Program and the Bangladesh

IEG and management will need to develop a formal framework and timeline for evaluating the outcomes of sector strategies.

Greater focus is being placed on impact evaluations, which are expected to provide information on what works in different country contexts.

Maternal and Child Health Program. An impact evaluation of selected India Rural Poverty programs is ongoing. These ex-post impact evaluations use advanced quantitative evaluation techniques (based on a quasi-experimental design) and demonstrate that impact evaluations can be undertaken even when baseline data are not available.

Trust Fund and Global Program Evaluation Are Not Fully Effective

The 2003 AROE noted that the evaluation framework for trust funds was partial, focused on compliance with Bank policy and donor conditionality, and needed to be strengthened. It suggested that management set guidelines for independent evaluations of large trust funds because this would be consistent with the existing practice that requires external evaluations of Development Grant Facility (DGF) grants larger than \$300,000. The Trust Fund Operations (TFO) and Global Programs and Partnerships (GPP) group are currently in discussions with IEG on ways to advance M&E for trust funds and GPPs.

An Implementation Completion Memorandum (ICM) is required for all activities in excess of \$1 million funded under trust funds. The ICM is a self-assessment instrument that monitors the degree of success of the trust-funded activities as measured by outputs, outcomes, impact, and Bank and recipient performance. The Trust Fund Quality Assurance and Compliance Unit reviewed 74 ICMs received during fiscal 2004 and found that many reports did not identify the impact of the trust-funded activities, but instead identified inputs and outputs such as workshops and reports. It was recommended that the Bank and donors initiate a targeted M&E program for larger country-based or sector-specific trust fund programs. It further suggested that OPCS and Concessional Finance & Global Partnership (CFP) develop standards that defined when M&E is required for trust fund programs, based not only on dollar amounts but also on the significance of the trust fund program. A targeted M&E program would help address the 2003 AROE recommendations that called for

stronger M&E of development outcomes for trust-funded activities.

Grants complement the Bank's lending and nonlending operations. Grant Completion Reports (GCRs) are required only when a program exits from DGF funding, a process that can take several years. Most of the evaluations of grant programs over \$300,000 are done by external consultants paid by program management, and are not regarded as truly independent.

In response to IEG's Phase I report on global partnerships and programs (IEG 2002), the Bank has implemented organizational and procedural changes in global program management. The DGF has instituted an external peer review process for new programs seeking grant support. In response to IEG's meta-evaluation of the CGIAR (IEG 2004e), Bank management has accepted the principle of independent oversight by assigning oversight of the CGIAR to the Bank's chief economist. These actions are intended to strengthen the Bank's global program portfolio. But further improvements are needed to enhance the linkages between global programs and country operations.

As part of the 2004 IEG Phase 2 review of the Bank's involvement in global programs and partnerships, IEG reviewed 26 Bank-supported global programs that accounted for 90 percent of the Bank's global program expenditures in 2002 (IEG 2004d). The review recommended that global programs be included in IEG's standard processes of evaluation and reporting to the Board. This would involve working with the Bank's global partners to develop and apply international standards for the evaluation of global programs. It also recommended that IEG review program-level evaluations conducted by Bank-supported global programs, similar to IEG reviews of other self-evaluations at the project and country levels.

In April 2005, IEG issued guidelines for its Global Program Reviews (GPRs). These reviews of global program evaluations are seen as parallel to existing review processes for the evaluation of investment projects and country programs, and will therefore be built on established models for ICR Reviews and Project

Performance Assessment Reports (PPARs) for investment projects, and CAS Completion Report Reviews and CAEs for country programs.

The GPR is based on a review of the external evaluation of the global program, and in some cases on fieldwork conducted by IEG. A key difference between GPRs and the reviews of projects and country programs is that global programs are partnership programs. Instead of being the responsibility of a Bank operational Region or network, global programs are the responsibility of the governing body of the program, and the Bank is only one member, though it often acts as the chair.

In the context of results orientation and accountability for results, it should be noted that in cases where the Bank's role is that of fiscal agent, the Bank is not accountable for program results. Global program evaluations are the property of the global program, which suggests the need for common standards for global evaluations. Also, global programs, unlike investment projects, are ongoing rather than completed.

IEG will initiate two pilot GPRs in fiscal 2005 that will be similar in scope to a PPAR. They will include lessons learned rather than recommendations. Like PPARs, each review will undergo a multi-tiered review within IEG. For the pilot phase, IEG plans to ask the Board's Committee on Development Effectiveness (CODE) to review the first two pilots in order to provide feedback to IEG regarding the new product.

Conclusions

The Bank is making progress in strengthening a results orientation in M&E. Preparation of the CAS Completion Reports, the IEG reviews of these reports that are now required for all CASSs, and the mainstreaming of the Results-Based CAS provide a solid framework for managing for results at the country level. The Bank is rightly moving its focus from instruments to the country.

Whereas the CAS has traditionally been built up from the instrument level, instruments are

increasingly being seen as designed to support the CAS. New CAS progress reports are expected to provide just-in-time learning.

The development and refinement of results frameworks for all sectors as part of the SSIU is a start in providing Bank staff with tools to establish the links between country and CAS objectives and to identify intermediate outcomes that can be used to monitor progress.

The revision of the PAD, introduction of the ISRs, consolidation of adjustment lending into DPLs with a stronger emphasis on results, and introduction of simplified ICRs for development policy lending enable the Bank to obtain a clearer picture of the results of its operations. These steps foster the identification and reporting of clear and monitorable indicators and the tracking of implementation progress.

While policies and procedures are being put in place, it will take some time before the Bank is able to effectively manage for results. Experience from developed countries, where results-based management was introduced in the late 1980s and early 1990s, indicates that there is a significant learning period for an organization as managers and staff expand their focus from managing inputs such as budgets and administrative matters to managing outputs, and then outcomes (box 3.1). The challenges in developing countries will be more pronounced because of data and capacity constraints. Bank management will need to align incentives to manage for results. It has taken an important step in this direction by incorporating portfolio oversight as an element in the annual reviews for all managers of country and sector management units.

IEG issued guidelines for its reviews of evaluations of global programs in April 2005.

While policies and procedures are being put in place, it will take time before the Bank is able to effectively manage for results.

Box 3.1: Lessons on Results-Based Management in the U.S. Federal Government

The Government Performance and Results Act of 1993 requires U.S. federal agencies to provide annual data on the outcomes of each major federal program. The primary focus has been to strengthen accountability for program results.

Several obstacles to using outcome data have been observed. Some managers may not be aware of the variety of ways in which they can use outcome data, or of which particular data can be used for specific purposes. Outcome data use is sometimes limited by data problems. Some data are old by the time they reach program managers; data that are two or three years old do not have the currency that program managers usually need to react. Some of the data may not be broken out in sufficient detail to be useful to particular

program managers. Many federal programs, such as those that work through other levels of government, require long periods of time before they can be expected to yield the major outcomes sought.

Although a foundation has been established, there are still major challenges in implementing results-oriented government. Inconsistent commitment to achieving results among top leadership can hinder the development of results-oriented cultures in agencies. Furthermore, in certain areas, federal managers continue to have difficulty setting outcome-oriented goals; collecting useful data on results; and linking institutional, program, unit, and individual performance measurement and reward systems. Finally, there is an inadequate focus on issues that cut across federal agencies.

Sources: Hatry and others 2004, pp. 5 and 13; GAO 2004.

Chapter 4: Evaluation Highlights

- IEG is requiring evaluations at the country, sector, and project levels to give greater attention to results-oriented M&E.
- IEG's Outreach and Communications Strategy is aimed at improving the communication and dissemination of IEG evaluation findings.
- IEG client survey results highlight the need for evaluation findings to be made more operational.



Improving IEG's Effectiveness

IEG evaluates its own products and services as part of a continuous process of self-evaluation. This chapter discusses IEG's response to recommendations for itself contained in last year's AROE and the results of the IEG client survey, which assesses Bank staff awareness about IEG studies and their perceptions of quality.

IEG's Response to 2004 AROE Recommendations

The 2004 AROE included two recommendations for IEG:

Recommendation 1: Strengthen attention to results-oriented M&E in IEG by requiring evaluations at the country, sector, and project levels to address systematically the three criteria for results-oriented M&E, and increase the weight that IEG ratings give to the extent of results-oriented M&E and managing for results in Bank programs. The three criteria for assessing results-oriented M&E are the results framework, the design of M&E, and the implementation of M&E.

At the country level, beginning January 2006, CAEs will note the extent to which the CASs have become results focused, including the definition and analysis of a reasonable quantity and quality of indicators, and the extent to which the CASs and the CCRs made good use of the outcome indicators to arrive at assessments of the overall

outcomes of Bank assistance. The CAEs will also examine the degree to which the Bank's assistance program included a focus on supporting the country's efforts to adopt a more results-oriented framework—for example, through defining outcomes indicators in PRSPs or other strategy documents of the country. The degree of Bank support provided through loans or grants to develop the capacity to improve or make use of existing systems of data gathering and analysis and other monitoring will be noted. Finally, the CAEs will assess the extent to which the Bank's efforts resulted in improvements in the quality, timeliness, and availability to end users of monitoring indicators.

At the sector level, IEG is increasing its emphasis on monitoring and evaluation and results frameworks in its sector and thematic evaluations. As an example, two major recent studies, of Community Development (IEG

CAEs will assess the focus of Bank assistance on supporting the country's adoption of a more results-oriented framework.

2005b) and HIV/AIDS (IEG 2005c), contain in-depth analyses and specific recommendations to help improve the quality and quantity of M&E in the Bank.

At the project level, IEG is introducing a separate section in PPARs and ICR reviews that will discuss design, implementation, and use of M&E.

At the project level, IEG is strengthening the focus on M&E and results frameworks. While most IEG PPARs and many ICR reviews comment, and often include lessons learned, on monitoring and evaluation, IEG is introducing a separate section in these reviews that will discuss design, implementation, and use of M&E. Monitoring and evaluation is receiving increased emphasis in IEG assessments. In some cases, inadequate M&E has been an important reason for downgrading Bank performance. Also at the project level, IEG is adopting a results framework approach in its impact evaluations by testing the results achieved by specific interventions.

At the global level, IEG's recently completed review of global programs contains a recommendation on the evaluation of Bank-supported global programs. IEG is piloting reviews of the Bank's self-evaluations of individual global programs.

Beginning in 2005, IEG instituted an award to recognize good practice in M&E as part of the annual IEG Awards Program. The purpose of the award is to emphasize the importance of M&E in ensuring the effectiveness of Bank assistance and to create incentives for good performance measurement and management.

IEG's Outreach and Communications Strategy is aimed at ensuring that IEG findings reach their target audiences at the right time, in the right form, and are easily accessible.

Recommendation 2: Communicate and disseminate findings and lessons of IEG evaluations more broadly, both inside and outside the Bank, in line with the new communications strategy. IEG's new Outreach and

Communications Strategy (IEG 2005d) is aimed at ensuring that IEG findings reach their target audiences at the right time, in the right form, and in a manner that the recipients can access easily.

The new strategy encompasses three approaches: embedding outreach and communications throughout the evaluation cycle, targeting underserved audiences, and accelerating learning from evaluation findings. IEG will capitalize on these three approaches by streamlining internal communications practices through greater consultation, earlier audience identification, and systematic integration of communications throughout the evaluation cycle.

IEG products will be improved through better content and design. Also, the targeting and delivery of IEG findings will be enhanced through better audience identification, adoption of advanced electronic tools, and strategic partnerships to promote dissemination.

Finally, learning and use of evaluation lessons will be accelerated through increased participation in learning events by Bank staff, greater sharing of IEG findings with operational teams, and greater use of public forums. Pilot initiatives are under way to promote more effective internal communication approaches and to shift outreach efforts from a supply-driven model to one that is demand-driven, customized, and timely.

Performance indicators will be used to help IEG assess the impact of the Outreach and Communications Strategy. In the annual client survey, performance indicators include awareness, readership, timeliness, and quality. In addition, audience feedback will be tracked regularly by monitoring changes in Web usage as measured by unique monthly visitors. Whether target audiences are being reached will be assessed by tracking Web users, the content they visit, and the time they spend. Targeted delivery of IEG products will be tracked by monitoring the percentage of evaluation consultees who received the final product. External dissemination events such as press activities and external briefings will be monitored for the number of references to IEG

products in the media and attendance at briefings by key stakeholders.

IEG Client Survey Results

As part of IEG's self-evaluation, it annually surveys its clients about the readership and perceptions of its major studies. This year, IEG conducted two surveys. The first focused on the Bank's internal clients and asked 2,120 of them about IEG products to find out about readership and awareness, perceptions of quality, the influence of the reports, and recommendations for improving IEG products. The second asked 926 external clients—government officials, researchers, and consultants—similar questions about IEG products. The following section highlights key findings. Detailed survey results are in Appendix D.

IEG Internal Client Survey Results

This year's survey targeted several kinds of IEG products, including eight Country Assistance Evaluations, the *Annual Review of Development Effectiveness*, PPARs, and evaluation studies. IEG surveyed a targeted sample of 2,120 internal clients, including Bank staff associated with the products and executive directors and their advisors. The response rate was 31 percent, an increase over the response rate for the 2004 surveys (21 percent).

Readership and awareness. Almost half (45 percent) of Bank staff who responded to the survey had read, participated in a discussion of, or attended an event on the evaluation for which they were surveyed. This is a marked increase over the levels of awareness found in previous AROE surveys (2004: 27 percent). Eighty-five percent of executive directors and their advisors reported that they had read, participated in a discussion of, or attended an event for a 2004 evaluation.

Quality. Two-thirds or more of Bank staff respondents rated as good or excellent the depth of analysis (71 percent), presentation of the evidence (68 percent), presentation of conclusions (73 percent), and relevance of lessons to the Bank's work (67 percent, with 21 percent rating this

aspect as excellent). Executive directors and their advisors rated the evaluations more positively, with at least 84 percent good or excellent ratings for each aspect.

Influence of the IEG study.

More than three-quarters (78 percent) of Bank staff respondents indicated that the evaluation they were asked about influenced their view of the subject areas either a great deal or somewhat. Among country directors and managers, 92 percent said they would use the findings of the evaluation in their work. Executive directors and their advisors (95 percent) indicated that IEG evaluations influenced their view of a subject area, with 40 percent reporting a great deal of influence.

Respondent recommendations to IEG. Based on the results of the survey and recommendations from respondents, IEG should make the findings in the evaluations more operational.

Other areas that need greater attention include strengthening the links in evaluations between conclusions and evidence. Also, IEG

needs to continue to engage in broader consultation with Bank staff to improve its evaluations. Finally, IEG's outreach can be enhanced with better notification regarding new documents and the targeted provision of summaries of IEG findings. IEG's Outreach and Communications Strategy is intended to be responsive to these suggestions. Appendix D has details of the survey results.

IEG External Client Survey Results

This year, in a survey of the Bank's external clients, 926 government officials, consultants, and researchers were

IEG found that 45 percent of respondents were aware of the relevant evaluation — a marked increase from previous years.

IEG needs to make its evaluation findings more operational and to better communicate the results of its work.

IEG's client survey found that two-thirds of respondents were aware of the evaluation in question.

asked about one of two IEG products that had been subject to targeted external dissemination: IEG's evaluation of Global Programs and Partnerships and IEG reports on M&E instruments and approaches to capacity building, which had been distributed to International Program for Development Evaluation Training (IPDET) alumni.¹

Readership and awareness. Two-thirds of respondents had read, participated in a discussion of, or attended an event on the evaluation product for which they were surveyed. The vast majority of external client respondents surveyed about the M&E reports heard about the product by e-mail or online, while 53 percent of GPP respondents heard about the evaluation in this way. Half of the GPP respondents indicated they had heard about the evaluation by being consulted during the evaluation process.

Quality. A large majority of respondents (94 percent) indicated that IEG's report was timely. Eighty percent of GPP respondents said the evaluation made full use of the information available, and 83 percent noted that they thought the evaluation used a methodology that was transparent and clear. Ninety-two percent of the IPDET alumni who responded

stated that they thought the evaluation paper presented information in a transparent and understandable manner. Although a majority of all respondents rated each of the quality dimensions positively, the presentation of evidence (71 percent good or excellent) and the link between conclusions and evidence (71 percent good or excellent) were dimensions of quality for which ratings were lower than for other dimensions—for example, depth or quality of analysis (82 percent good or excellent).

Influence. Seventy-four percent of respondents said they had used or would use the findings from the IEG products. About two-thirds of GPP respondents said that the study influenced their view a great deal or somewhat. Variation was noteworthy among respondent groups with regard to the influence and use of the IEG products. Eighty-eight percent of IPDET respondents said they had or would use the findings of the IEG evaluation product in their work.

Recommendation for IEG

As with the internal survey findings, respondents to the external survey most often selected the recommendation for IEG to make findings more operational.



Conclusions and Recommendations

The Bank is making progress in strengthening the results focus of its operations at the country, sector, and instrument levels. The increasing emphasis on outcomes will lead to joint accountability for outcomes.

As countries assume more responsibility for M&E and performance measurement, they will have to develop both monitoring systems to capture outcomes and evaluation capacity to process and use the information. This will require continuing Bank efforts to strengthen M&E capacity at the agency and country levels, harmonizing its efforts with those of the country and other development partners. As part of CAS preparation, the Bank should systematically assess the country's capacity to generate, collect, and analyze performance information. It should identify gaps and support capacity development as needed.

This shift will also require governments to become more aware of the need for performance information and to develop knowledge of its uses. Increasing the results focus of CPPRs and expanding them to include AAA would allow the Bank and the government to focus discussion on the performance of the Bank's overall program and its contribution to achieving country development and Bank CAS objectives. An expanded, results-focused CPPR would help link the CAS outcomes to project

and sector outputs and outcomes. It would also inform and complement the CAS Progress Report and the CCR. The 2003 AROE noted that management was replacing the CPPR with a Country Portfolio Review, but this has not been done. It may now be appropriate to revisit the CPPR as a tool for managing country programs.

Recommendations

For Management

- Incorporate a diagnosis of the country's performance measurement capacity and recommendations for action to strengthen that capacity in the CAS. The diagnosis and actions should be customized to country capacity and focused on critical information needed for the government to implement its program.
- Expand use of the Results-Based CAS as a tool to manage country programs and strengthen the link between the Results-Based CAS and the country program by providing country teams with guidance designed to increase the results orientation of CPPRs.

For IEG

- Review the results of pilot initiatives to promote more effective outreach and communication of IEG products and consider lessons learned in
- Subject to CODE/Board endorsement of the new procedures, institute a formal system for tracking and managing IEG recommendations.

the full implementation of the communication strategy.

APPENDIXES

APPENDIX A: REVIEW OF CAS RESULTS MATRICES

Objectives, Scope, and Methodology

Objectives

The objective of the review is to evaluate the effectiveness of the CAS results matrix as a tool for country-level Bank management. The results matrix is a key element of the Results-Based CAS. The matrix separates out and defines country development goals, CAS outcomes, Bank intermediate outcomes, and Bank programs (as columns) for each goal (as rows). By using this matrix, Bank operations are expected to be more explicitly linked to overall country goals, an important change from the Program Matrix of the previous (traditional) CAS. The review considered both the usefulness of the format for presenting the information needed to manage the Bank's country program and the actual formulation of each matrix for each CAS. A sample results matrix is presented as Attachment 1 to this appendix.

Scope

Twenty-four CASs presented to the Board in fiscal 2004 and 2005, as of end-January 2005, were reviewed. They were categorized as follows:

- Six Results-Based CAS Pilots—Armenia, Brazil, Cameroon, Mozambique, Ukraine, and Zambia
- Thirteen Results-Oriented CASs that employed some form of the results-based approach, although they were not formal pilots
- Five traditional CASs that contained sufficient information for analysis, although they followed the traditional CAS format.

During this period, nine CASs were presented to the Board using the traditional CAS format. Of these, five contained sufficient results informa-

tion to be evaluated and were included in the analysis as a baseline or control group for comparison with the results-oriented CASs.

Methodology

The results matrices were examined to assess: (a) how the structure and content of the framework contributed to the Bank's results focus and (b) how effective the performance measures would be.

A maturity model¹ was used to evaluate the matrix framework and the performance measures. Assessment criteria were established and defined, and three levels of maturity, from Level 1 (lowest) to Level 3 (highest), were defined for each criterion. Level 0 describes a CAS matrix that could not be evaluated in that criterion for lack of data. The methodology is presented in Attachment 2.

The approach used to assess the framework and the measures is described below.

- **Analysis of the Matrix Framework:** Three criteria were developed to gauge the overall maturity (strength) of the framework: (a) the structure of the logical sequence of the matrix (**logic chain**), (b) the content of the matrix (**results chain strength**), and (c) the representation of partner programs (**linkage to other donors**). For the logic chain, the columns and rows of the results matrix were reviewed to determine how effectively links between country goals, Bank outcomes, and Bank operations were portrayed. For results-chain strength, the links between country goals and Bank operations for each goal were examined to determine how easy it was to follow the relationship between the goals, measures, and programs.

- Analysis of Performance Measures:** The performance measures in the matrices were assessed on the following criteria: defined, baselined, targeted, and aligned² (see Attachment 2 for definitions). These criteria were applied to three sets of performance measures: (a) **country-level measures**, (b) **CAS outcomes**, and (c) **intermediate indicators and milestones**. Each criterion pertains to qualities of individual measures. In order to determine maturity for the whole set of measures in the matrix, maturity thresholds were established based on the proportion of measures that met each criterion: that is, Level 1 if less than half of the measures apparently met the criterion; Level 2 if more than half but less than three-quarters of the measures met the criterion; and Level 3 if more than three-quarters met the criterion.

comes to intermediate outcomes and to specific Bank programs. All of the Results-Based CAS pilots and results-oriented CASs were rated at higher levels of maturity. One traditional CAS, Bolivia, used its program matrix to articulate the relationship between country goals and Bank operations (figure A.1).

The Bank's standard results matrix format has been a good starting point for many Bank country teams. Several have adjusted the standard format to better articulate their country strategy and the relationship between country and CAS goals, and between Bank CAS outcomes and operations.

For example, many CAS matrices introduced some structure in the rows of the matrix, organizing the column information into related sets of goals and programs. This makes it easier to comprehend the linkages between different country development goals and how they jointly contribute to a higher objective, and to identify the potential synergies between Bank operations that support individual country development goals.

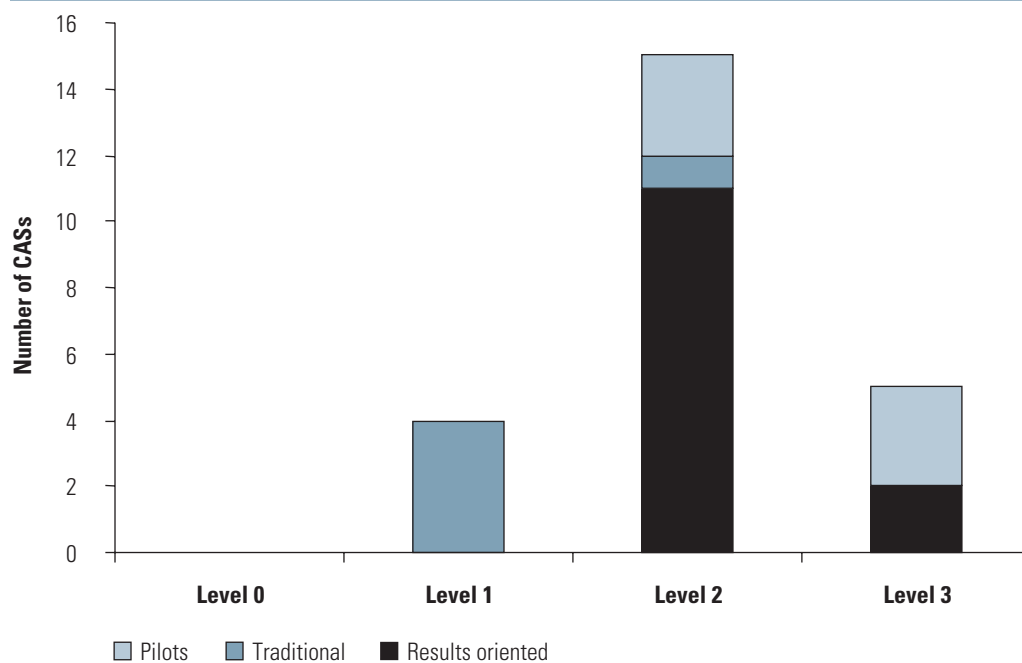
The review also identified some weaknesses. A typical problem was the lack of a structure

Results of the Matrix Structure Evaluation

Logic Chain

Not unexpectedly, the results matrix format does a better job of relating country goals to CAS out-

Figure A.1: Logic Chain



within the columns that made it difficult to follow the logical links between country goals, CAS outcomes, intermediate outcomes, and operations. For example, if one country goal was followed by 10 CAS outcomes, another 20 intermediate outcomes, and 25 programs, with no explicit or implicit structure, it was difficult to understand the interactions between actions and results. Another weakness was a lack of understanding of the use and meaning of the matrix. For example, some CAS reports did not include any measures in the CAS outcomes column. Others mixed long-, medium-, and short-term outcomes between the CAS and intermediate columns.

Results Chain Strength

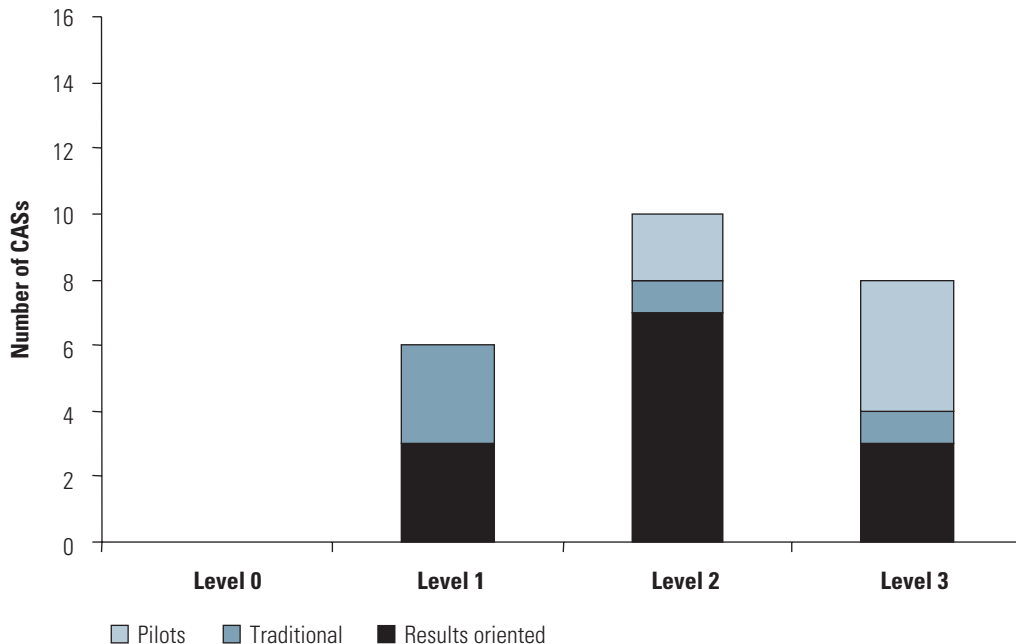
Results-Based CAS pilots did a better job of articulating the results chain, allowing the reviewer to follow the link between Bank operations, intermediate outcomes, CAS outcomes, and country development goals for individual goals. The results-oriented CAS formats, especially the pilots, exhibited higher levels of maturity; more traditional CAS reports were lower on the maturity scale (figure A.2). A notable exception was the

traditional CAS format used by Bolivia, which was evaluated as a Level 3 results chain. In this case, the country team used an enhanced version of the traditional CAS program matrix to draw clear results relationships between programs and Bank and country goals.

Countries with Poverty Reduction Strategy Papers (PRSPs) tended to have better results chains. The education, health, and environment sectors were relatively better at creating a compelling, results-oriented narrative through the contents of the results matrix. Even in countries where the matrix itself was not well developed, or where the other sectors were less mature, the programs in these sectors were frequently described effectively in results-oriented terms. In contrast, anti-corruption and governance areas were often less mature, probably reflecting the greater difficulty of establishing clear, incremental results chains from specific Bank activities to country-wide results.

More CASs (eight) were evaluated at the highest level of maturity, Level 3, for results chains, compared with the logic chain criterion (five CASs). This indicates that some CAS matrices were able to overcome structural deficiencies in

Figure A.2: Results-Chain Strength



the presentation of the results matrix and to show the logical relationship of country goals, Bank outcomes, and Bank activities. Typically this is due to a well-developed underlying results chain, representing a mature understanding of the relationships between activities and results in particular programmatic areas.

Linkage to Other Donors

The CAS matrices generally did a poorer job of showing the interaction between Bank programs and those of other partners. There were virtually no differences between the Results-Based CAS formats and the traditional formats (figure A.3). Although not statistically significant, the traditional CAS formats were more likely to have higher levels of maturity under this criterion. The format of the matrix itself is a constraint. A two-dimensional matrix is limited in space, which may make it difficult to provide sufficient information on partner activities alongside effective descriptions of Bank results chains. Several CASs used different approaches to display the relationships with partners. For example, the Paraguay CAS included a separate “Partnership

Matrix” (with strategic areas as columns and the different development agencies as rows) showing the estimated weighted effects of partner activities on country goals.

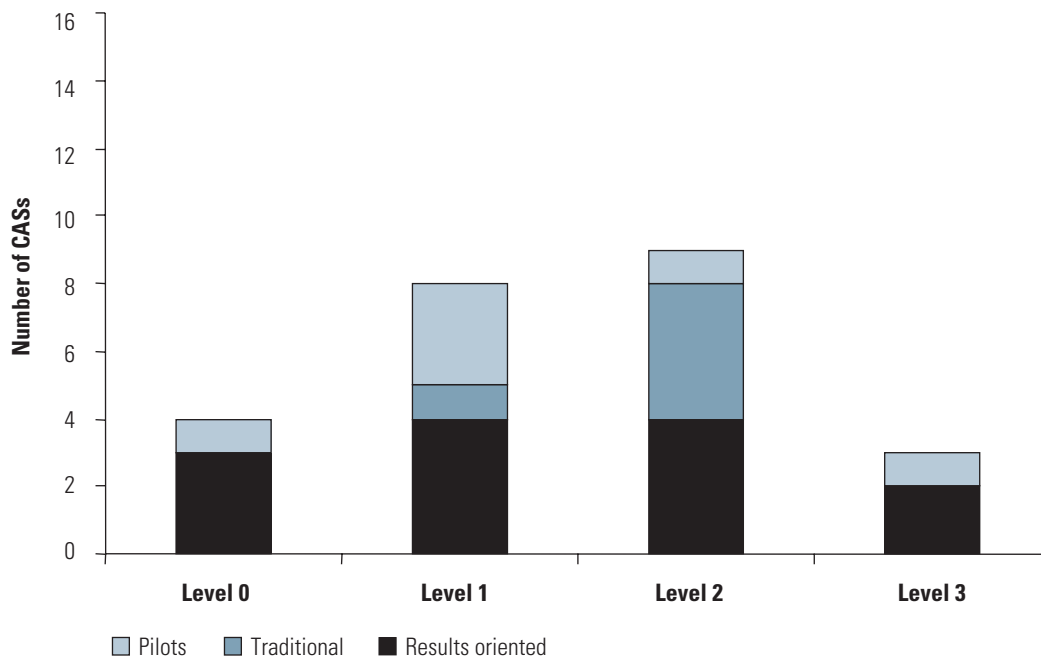
Results of Performance Measure Evaluation

Results-Oriented Formats More Mature than Traditional Formats

The review of the performance measures found that those in results-oriented CAS formats, especially the pilot countries, generally were better developed than those in traditional formats. Examining the measures across the three levels (country, CAS, and intermediate), countries with PRSPs tended to have more mature country-level performance measures. Some traditional CAS reports are rated Level 2 and 3 at the country level for this reason.

Performance measures were generally well defined: that is, a measure is clearly identifiable as a performance measure as distinct from a goal or activity. Most CASs demonstrated high maturity for all three levels of measures, with a

Figure A.3: Linkage to Other Donors



large number of CASs (19, or 79 percent) rated as Level 2 or 3 for country performance measures (figures A.4, A.5, and A.6). Part of the explanation is the use of widely accepted measures at the country level, such as the Millennium Development Goals.

Similar patterns were noted for the alignment of measures. The measures generally were presented at the right levels in the results chain (figures A.7, A.8, and A.9). For example, long-term, national health, or economic measures, such as growth of GDP or reduction in national infant mortality rates, would be appropriately aligned in the country level, but misaligned at the intermediate level. The PRSP and MDGs served as a good starting point for measurement at the country level for both results-oriented and traditional CAS formats, but are often out of place at lower levels of the results chain.

There is a fall-off in alignment maturity through the CAS and intermediate levels. Typically, this indicates that a CAS report, while possessing the appropriate measures for the country level, inappropriately relied on country-level measures in the CAS and intermediate levels.

Maturity Level 0 typically indicates that there were no measures at all, or at least no appropriate measures at the CAS or intermediate levels.

Difficulties in Developing Bank Outcome Measures

A comparison of measures across the three levels—country, CAS, and intermediate – indicated that the Bank finds it more difficult to develop measures at the CAS and intermediate levels. In many cases, country measures were aligned, and had baselines and targets, but CAS and intermediate measures were missing these qualities. Since the general trends were comparable across the three criteria (aligned, baselined, and targeted), the results of these assessments were aggregated. When all CAS reports are summed across these three measurement criteria, there were 21 cases of measures at Level 0 at the CAS level and 19 at the intermediate level, but only 7 at the country level (figure A.10). Conversely, there were 27 cases of Level 3 measures at the country level, 10 more than at the intermediate level and 11 more than at the CAS level. The higher overall levels of maturity for country-level

Figure A.4: Measurement Definition: Country Level

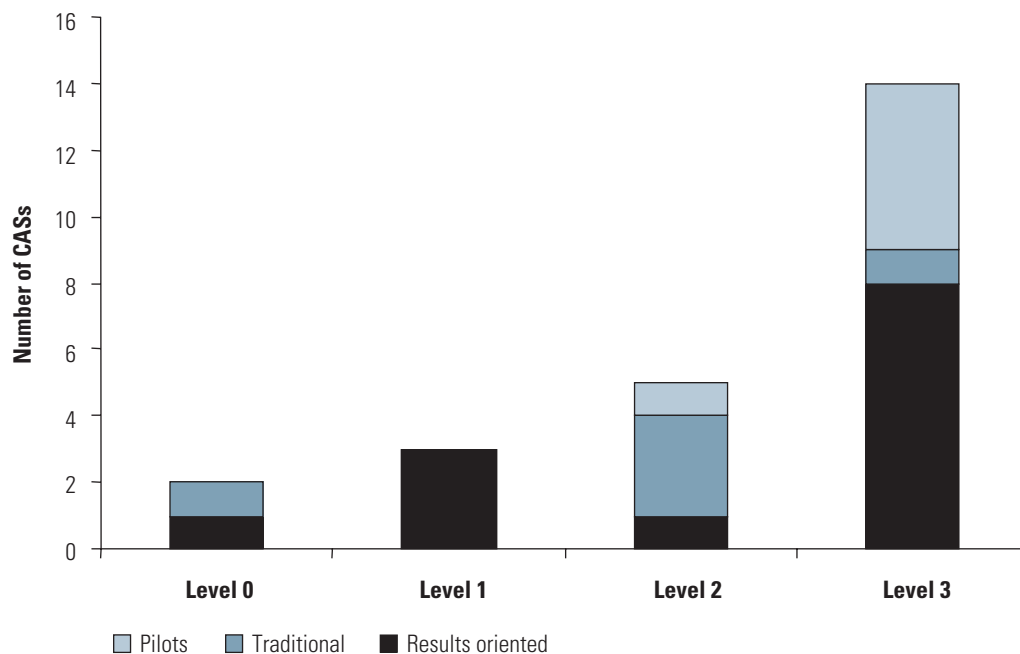


Figure A.5: Measurement Definition: CAS Level

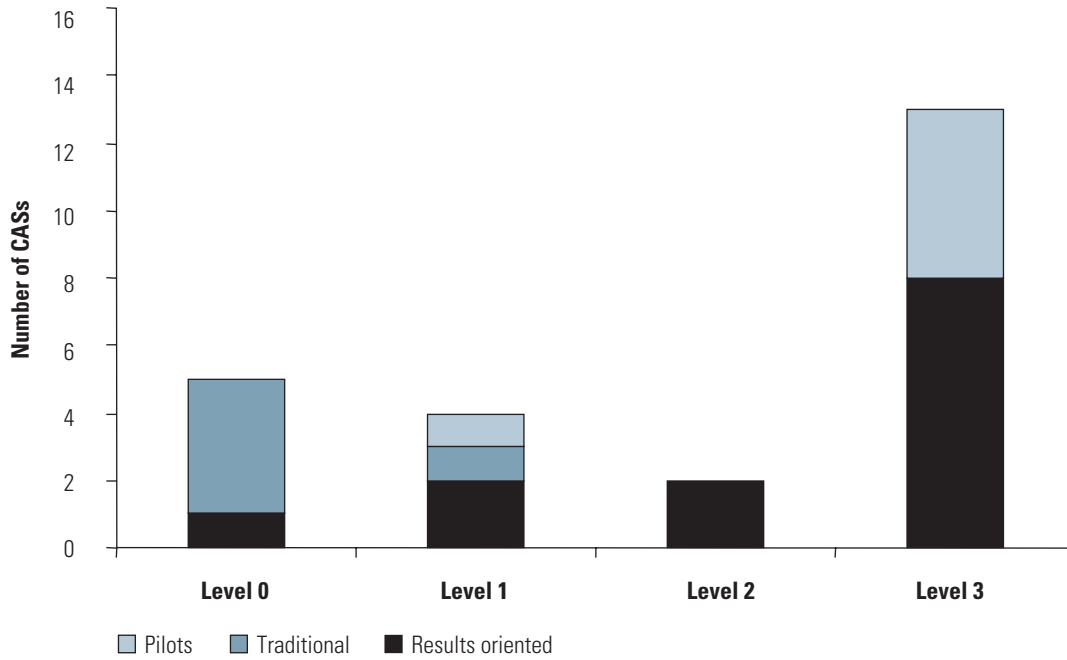


Figure A.6: Measurement Definition: Intermediate Level

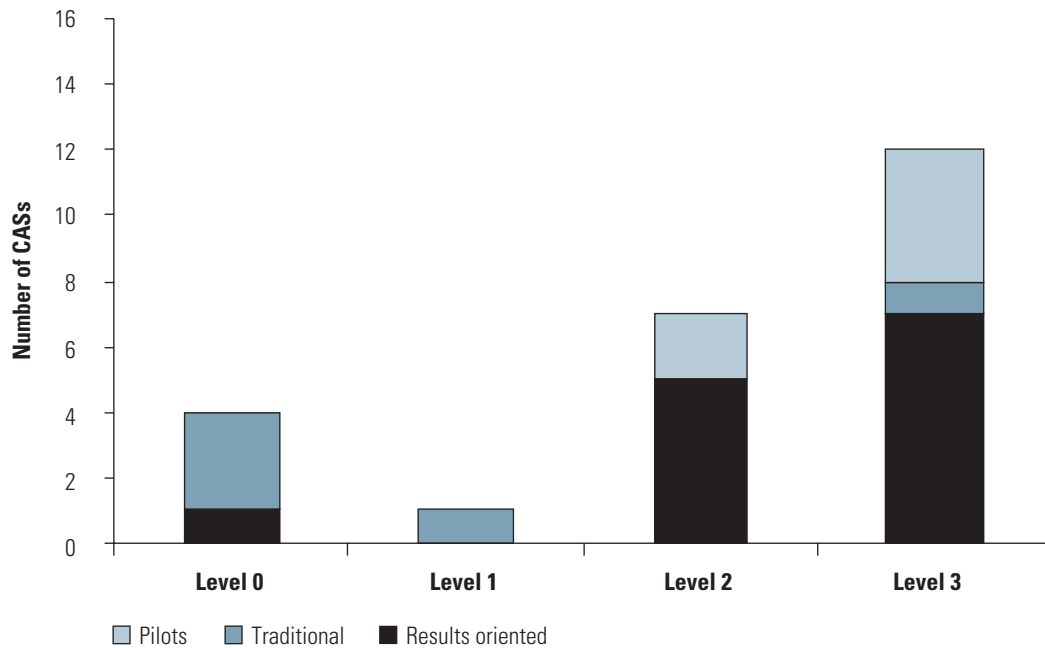


Figure A.7: Measurement Alignment: Country Level

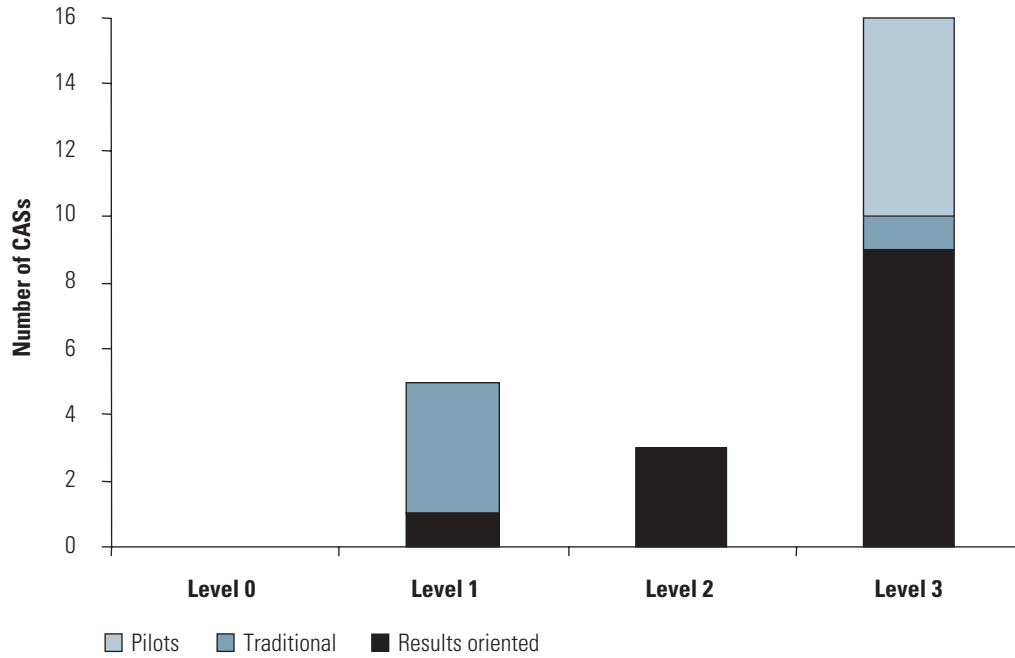


Figure A.8: Measurement Alignment: CAS Level

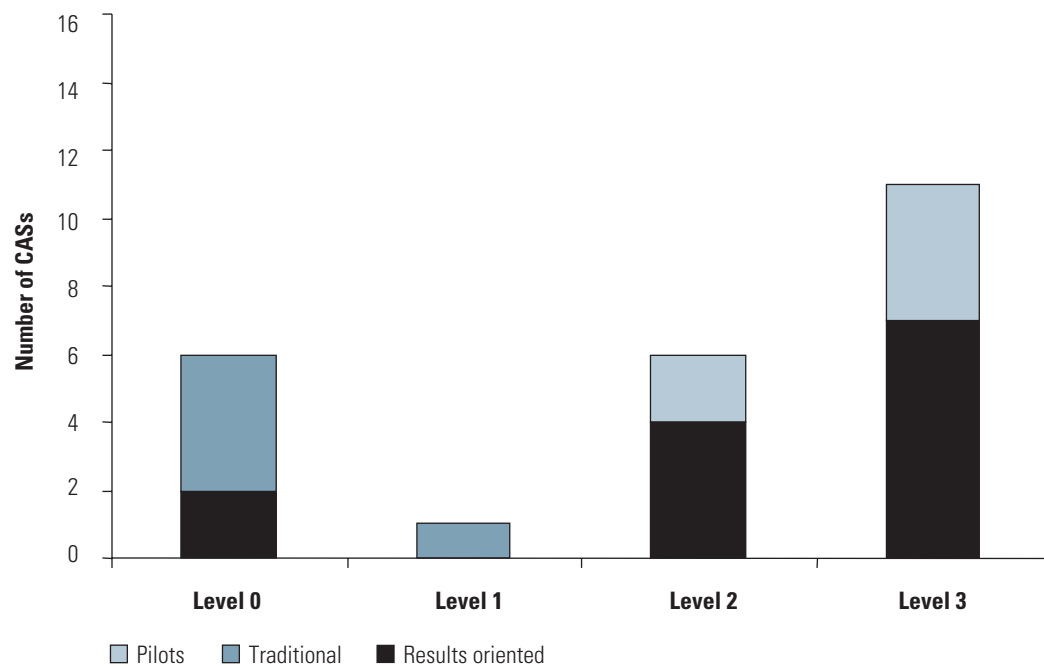
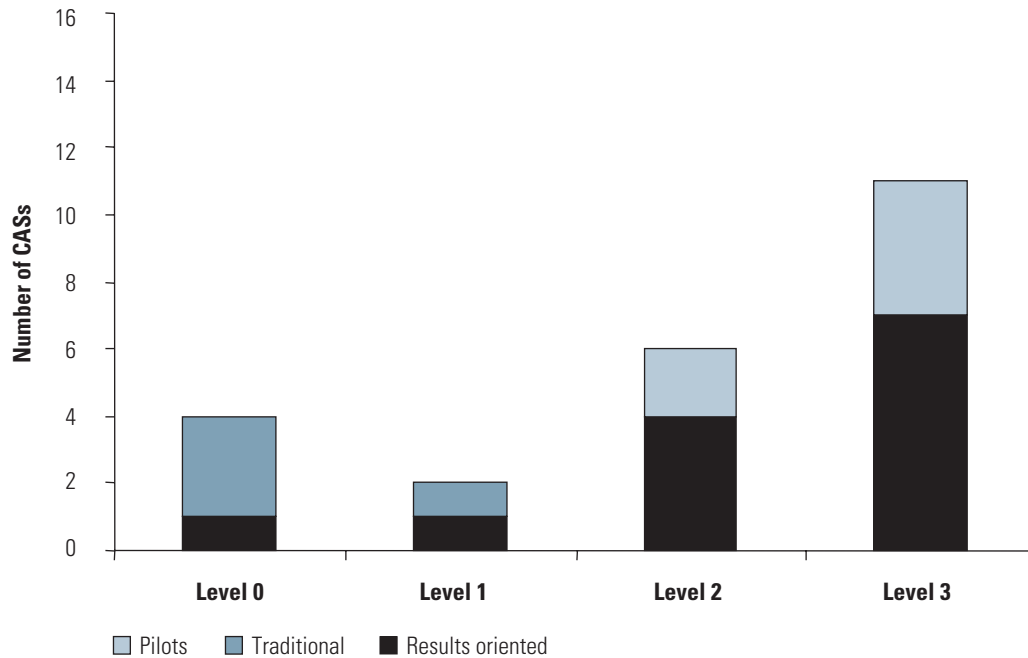


Figure A.9: Measurement Alignment: Intermediate Level

measures reinforces the idea that the country-level measures, typically founded on MDG or PRSP measures, are more developed, while the measures at the CAS and intermediate levels are relatively underdeveloped.

Problems in Baselines and Targets Point to Difficulties in Data

Performance measures were typically well defined, that is – most measures did a good job of specifying the evidence that would allow the Bank to understand progress (as discussed in the Logic Chain section, above). However, measures across the sample were much less likely to have baselines, an indication either that current data do not exist or that the data were not available to the Bank when the report was finalized. Also, many measures lacked targets. For example, “Increase living standard of working population measured as increase in real salaries” is a well-defined measure because it indicates specifically what evidence should be used to determine the desired effect. But it lacks both a baseline and a defined target. Measures such as this were more typically found at the CAS and intermedi-

ate levels, but were also prominent at the country level.

The pattern was generally consistent across country, CAS, and intermediate levels, although the differences were greater at the CAS and intermediate levels, as noted above. The evaluations for targeting, baselining, and definition for all three levels were added up and compared by maturity level (figure A.11). Measures were better targeted, with close to three-quarters (73 percent) assessed at Level 2 and 3. The opposite is true of baselining and targeting. Measures were slightly more likely to have targets than baselines, but this area also demonstrated a lack of maturity relative to the definition of the measures.

The Number of Measures

Results matrices often contained a large number of performance measures, raising doubts about the usefulness of many measures as a management tool. The average number of measures (excluding the traditional CASs) was 22 for country development, 39 for CAS outcomes, and 40 for Bank intermediate measures for the CASs reviewed for this report (table A.1). The number of measures

Figure A.10: Comparison by Type of Measure

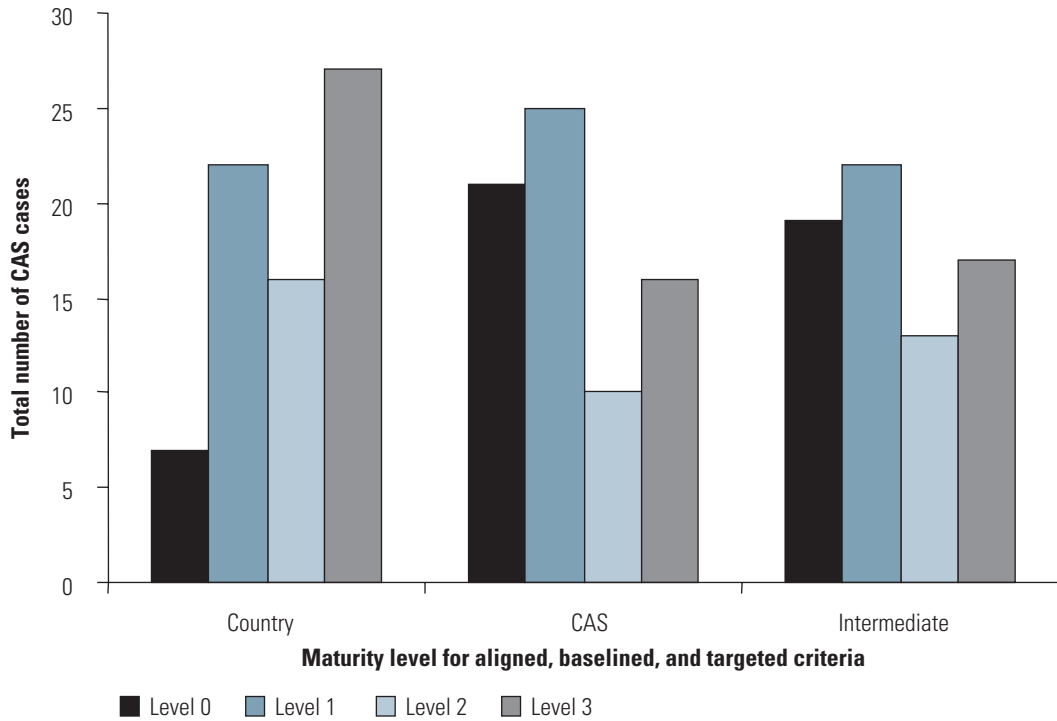


Figure A.11: Evidence of a Data Gap

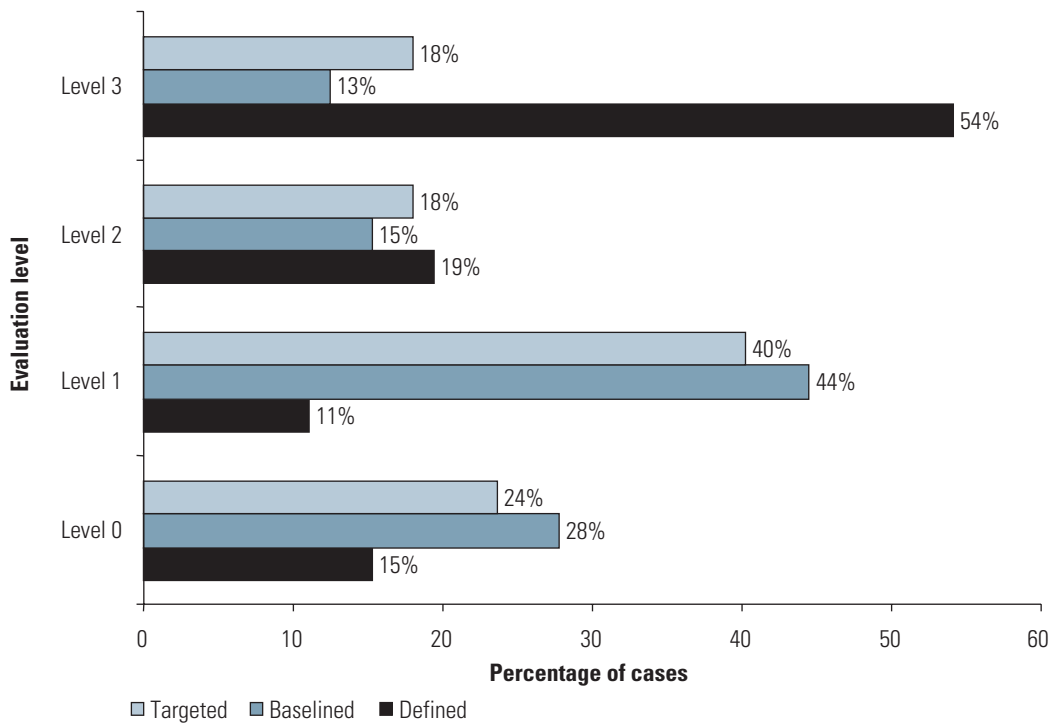


Table A.1: Number of Performance Measures by Type

	Average	Low	High
Country	22	7	38
CAS outcomes	39	15	80
Intermediate outcomes	40	14	97

should reflect country program needs, so that large and complex programs may require a large number of measures. While there are no criteria for a priori determining a maximum, a large number of measures would tax the government's ability to collect and analyze the information. Some CASs have addressed this by showing a large number of measures in their results matrices but presenting a smaller number of core measures in the main text of the CAS.

Attachment A.1: CAS Results Matrix—Definitions and Excerpt from Cameroon CAS

Country development goals

Definition: Longer-term or higher-order development objectives identified by the country. Usually not achievable in the CAS period or based solely on the CAS program. Only those to which CAS outcomes will contribute are included.

Issues and obstacles

Definition: Critical issues and obstacles to achieving country development goals, providing the logical link to CAS outcomes.

Reduction in rural poverty

- Decrease in the poverty incidence in rural areas from 49.9% (2001)
- Maternal mortality rates decrease from 430 per 100,000 (1998) disaggregated by rural area
- Child mortality decreases from 77 per 1,000 (1998) disaggregated by rural area
- 100% primary school completion rates by 2015, from 56% (2001) disaggregated by rural area

Determinants of rural poverty

- Lack of education among heads of households is a key determinant of rural poverty
- Relative isolation (enclavement) due to a lack of physical infrastructure (mainly roads) and no adequate support for road maintenance
- Isolation from social services (health and education especially in rural areas)
- Low level of coordination between ministries in charge of rural water, and low local capacity and skills for water supply infrastructure
- 20% of the poor have no financial social coverage

Source: World Bank 2005a.

CAS outcomes	Milestones	Bank program (and partners)
<p><i>Definition:</i> Results achievable in the CAS period and that the Bank expects to influence through its interventions. Indicators of each outcome are included, with baselines and targets.</p>	<p><i>Definition:</i> Progress markers of implementation of the Bank's program; outputs, actions, or outcomes expected to be realized during CAS implementation.</p>	<p><i>Definition:</i> Ongoing and planned lending, grants, and guarantees; analytical and advisory activities. Includes IBRD, IDA, IFC, MIGA. Partners included if cofinancing or other support of same CAS outcome.</p>
<p>Improved delivery of basic services (sustainable management of service delivery at the community level for rural communities in four targeted provinces):</p> <ul style="list-style-type: none"> • Percent of rural communities in the four provinces with access to basic socioeconomic services (baseline in fiscal 2004), including: <ul style="list-style-type: none"> – Increase in access to water (baseline: 58%) – Decrease in time to get to school. – Increased use of rural health care clinics • Percent of communities in the four provinces that have functioning elected committees 	<ul style="list-style-type: none"> • 50% of the 4,000 rural communities targeted have fully implemented one micro-project (roads, schools, hospitals, etc.) by 2006 • Over 4,000 rural community-based action plans designed and implemented between 2003 and end 2005 • Increased involvement of women and vulnerable groups in decision making, definition of priority action plans, and implementation oversight 	<p>Lending - Ongoing</p> <p>Ag Serv. and Ext. (PNVRA) ; Cameroon Petroleum Environment Capacity Enhancement (CAPECE) ; Douala Infrastructure</p> <p>Lending - Proposed</p> <p>FY04 : PNDP (community development). FY06 : PRSC - Rural development</p> <p>AAA - Ongoing</p> <p>Expenditure tracking and beneficiary assessments (FY03)</p> <p>AAA - Proposed</p> <p>FY04 : Health Sector Reform Note; Gender strategy (IDF) FY06 : Poverty Assessment</p> <p>PARTNERS</p> <p>Cofinancing of PNDP by AFD, IFAD, KfW and Netherlands cooperation</p>

Attachment A.2: Detailed Maturity Model and Evaluation Criteria

Criterion	Definition
Logic chain	The results matrix schematic structure effectively portrays the chain of causal assertions connecting country and Bank goals and activities. This would be expressed in a results chain that is the cascade of causal assertions between outcomes (country, medium, and short term), outputs, activities, and inputs. The results chain should be visible in the results matrix such that the links between country goals, Bank goals, and Bank activities are clear.
Results chain strength	The causal assertions made in the results chain are logically related and sensibly support country and Bank goals. The goals and activities of the Bank should logically connect to and support the goals of the country. Activities should be sensibly aligned to goals that they will affect.
Linkage to other donors	The activities of other donors and lenders, to the extent that those activities are reflected in the text of the CAS, should be referenced in the appropriate place in the results matrix.
Defined	The measure is clearly identifiable as a performance measure, distinct from a goal or program activity.
Baselined	The measure includes the current level of performance as a baseline.
Targeted	The measure includes a specified, time-bound performance target.
Aligned	The measure provides the appropriate level of measurement given its location in the results matrix. Outcome measures describe the accomplishment of changes in country practices or behaviors (for Bank outcomes), while output measures provide information on services delivered by the Bank in order to accomplish outcomes. Inputs and activity measures are used only to provide leading indicators of output or outcome measures where results are hard to measure or slow to develop.

Framework level			
Level 0	Level 1	Level 2	Level 3
	The matrix portrays the appropriate elements of a results chain, but the schematic portrayal does not effectively link the elements, or some elements are present and linked but significant elements are lacking.	The matrix contains and shows most of the appropriate elements of the results chain, but is lacking one or two critical elements or portrays some elements in a way that is unclear.	The matrix portrays the elements of the results chain in a way that clearly and directly demonstrates the linkage between country goals and Bank's results, as well as associated outputs, activities, and inputs.
Absence of evidence in the criterion area.	Most Bank goals and activities seem misaligned with the country goals, or a profusion of goals and activities buries the logical, causal relationships	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.	Most pillars or country goals are logically supported by succinctly stated Bank activities and goals.
	Other donor and lender activities are described in the CAS text but mostly ignored in the results matrix, or the activities of other organizations are included but misaligned.	Some country goals have supporting activities effectively associated but others do not, or some activities are placed appropriately but some key activities described in the text are missing or improperly aligned.	Almost all of the key activities of other donors and lenders included in the CAS text are cited in the appropriate country or Bank goal of the results matrix.
Very few measures in the matrix demonstrate the defined qualities.	Fewer than half of the measures demonstrate the defined qualities.	More than half but fewer than three-quarters of the measures in the matrix show the defined qualities.	More than three-quarters of the measures in the matrix exhibit the defined qualities.

APPENDIX B: COUNTRY DIRECTOR SURVEY RESULTS ON THE USE OF MONITORING AND EVALUATION SYSTEMS AND INFORMATION AT THE COUNTRY LEVEL

This appendix summarizes the results of a survey of selected country directors carried out as part of the 2005 AROE in March 2005. The objective of the survey was to obtain information on the use and usefulness of monitoring and evaluation systems and information at the country level for Bank country program management. The questionnaire covered the processes to review progress toward CAS outcomes, information systems, and the availability and usefulness of management information. The survey was sent electronically to country directors of countries that had submitted Country Assistance Strategies (CASs) to the Board in fiscal 2004 and 2005 (as of end-January 2005). These country directors were selected because most of these CASs followed or incorporated elements of the results approach.

A total of 26 country directors were surveyed.¹ Responses were received from 16 (62 percent) of the country directors (3 from staff designated by the country director). The numbers received and response rate² by Region and country type are shown in table B.1.

Progress toward CAS Objectives

The country directors were asked what regular processes they had in place to monitor and review progress on the CAS: (a) within the Bank, (b) with the government, and (c) with other donors. All the respondents replied that they regularly review progress toward CAS objectives with their country teams through work programming, stock-taking (such as Country Portfolio Reviews), or regular country team meetings. Most (87 percent) respondents discussed progress toward CAS objectives with country counterparts. These discussions were conducted as part of regular consultations and program and portfolio reviews. In low-income countries, and Poverty Reduction Strategy Paper (PRSP) and Credit (PRSC) reviews were vehicles for these discussions. In one middle-income country, the country director expected to use programmatic adjustment loans (a series of Development Policy Loans) as the primary vehicle for dialogue with the government. Fewer respondents (64 percent) replied that they discuss progress toward CAS objectives regularly

Table B.1: Summary of Responses

By Region	Received	Rate (%)	By country type	Received	Rate (%)
Africa	7	78	Middle income	8	53
East Asia and Pacific	1	50	Low income	8	73
Europe and Central Asia	5	71	Total	16	62
Latin America and the Caribbean	1	20			
Middle East and North Africa	1	100			
South Asia	1	50			
Total	16	62			

with other donors. The discussions were held as part of joint Bank and donor reviews with the government, such as the PRSP review, or regular donor consultations.

The respondents discussed results, both of the CAS and Bank operations, with their country teams. When asked to rate the extent to which they took up various operational items with their country teams (with 1 equaling “not at all” and 5 equaling “very much”), the country directors in the survey responded that CAS outcomes and project approvals were the most frequently discussed items, both with an average of over 4 on a 5-point scale (figure B.1).

The averages mask differences in approach: some country directors placed greater emphasis on project deliveries, while others focused more on CAS outcomes and less on the particulars of the work program. For example, one country director agreed to milestones toward CAS objectives with his country team, and these milestones became the focus of reviews during the fiscal year. Most country directors adjusted

project and ESW deliveries following the discussions with their country teams.

Respondents’ discussions with the government were more focused on projects. The country directors in the survey responded that the status of the project portfolio was the main item of discussion, followed by project readiness (the quality of project preparation) and CAS outcomes (Figure B.2). This result most likely reflects the venue for these discussions: that is, PRSP/PRSC reviews, programming discussions, and Country Portfolio Performance Reviews.

Information Systems

While many respondents rely on the Bank’s SAP/Business Warehouse for information to manage their country program, many have developed their own systems to consolidate necessary information from Bank systems or to supplement it to meet their own requirements. When asked to rate whether they used the Bank’s SAP and Business Warehouse for information in conducting their reviews with the country and the

Figure B.1: Discussions with Country Team

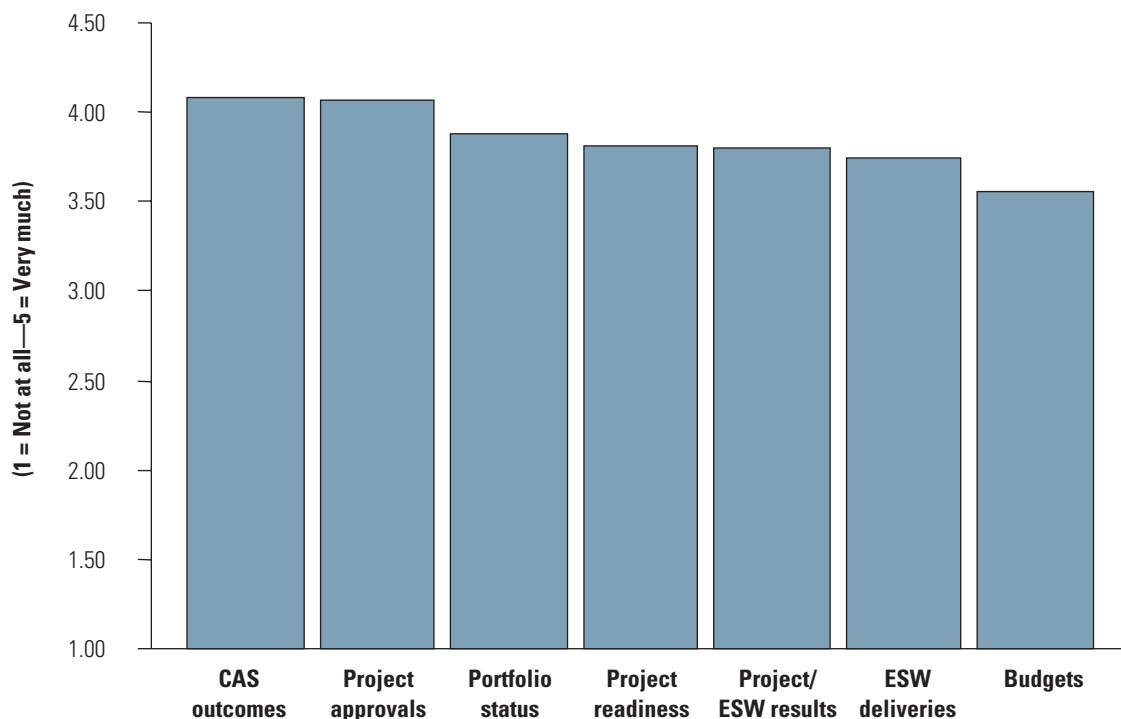
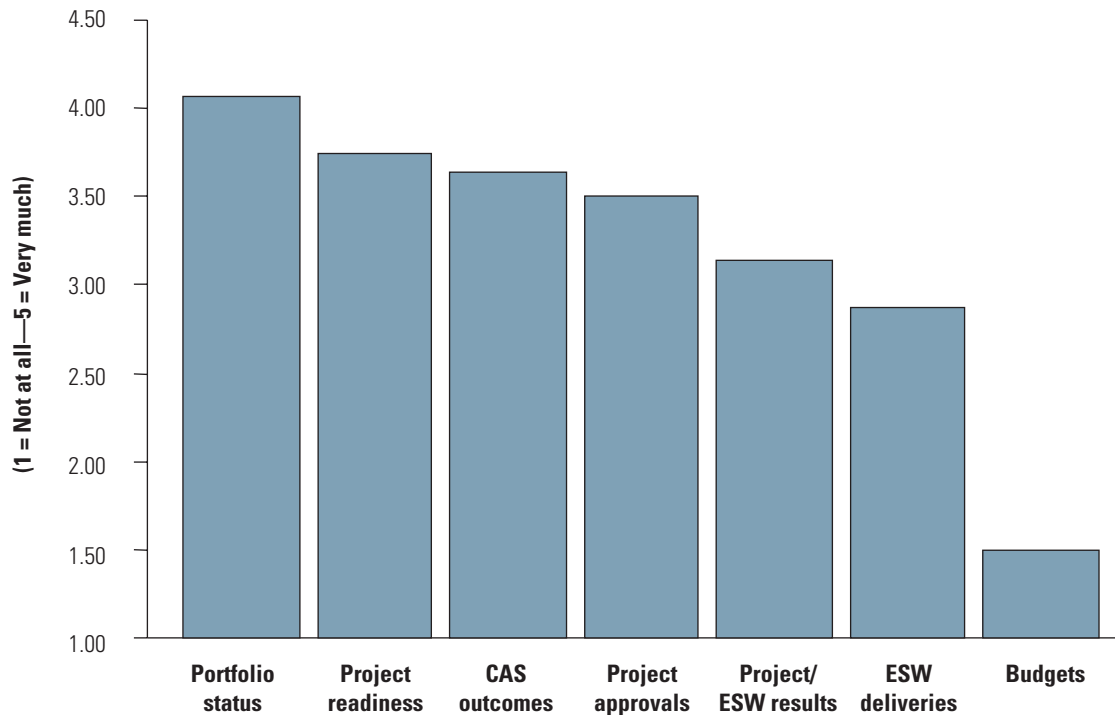


Figure B.2: Discussions with Government

government, the responses ranged from 1 (not at all) to 5 (very much), with a mean of 3.5 (with a coefficient of variation of 35 percent). Sixty-seven percent of respondents indicated that they were able to obtain the information needed from the Bank's information systems. However, 12 respondents (75 percent) have packaged the information to better meet their needs. Some have consolidated data from different Bank systems into their own spreadsheet or have included outcome information, such as that from PRSP reviews, which is not available from the Bank.

The respondents have yet to develop common systems or processes with the government or donors to jointly develop or share information for monitoring progress on meeting CAS objectives. Five Country Directors in the survey responded that they have a common system or process with the government, and only one responded that they have a system with donors. Several respondents indicated that they were in the process of developing a joint system with the government – for example, through an IDF grant. The lack of government capacity or lack of in-

terest was often cited as the reason for not developing a joint monitoring system. Some respondents commented on the effort needed to develop a common system: “It takes time and effort and requires sustained government attention as well as Bank dedication to the task.” Lack of interest was the most common reason given for not developing joint systems or processes with other donors. Several country directors in the survey mentioned that they intended to develop a common system or process with other donors, most often through the PRSP.

Availability and Usefulness of the Management Information Available

The country directors were asked to rate, from the following list of five possible uses of information, which they considered important for managing the Bank's country program, and which they thought were satisfactory in terms of the information currently available to them.

- Communications—describing Bank country performance

- Accountability—assigning responsibility for success or failure for Bank operations
- Learning—developing alternative strategies or approaches based on success or failure
- Business planning—allocation and adjusting human and financial resources
- Modifying operations—redesigning projects, revising strategies.

They also were given the opportunity to raise other important uses of information not included in the above list, but no other major uses were identified.

The respondents rated learning as the most important use of information (figure B.3). They were in relative agreement on this point (with a coefficient of variation of 15 percent). Learning was followed by business planning and modifying operations. Accountability was considered to be less important. Several commented on the difficulty of determining the Bank’s contribution (or lack thereof) to an outcome as a reason for giving accountability a lower rating.

Overall scores for satisfaction with the information the respondents have available were lower: the highest was 3.7 for modifying operations (figure B.4).

General Comments

When asked to comment on the available information and the need for improvements, several respondents mentioned the need for a more systematic approach to making country-level information available,

- “Standardized results measurement equivalent to key macro indicators should become part of the normal reporting from sector departments.”
- “It would be ideal to have the system generate monthly country specific reports on the portfolio performance, quality, disbursements, etc., to avoid all the manual handling of information.”
- “Information is scattered and not readily available in one place. Suggested Country Director

Figure B.3: Important Uses of Information

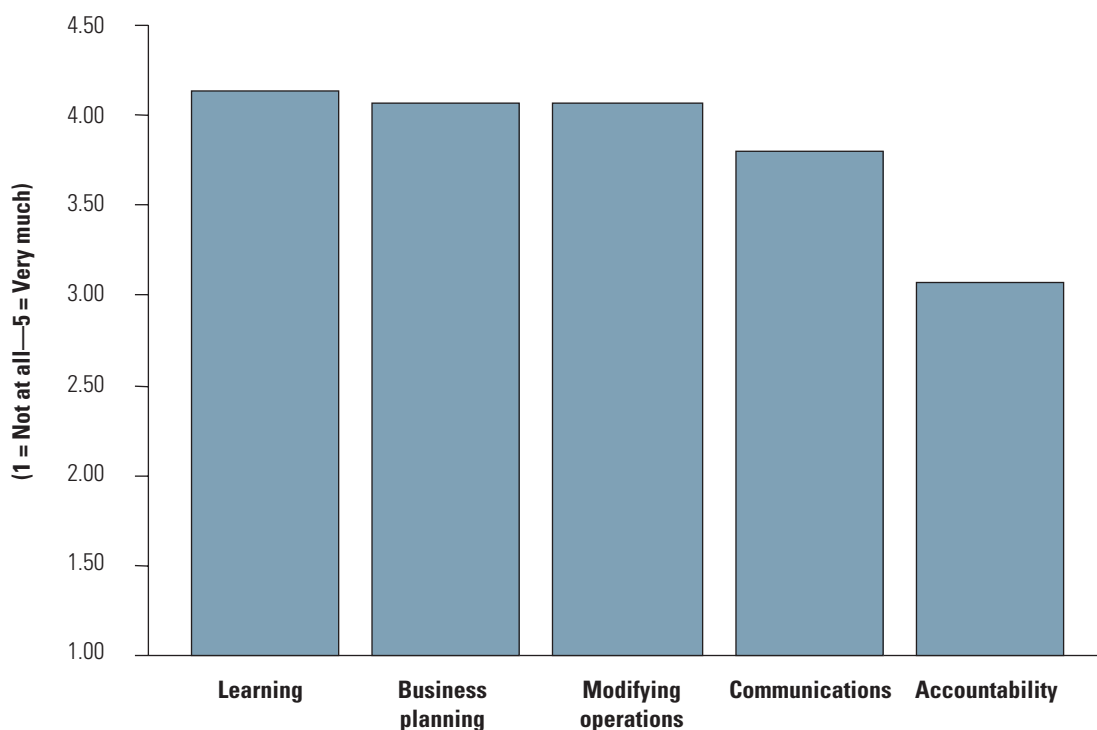
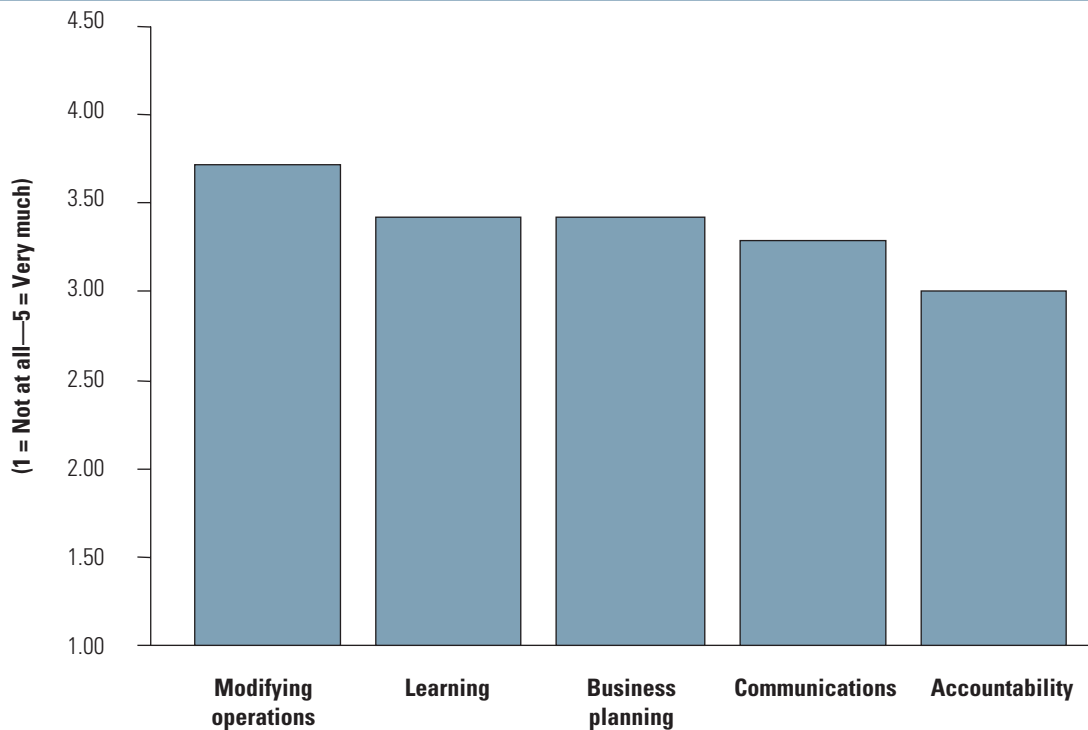


Figure B.4: Satisfaction with Information

Portal would be a welcome introduction as it would allow [us] to organize all of the information in one place.”

- “Quicker feedback on ground level impact would help.”
- “[Information] adequate, but presentation is not always user friendly; information relevant to manage the portfolio had to be developed ‘in-house’ since most of what comes from DC is hard to use for action.”

Conclusions

In sum, respondents reported that they see reviewing progress toward CAS objectives as part of their internal program reviews with their country teams and with the government. Project implementation and readiness feature more prominently in their discussions with the government, but this may be influenced by the context, PRSP/PRCS reviews, and portfolio reviews.

Generally, the respondents indicated they were able to obtain information needed to manage their country programs from Bank systems, but many have developed their own systems to

synthesize information available from Bank systems and from operations (such as PRSP reviews). Several commented on the need for a more systemic approach to data collection at the country level.

But only five (31 percent) of the respondents indicated that they have a common system or process with the government to jointly develop or share information for monitoring progress on meeting CAS objectives. Most existing systems and processes involved the PRSP. Several respondents indicated that they were developing a joint system with the government through an IDF grant or other means. Lack of government capacity or lack of interest were often cited as the reasons for not developing a joint monitoring system.

The respondents saw learning, business planning, and modifying operations as relatively more important uses of information compared with communication and accountability. That is, they report seeing information as useful for managing and improving their program more than for reporting and enforcing accountability.

APPENDIX C: MONITORING AND EVALUATION FRAMEWORK IN THE BANK

Table C.1: Framework for Monitoring Bank Operations

Type of intervention or financing instrument	Defining outcomes and monitorable indicators	Review and quality assurance	Tracking and reporting on implementation progress	Managing implementation progress
Loan/credit/grant operations	Project Appraisal Document (PAD) (revised) or Memorandum & Recommendation of the President	Management and peer reviews, Quality Enhancement Reviews (QER), and annual QAG review of quality at entry ^b	Implementation Status and Results Report (ISR) (new)	Mid-Term Review. Country Portfolio Performance Reviews (CPPR), annual QAG reviews of supervision (QSA)
ESW/AAA	Concept Paper	Peer and management review, QER (few), QAG ESW reviews QAG Country-level AAA Assessments (new)	SAP monitoring	Peer review, management review, QAG ESW reviews
Trust funds ^c	Initiating Brief for trust funds	VPU, TFO review	Trust Fund Status Reports (new) Implementation Completion Memorandum (ICM), (new)	
Sector strategy	Sector Strategy Implementation Update (SSIU) (new)	QAG, network, and OPCS review	Sector Strategy Implementation Update (SSIU)	Sector Board (SB) reviews
Country program ^d	Results-Based Country Assistance Strategy (Results-Based CAS) (new)	SB, network, and OPCS review	CAS Progress Report (CASPR) IDA14 Results Framework (new)	CPPR, management reviews
Grants/global programs	Grant proposal	Task manager review	Financial reporting	Task manager review

Note: ESW = economic and sector work; SAP = Systems, Applications, and Products in Data Processing; TFO = Trust Fund Operations.

a. Grants and Global Programs and Knowledge Initiatives are discussed in Chapter 3. One other means of financing not discussed here is guarantees.

b. QAG's quality at entry assessment is undertaken after project approval

c. Trust funds are a source of funding that is used to supplement Bank financing of various Bank and recipient development activities.

d. Country program support includes CAS, PRSP, JSA, HIPC, and other development services.

Table C.21: Framework for Evaluation in Bank Operations

Type of intervention or financing instrument	Self-evaluation		Independent evaluation (IEG)	
	Individual activity	Aggregate level	Individual activity	Aggregate level
INSTRUMENT				
Loan/credit/grant operations	Implementation Completion Reports (ICRs)	Lending retrospectives	Evaluation summaries, Project Performance Assessment Reports, Impact Evaluations	Annual Review of Development Effectiveness (ARDE)
Economic and sector work/analytical and advisory services	Activity Completion Summaries (ACS) for tasks > \$50,000	Regional retrospectives (<i>ad hoc and rare</i>)	Periodic studies (such as Public Expenditure Reviews, Poverty Assessments)	Sector and thematic studies, implementation reviews, and Country Assistance Evaluations
Trust funds	Implementation Completion Memorandum (ICM) > \$1 million	Annual Review of Trust Fund Portfolio Performance, some donor reports		
Sector strategy	n.a.	n.a.	Sector and thematic studies	Annual Report on Development Effectiveness
Country program	CAS Completion Reports (<i>new</i>)	CAS Retrospectives	Country Evaluation Summaries (<i>new</i>), CAEs	Annual Report on Development Effectiveness, Regional IEG Evaluations (CFA, transition economies)
Grants/global programs	Grant Completion Reports (GCR)	DGF evaluations for >\$300,000	Global Program Reviews (pilot)	IEG evaluation of global public programs
KNOWLEDGE INITIATIVES				
Knowledge management	For few knowledge sharing activities			Evaluation of Knowledge Sharing (IEG) 2003a)
DEC research	Research Project Completion Reports	2001 Research Support Budget (RSB) evaluation		
World Bank Institute training and capacity building	Individual learning events	World Bank Institute evaluations of selected programs		

APPENDIX D: IEG CLIENT SURVEYS

Since 1999, IEG has surveyed its clients annually to gain 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, the influence of the report on Bank processes and policies, and recommendations made to IEG on improving the quality and relevance of the reports. This year IEG conducted two surveys: one of internal clients in the Bank and a second of external clients. This appendix reports on the findings from these two surveys.

Results of Survey of Internal Clients

This year an internal survey of 2,120 Bank staff sought to elicit their perceptions of several kinds of IEG products. The following IEG products were the subject of the survey:

- Country Assistance Evaluations (CAEs):
 - *Armenia: Country Assistance Evaluation*
 - *Bosnia and Herzegovina: Post-Conflict Reconstruction and the Transition to a Market Economy, An IEG Evaluation of World Bank Support*
 - *Brazil: Forging a Strategic Partnership for Results, An IEG Evaluation of World Bank Assistance*
 - *China: An Evaluation of World Bank Assistance*
 - *Croatia: Country Assistance Evaluation*
 - *Jordan: Supporting Stable Development in a Challenging Region*
 - *Rwanda: Country Assistance Evaluation*
 - *Republic of Tunisia: Country Assistance Evaluation*
- *The 2003 Annual Review of Development Effectiveness: The Effectiveness of Bank Support for Policy Reform*
- Project Performance Audit Reports (PPARs)
- Evaluation products:
 - *Books, Buildings, and Learning Outcomes: An Impact Evaluation of World Bank Support to Basic Education in Ghana*
 - *Evaluating The World Bank's Approach to Global Programs: Addressing the Challenges of Globalization*
 - *The Poverty Reduction Strategy Initiative: An Independent Evaluation of the World Bank's Support Through 2003*
 - *Economies in Transition: An IEG Evaluation of World Bank Assistance.*

Two primary audiences were targeted by this year's survey:

- Bank staff groups associated with the products
- Executive directors (EDs).

Target respondents were surveyed only once, even if their work was relevant to more than one assessed product. The following is a summary of the results. This year's results are compared with results from previous years and similar types of respondents, where possible.

Response Rates

The surveys were sent electronically to 2,120 staff over a two-week period in April 2005. The response rates ranged from 23 percent (for executive directors) to 46 percent (for the Armenia CAE), and averaged 31 percent across all survey products. This represents an increase in response rate relative to the 2004 surveys (21 percent). The response rate was boosted through the use of multiple reminder notifications. Reminders directly targeted non-respondents.

Figure D.1: Response Rates

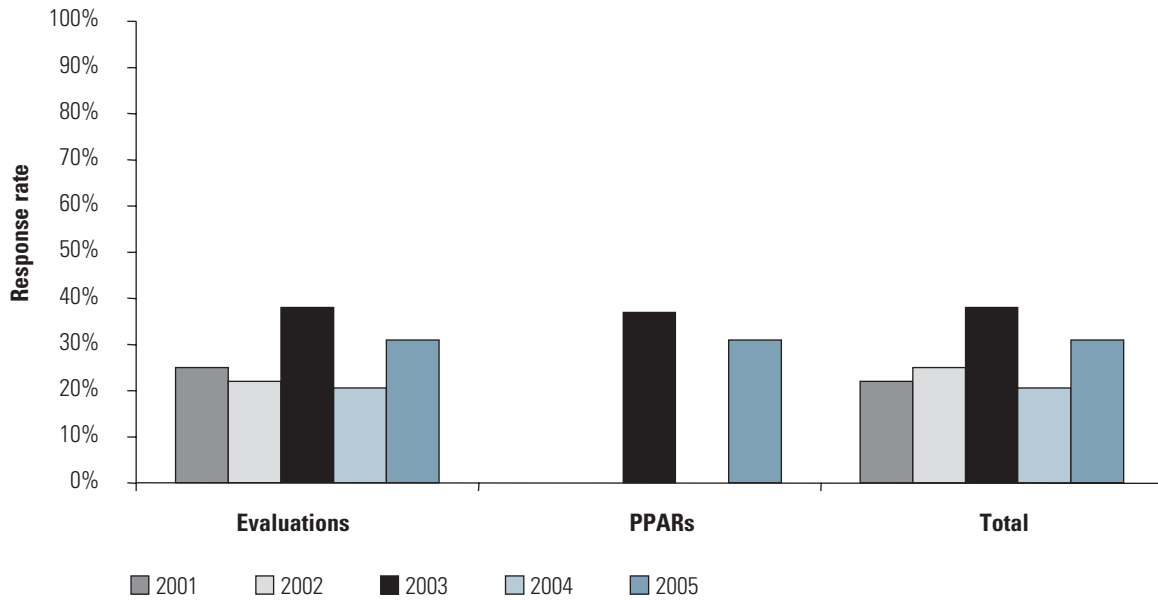
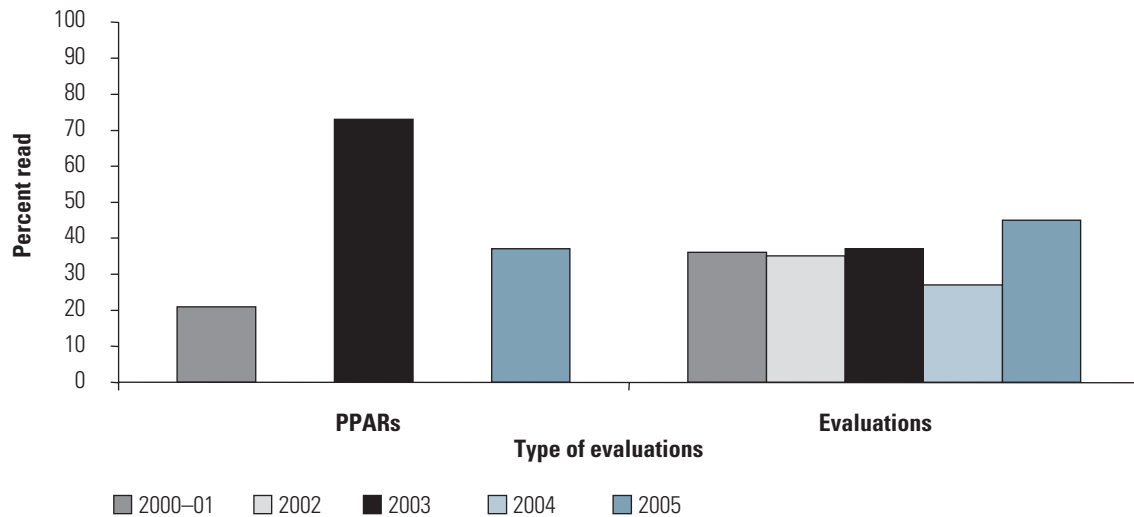


Figure D.2: Readership and Awareness of IEG Evaluations



Survey response rates are not directly comparable to those earlier years, however. Prior to this year’s surveys, 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 survey, individuals were limited to a single record but

could re-enter that record until they had completed the survey to their satisfaction.

Readership and Awareness

Almost half (45 percent) of non-ED respondents had read, participated in a discussion, or attended an event on the evaluation for which

they were surveyed. This is a marked increase over the levels of awareness found in previous AROE surveys. An additional 13 percent of respondents indicated that they were aware of the findings, even though they had not read the study or participated in any dissemination events. The level of awareness varied by product readership, from a low of 15 percent to a high of 79 percent awareness. Executive directors were asked if they had read, participated in a discussion, or attended an event for any 2004 evaluation, and 85 percent said that they had.

For non-ED respondents who were aware of the evaluation, half (52 percent) heard about it online or via email. Thirty-five percent were consulted during the evaluation. For executive directors, three-quarters (74 percent) of them heard about these evaluations at CODE meetings, and almost two-thirds (64 percent) through print copies.

Over three-quarters of non-ED respondents said that their evaluations were timely (89 percent), made full use of available information (74 percent), and used a methodology that was transparent and clear (78 percent). Similar to other respondents for a specific evaluation, executive directors reported positively for large

percentages for evaluations in general for these three measures (80 percent, 81 percent, and 82 percent respectively). A comparison with prior years shows that evaluations have consistently been timely.

Respondents were asked to assess several aspects of their evaluation (or, for executive directors, evaluations in general). Figure D.4 shows the proportion of respondents that rated each aspect of the evaluations as either excellent or good.

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

- Depth of analysis (74 percent good or excellent)
- Presentation of the evidence (70 percent)
- Link between conclusions and evidence (62 percent, but 10 percent rated this aspect as poor)
- Presentation of conclusions (75 percent)
- Relevance of lessons to Bank's work (69 percent, with 21 percent rating this aspect as excellent).

In general, respondents to this year's surveys rated the evaluations the same or just slightly less

Figure D.3: Timelines of IEG Sector and Thematic Evaluations

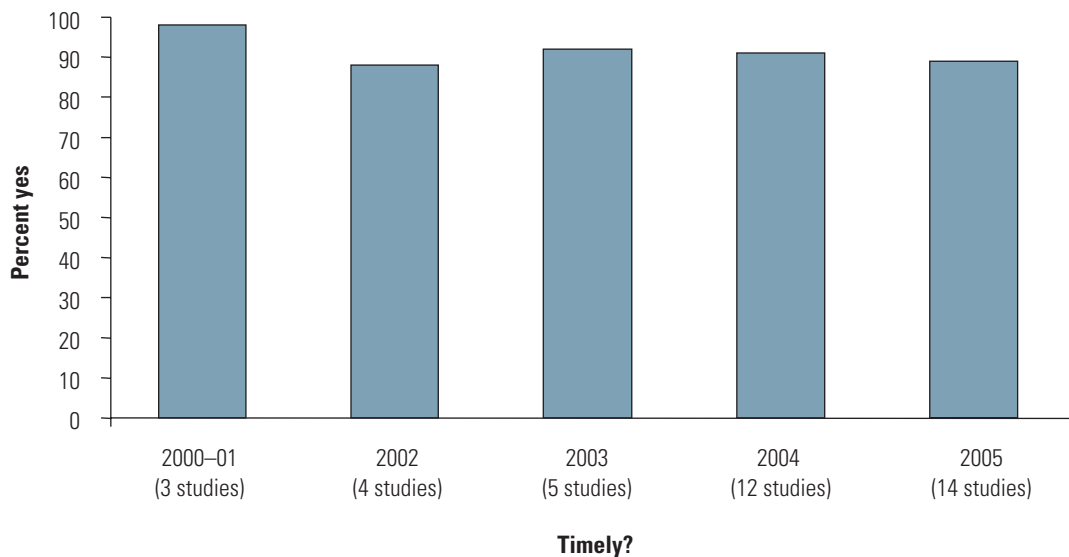
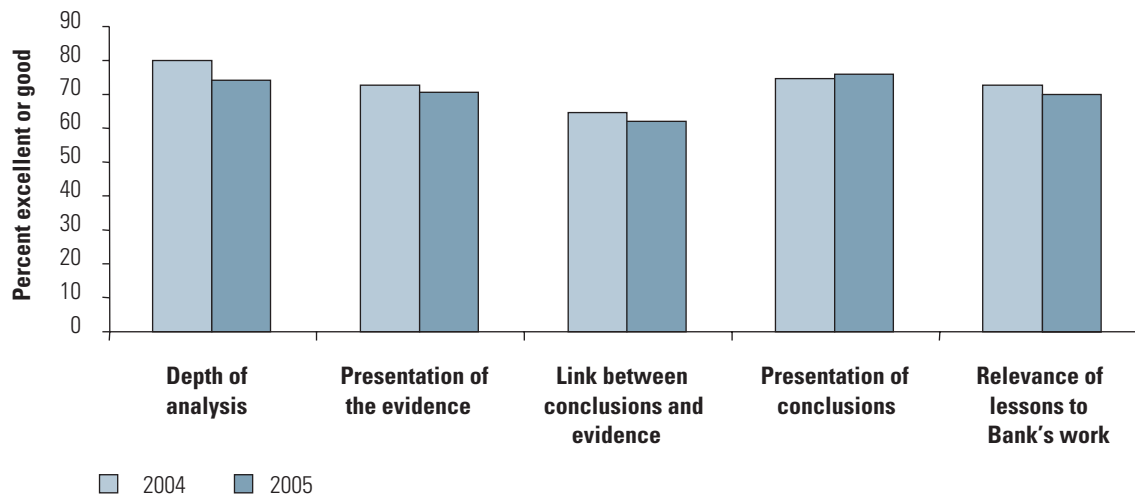


Figure D.4: Quality of Studies



positively than for the 2004 surveys, with less than a three percentage-point average movement from excellent or good to fair or poor across the five rating scales.

Overall, the executive directors rated evaluations more positively than did the non-ED respondents, with at least 84 percent good or excellent ratings for each aspect. Most positively rated was depth of analysis (95 percent good or excellent, including 24 percent excellent). The only aspect rated as poor by any ED respondents was the link between conclusions and evidence. This aspect also received the lowest excellent rating (8 percent) among executive directors.

Influence of IEG Evaluations

Over three-quarters (78 percent) of non-ED respondents said that their evaluation influenced their view of the associated subject area either a great deal (19 percent) or somewhat (59 percent).

Among executive directors, almost all (95 percent) said that IEG evaluations influenced their view of a subject area, with 40 percent reporting a great deal of influence.

Only 12 percent of non-ED respondents said they had not and would not use the findings of the evaluation in their work. Seventy percent said they had or would use the evaluation's findings, while the remaining 18 percent did not know.

Executive directors were asked if they had used, or planned to use, the findings from any IEG evaluation in their work. Ninety-five percent said they had or would use the findings, with the remaining 5 percent indicating "don't know."

In a pair of related questions, non-ED respondents were asked to what extent their evaluation influenced the Bank's country or sector strategy, and to what extent their evaluation influenced the Bank's lending and nonlending services. The respondents' answers regarding influence were similar overall for strategy and services. Almost half (47 percent) of the respondents in each case said that the evaluation somewhat influenced the country or sector strategy or the lending and nonlending services.

Results by Demographic Profile

IEG checked for differences in perceptions of IEG work by staff level (manager versus non-manager) for Sector and Thematic evaluations, and by location (headquarters versus field-based) for both sector and thematic and PPAR evaluations. The findings:¹

- Indecision or lack of knowledge about using the evaluations was greatest among respondents who were of level F and below (31 percent versus 16 percent for those of level G or higher). Among country directors and man-

agers, 11 of 12 respondents (92 percent) said that they would use the findings of the evaluation in their work.

- With regard to evaluation influence on country or sector strategy, more lower-level respondents (12 percent of those level G and under) than higher-level respondents (4 percent of level H and above) said there was a great deal of influence.
- Respondents in country offices recommend more training or educational materials than those in Headquarters as a means of improving outreach.

non-ED respondents). Nearly half (46 percent) of non-ED respondents also suggested that broadening consultation with Bank staff could improve their evaluation.

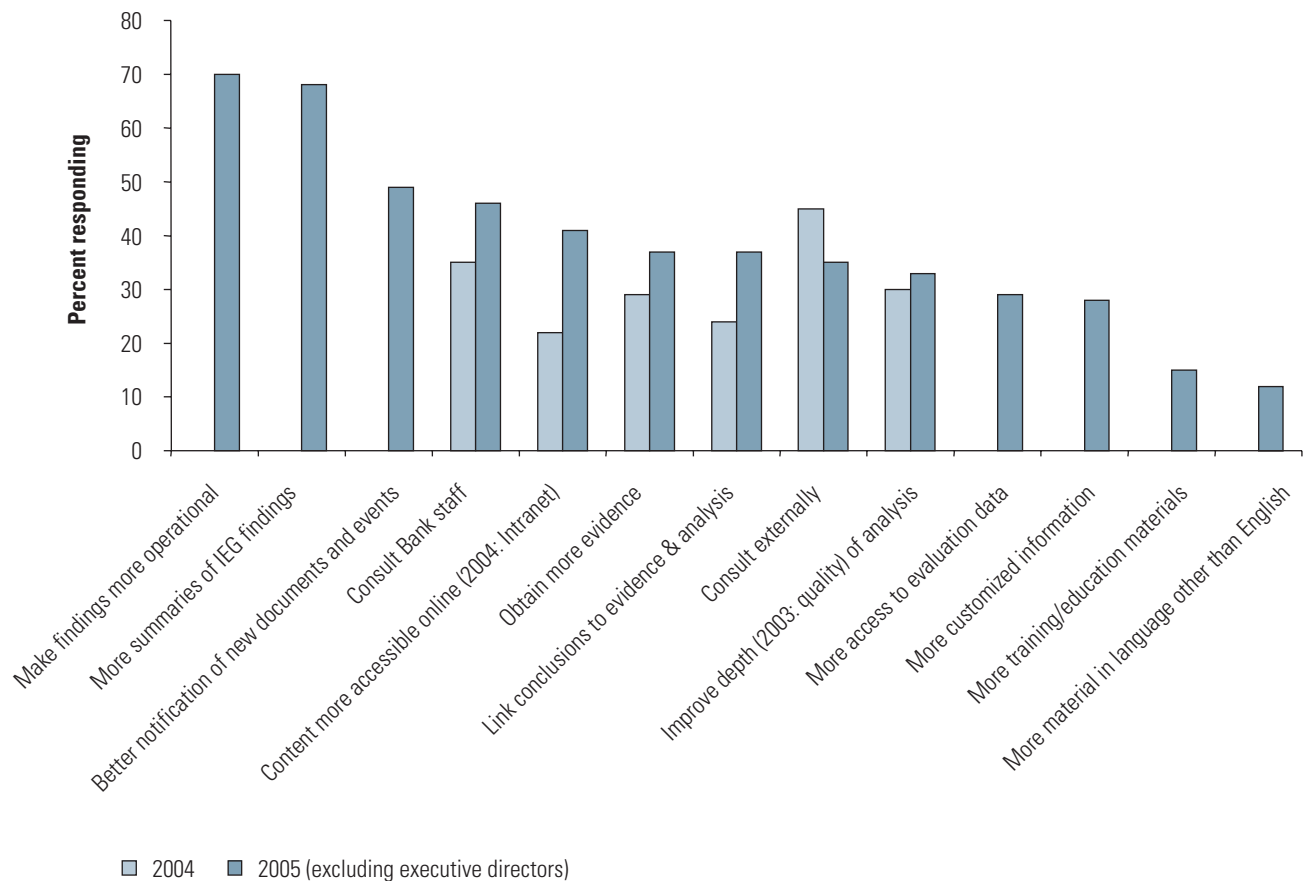
Executive directors were asked to select from the same set of recommendations for IEG evaluations in general. A full 87 percent of ED respondents said they would make the findings more operational. The ED respondents differed from non-ED respondents in their lack of enthusiasm for broadening consultation with Bank staff, with only one in five (21 percent) selecting this potential improvement.

ED and non-ED respondents were asked to select recommendations that they thought would improve IEG’s outreach. Respondents generally called for summaries of IEG findings to be provided and for better notification of new documents and events.

Recommendations from Respondents

When asked what recommendations they would make to improve their associated evaluation, the recommendation most often selected was to make findings more operational (70 percent of

Figure D.5: Recommendations for IEG



Results of Survey of External Clients

This year's external survey queried 926 external clients, including government officials, consultants, and researchers. They were asked about two IEG product types:

- Two evaluation papers associated with the International Program for Development Evaluation Training (IPDET):
 - *Monitoring and Evaluation: Some Tools, Methods, and Approaches*
 - *Influential Evaluations: Evaluations that Improved Performance and Impact of Development Programs*
- One evaluation associated with Global Programs and Partnerships (GPP):
 - *Addressing the Challenges of Globalization: An Independent Evaluation of the World Bank's Approach to Global Programs*

A pool of potential survey respondents was selected for each product type: IPDET and GPP. The IPDET pool was randomly split into two equivalent-sized subgroups corresponding to each individual evaluation paper: IPDET1 for the *Monitoring and Evaluation* (M&E) paper and IPDET2 for the *Influential Evaluations* paper. The following is a summary of the results.

Response Rates

The surveys were sent electronically to the reported e-mail addresses of 926 external clients, including government officials, researchers, and consultants during a two-week period in April 2005. The overall response rate was 31 percent. Disaggregated, the response rates were 28 percent for GPP; 34 percent for IPDET1; and 33 percent for IPDET2. The response rates were similar to those for the 2005 internal client survey. Response rates were boosted through the use of multiple reminder notifications that directly targeted non-respondents.

Readership and awareness. Two-thirds (66 percent) of all respondents read, participated in a discussion, or attended an event on the evaluation or evaluation paper for which they were surveyed (see figure D.7). The level of awareness varied somewhat by population, from 58 percent awareness for GPP to 73 percent awareness for IPDET2.

The vast majority of IPDET respondents (86 percent of IPDET1 and 94 percent of IPDET2) heard about the evaluation paper through email or online. Also, half (53 percent) of GPP respondents heard about the evaluation this way. Half of GPP respondents also said they heard

Figure D.6 : Response Rates

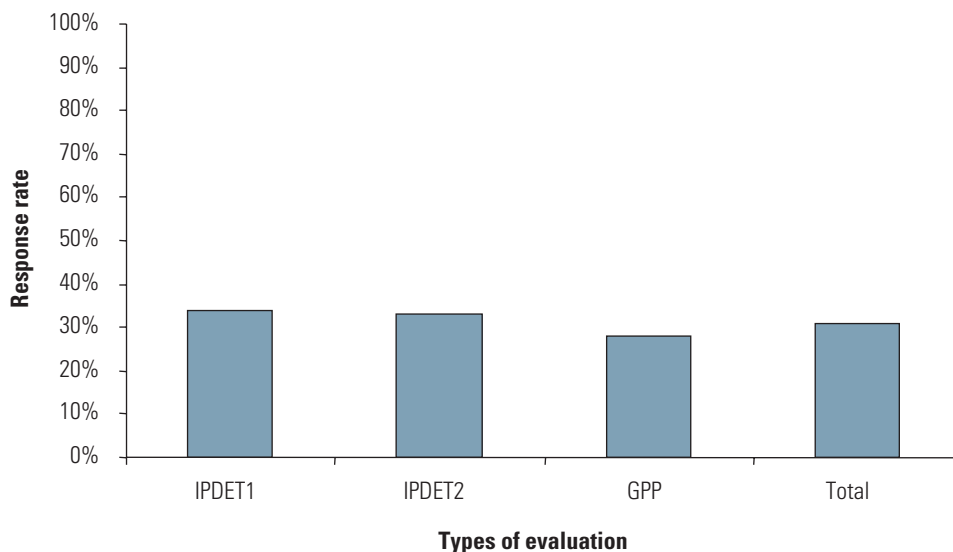
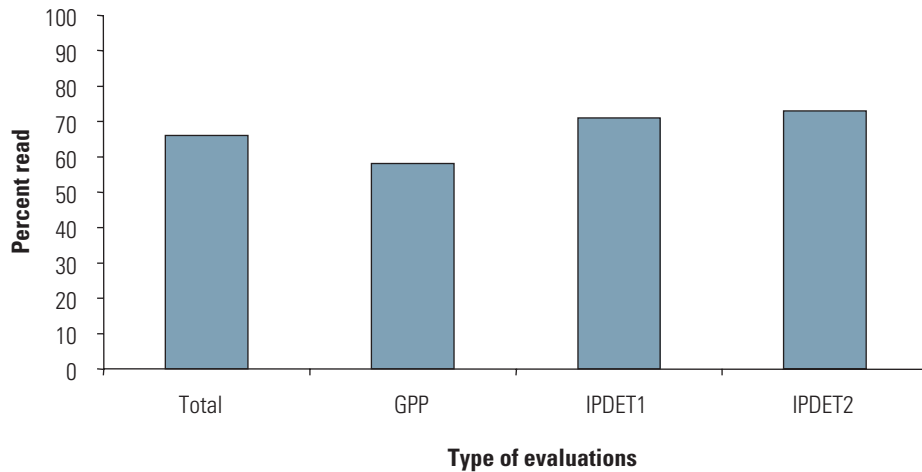


Figure D.7. Readership and Awareness of IEG Evaluations



about it by being consulted during the evaluation. IPDET respondents were asked if they were consulted about the evaluation paper during its preparation and only 2 percent said they had been consulted.

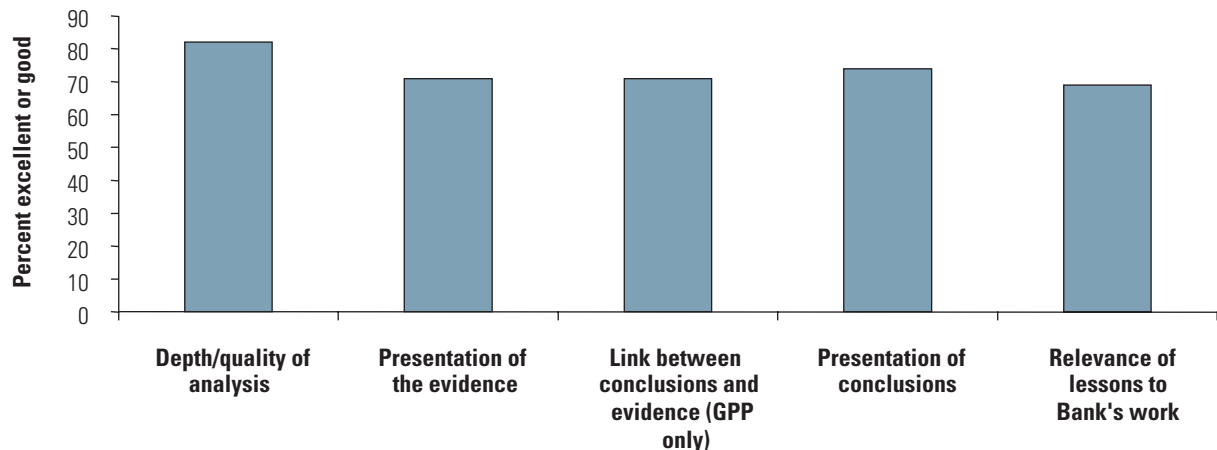
Quality of IEG Evaluations

A large majority of respondents (94 percent) indicated that their evaluation (GPP, 94 percent) or evaluation paper (IPDET1, 96 percent; IPDET2, 93 percent) was timely. GPP respondents were asked if the evaluation made full use of available information, and 80 percent of them

said it did. Eighty-one percent of IPDET1 and 89 percent of IPDET2 respondents said their evaluation paper described M&E techniques appropriately. Eighty-three percent of GPP respondents said they thought the evaluation used a methodology that was transparent and clear. Similarly, 92 percent of IPDET respondents said they thought the evaluation paper presented information in a transparent and clear manner.

Respondents were asked to assess several aspects of their evaluation (or, for executive directors, evaluations in general). Figure D.8 shows the proportion of respondents that rated each

Figure D.8: Quality of Studies



aspect of the evaluations either excellent or good.

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

- Depth (GPP) or quality (IPDET) of analysis (82 percent good or excellent)
- Presentation of evidence (71 percent good or excellent)
- Link between conclusions and evidence (GPP only, 71 percent good or excellent)
- Presentation of conclusions (74 percent good or excellent)
- Relevance of lessons to your work (68 percent good or excellent), including three-quarters (75 percent) of IPDET respondents.

Influence of IEG Evaluations

GPP respondents were asked the extent to which the evaluation influenced their view of the subject area. One in five (20 percent) said that it influenced their view a great deal, and another 48 percent said it influenced their view somewhat. A quarter (26 percent) said it had little influence, while 7 percent (just 3 of 46 respondents) said it had no influence.

Only 9 percent of respondents (including none of the IPDET government officials) said they had not and would not use the findings of the evaluation or evaluation paper in their work. Seventy-four percent said they had or would use the findings. The remaining 17 percent did not know. Of the populations responding, indecision or lack of knowledge about using the evaluation was highest for IPDET2 (*Influential Evaluations*) respondents (23 percent).

Results by Demographic Profile

For IPDET respondents, IEG checked for differences in perception of IEG work by position. The findings:²

- Among IPDET respondents, evaluation paper awareness was higher for researchers (76 percent) and consultants (77 percent) than for government officials (65 percent).
- IPDET government officials were more likely to share information or findings with colleagues

(57 percent) than were researchers and consultants (40 percent).

- While 80 percent of IPDET government officials thought their evaluation paper was good or excellent in presenting evidence, this was true for just 59 percent of IPDET researchers.
- Forty-two percent of IPDET government officials rated their evaluation paper to be excellent at presenting conclusions, versus just 19 percent of IPDET researchers.
- Indecision or lack of knowledge about using the evaluation was greatest among IPDET researchers (32 percent).
- Eighty-nine percent of IPDET government officials and 88 percent of IPDET consultants said they had or would use the findings of the evaluation paper in their work, compared with just 59 percent of IPDET researchers.

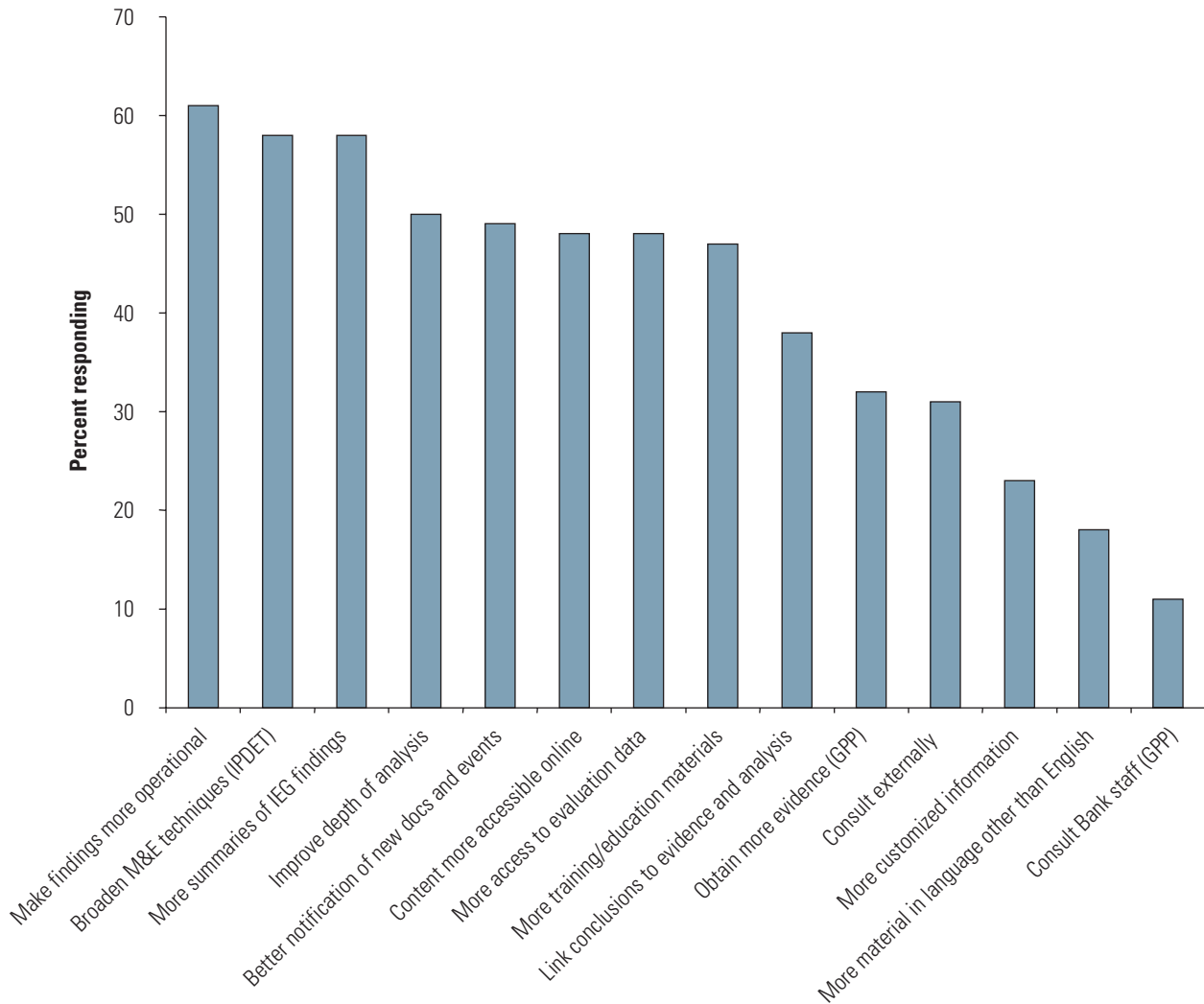
Recommendations from Respondents

When asked what recommendations they would make to improve their associated evaluation, the recommendation selected most often overall was to make findings more operational (61 percent, including 78 percent of IPDET government officials) (figure D.9). Additionally, 58 percent of IPDET respondents (including 75 percent of consultants but just 44 percent of government officials) recommended broadening the range of M&E techniques discussed. Half (50 percent) of those surveyed recommended improving depth of analysis. Broadening external consultation was selected by 62 percent of GPP respondents, but by only 19 percent of IPDET respondents. Respondents were also asked to select recommendations that they thought would improve IEG's outreach. The potential improvement selected most was for summaries of IEG findings to be provided. Providing more training or education materials was the most selected recommendation among IPDET respondents, but only 15 percent of GPP respondents selected it.

Implications for IEG

The survey results highlight the need for IEG to continue to improve the communication and dissemination of its evaluation products within

Figure D.9: Recommendations for IEG



the Bank and among its external clients. The new Outreach and Communications Strategy being launched by IEG is aimed at improving IEG's performance in this area. The survey findings indicate that clients would find summary information about IEG products useful. Also, increasing the use of online distribution of in-

formation appears to be important for dissemination, although effective targeting of audiences will be essential. In addition, the survey findings indicate that IEG needs to do a better job in relating its findings to operations. Finally, IEG needs to strengthen the linkages between evidence and conclusions.

APPENDIX E: MANAGEMENT RESPONSE

Introduction

The *2005 Annual Report on Operations Evaluation* (AROE) focuses on the country as the unit of account for monitoring, managing, and evaluating performance. The report takes stock of how well key Bank instruments and M&E procedures are linked to country goals and capacities, as well as to core elements of results-based management within the Bank and at the country level.

Management not only agrees with this focus, but has, in fact, been proactive in helping to lead the development community in that direction—internationally with the Comprehensive Development Framework and Poverty Reduction Support Paper (PRSP) concepts and leadership on the results agenda; internally in the gradual strengthening of the Country Assistance Strategy (CAS), including work on the country business model (Development Committee 2000), leading up to the Results-Based Country Assistance Strategy (RBCAS); and in its support to countries.

This year's AROE (which comes slightly more than six months after RBCASs were mainstreamed in the Bank) reviewed 24 CASs submitted to the Board in fiscal years 2004 and 2005 and reflects interviews with a sample of 26 country directors. The questions posed in this AROE are relevant to management's views of strategic priorities. The findings confirm management's sense of the advantages and obstacles to strengthening both Bank and country-level will- ingness and capacity to manage for results.

Recent Progress

Although the RBCAS approach is new, a careful review by management of seven pilot country cases has shown that it constitutes a significant step forward in taking results more seriously

(World Bank 2005a). The new approach helps distinguish between long-term country development goals and intended outcomes to which the CAS program directly contributes; strengthens the use of lessons learned from the last CAS; improves alignment of the Bank's program with the country's priorities; sharpens the design of the Bank program; and mobilizes country teams around a common vision of delivering results on the ground.

Another notable recent development is self-evaluation of the country assistance program in a CAS Completion Report (CASCR). The focus of the CASCR is on the achievement of CAS outcomes and on Bank performance in furthering CAS outcomes. It includes a discussion of projects and analytic and advisory work and is prepared at the end of the CAS period in time to deliver useful lessons for the next CAS. All CASCRs are independently evaluated by IEG, and each IEG assessment is submitted to the Bank's Board in advance of the Board's discussion of the next CAS document for the country.

AROE Analysis

The AROE reviews the use and efficacy of M&E for country program management. It also catalogues the measures taken to strengthen the results-orientation in M&E at the country, sector, and product levels within the Bank, updating the overview in the 2003 and 2004 AROEs. The analysis is a timely input for the next steps in the evolution of the country business model and internal implementation of the Results Agenda.

Measuring and Managing Performance at the Country Level

For the most part, the AROE uses the same concepts for measuring and managing performance

at the country level that management has used in setting out the RBCAS and the Results Agenda. These concepts are: strategic planning, performance measurement, and performance management. Management welcomes the finding that of the 24 CASs reviewed for the AROE, 19 attempted to establish causal linkages between country development goals, CAS outcomes, and Bank operations. It also welcomes the finding that the Bank's internal information systems, such as the Business Warehouse, allow Bank managers and staff to monitor budget use, product deliveries, and the quality of the portfolio at the country level.

IEG's Country Program Rating Methodology

The AROE cites an IEG review that found that one-third of country programs were rated unsatisfactory. Management notes that one-third unsatisfactory outcomes in country programs do not translate into an unsatisfactory Bank performance in one-third of country programs. IEG's methodology distinguishes between the performance of a country program and the performance of the Bank in support of that program. In IEG's methodological framework, the outcome of the Bank's assistance program is determined not only by the Bank's performance but by the joint impact of four agents (a) the client, (b) the Bank, (c) partners and other stakeholders, and (d) exogenous shocks (including events of nature and international economic shocks).

Also, despite extensive discussions, management and IEG have until recently not found agreement on the country program rating methodology, in part because before RBCASs were introduced, country programs were not designed to be evaluated. IEG has long rated the achievement of objectives, but objectives and strategies to achieve them change from one CAS to the next in response to changing country circumstances, and it was not clear to management *which* objectives IEG rated. Management and IEG are working to establish a clear understanding on the basis for ratings in the context of RBCASs and CASCRs, and management expects greater convergence of views in the future.

Baselines and Targeted Performance Measures in CASs

The AROE points out that lack of baselines and targeted performance measures limits the effectiveness of the Results Matrix as a monitoring, management, and evaluation tool. The report appropriately recognizes that the absence of baselines and performance measurement most often reflects lack of capacity and systems at the country level.

Management agrees that baselines and targets are central for managing for results, but notes that the RBCAS has only recently been mainstreamed, and that RBCASs are increasingly drawing attention to capacity gaps at the country level. It will take time before the Bank and our partners develop the systems necessary to ensure that baselines and targeted performance measures are in place—probably a whole RBCAS cycle for most countries, or longer. In the early stages of the RBCAS more attention has been focused on overcoming conceptual difficulties, such as identifying those outcomes to which the Bank could contribute, but not exclusively, and distinguishing between longer-term country objectives and CAS outcomes. A key issue has been striking a balance between choosing relevant outcomes to be monitored and availability of baseline data to monitor outcomes. Management agrees with AROE's observation that "availability of data, however, should not be a main reason for selecting measures. It may be necessary to include indicators for which data collection systems will need to be developed during the CAS period." Management does not want to create incentives for outcomes to be chosen only because baseline data are available, or for the goal to become meeting targets—as opposed to achieving results.

Capacity and Interest in Countries to Manage for Results

Helping country partners develop the interest and institutional capacity needed to manage for results is one of the three pillars of the results agenda, and central to the success of the entire agenda. Since RBCASs are based on the client country's own objectives as articulated in PRSPs

or similar documents, they reflect country ownership. The Bank has also actively supported countries in strengthening their national strategic planning (including for poverty reduction strategies) and in building results-based public sector management, statistical capacity, and M&E systems. Specific country-level activities by the Bank include the following:

- Strengthened advice from the Poverty Reduction and Economic Management Network on the design, monitoring, and evaluation of poverty reduction strategies.
- Demand-driven support from the Bank's Monitoring and Evaluation Improvement Program for countries to assess their institutional readiness and strengthen results-based approaches to public sector management.
- Development of the Statistical Capacity Building Program—approved by the Board in March 2004—which provides a sectorwide approach to building capacity based on a national statistical plan for providing reliable and timely data on countries' core development outcomes as articulated in their PRSPs or national development strategies. By the end of fiscal year 2005, five countries had accessed this program. In addition, the multidonor Global Trust Fund for Statistical Capacity Building, managed by the Bank, has supported the completion of National Statistical Development Strategies in 13 out of 56 countries with a PRSP, and a further 19 are in preparation.
- Development of a system for measuring statistical capacity at the country level. Starting in fiscal 2006, this measure is a key performance indicator in the Regional Strategic Performance Contracts (SPCs), signaling to staff the importance of working with clients to build country-level capacity.
- Mainstreaming of RBCASs, which also aim at strengthening client countries' capacity for managing for results and their monitoring and evaluation systems. The Development Economics Vice-Presidency in the Bank surveyed 50 recent CASs for IDA countries with the aim of identifying concerns about statistics and strengthening statistical capacity (IDA 2004).

Compared with conventional CASs, the five pilot RBCASs for IDA countries paid more attention to deficiencies in statistics and proposed more concrete plans to address those deficiencies. Both pilot CASs for IBRD borrowers proposed strengthening of monitoring and evaluation capabilities.

Country Ownership and the RBCAS

While the Bank can and should play a supportive role in developing capacity, the AROE tends to downplay the role of the country in managing for results. Management views performance measurement and management by countries themselves as central to the results agenda, because unless countries themselves manage for results, neither the Bank nor other donors can effectively manage for results. The most appropriate place to “(a) determine the information needed to monitor progress toward country goals and manage the Bank country program; (b) set targets and establish baselines; (c) identify the sources of information; and (d) establish responsibility for collection and analysis of the data” is not the RBCAS but countries' own growth and poverty reduction strategies. The RBCAS can, however, “determine areas where additional capacity building and assistance may be required and addressed through the CAS.” One of the commitments in the March 2005 Paris Declaration by partner countries was that they would “[e]ndeavor to establish results-oriented reporting and assessment frameworks that monitor progress against key dimensions of the national and sector development strategies; and that these frameworks should track a manageable number of indicators for which data are cost-effectively available” (box 2.3).

Monitoring at the Sector Level

The Sector Strategy Implementation Update (SSIU) provides an integrated assessment of the implementation of all sector (and thematic) strategies, trends in analytic and advisory services, lending commitments, and partnerships by sector/thematic areas. It draws on Implementation Completion Reports to pull together selected output-level results, such as kilometers

of urban roads built/rehabilitated and the number of rural facilities built/rehabilitated, organized by thematic area. Central elements of the results reporting and learning system that is under development will likely include progress on sectoral outcomes in countries where the Bank is active (or has focused its strategy), and lessons learned on (a) best practice in achieving results, (b) measurement issues, and (c) approaches to strengthening capacity for managing for results at the sector/country level (including analysis of the incentive environment that determines demand for information and statistical capacity building).

Trust Funds and Global Programs

Management has consistently noted that trust funds are a financing vehicle to support activities. What matters is that the activities be evaluated as activities, not as trust funds. There is a well-established agenda for M&E of activities across the Bank. Many trust funds are financing vehicles for partnership programs. Out of 115 partnership programs, 59 are trust-funded and fall under the new business process that requires a description of objectives of the proposed work program, as well as performance indicators. This is part of the new control of Global Program and Partnership (GPP) quality at point of entry. Management is moving toward a situation in which all GPPs will have to be independently evaluated every three to five years, regardless of funding source. A clear and strong differentiation of the Bank's roles as trustee and fiscal agent is important in the context of results orientation and accountability for results. When the Bank is strictly a fiscal agent, it cannot be accountable for program results. For example, the Bank is a fiscal agent for the \$3 billion Global Fund to Fight AIDS, Tuberculosis and Malaria; however, because the Bank is neither involved in the Fund's secretariat nor has a vote on its Board, it cannot be held accountable for the Fund's outputs, outcomes, or impacts.

AROE Recommendations

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

- Incorporate in the CAS a diagnosis of the country's performance measurement capacity and recommendations for action to strengthen that capacity. The diagnosis and actions should be customized to country capacity and focused on critical information needed for the government to implement its program.
- Expand use of the Results-Based CAS as a tool to manage country programs and strengthen the link between the Results-Based CAS and the country program by providing country teams with guidance designed to increase the results orientation of CPPRs.

Management Views

Management agrees that a diagnosis of the country's performance measurement capacity and recommendations for action to strengthen that capacity is appropriate, and already recommends that an assessment of M&E and results capacity be undertaken during CAS preparation. However, management does not agree on the need to mandate a formal document required as a CAS annex for all countries and CASs. There are cases where sufficient knowledge on country performance measurement capacity is already widely available or can easily be gained from work done by partner agencies. Regarding the second recommendation, management has already introduced the RBCAS and CASCR to strengthen country program management. In the RBCAS framework, the CASCR forms a crucial bridge between the downstream assessment of progress toward CAS outcomes and the upstream assessment of what is required in the next CAS to support the achievement of national development goals. Results-focused CPPRs can support country program management as input to the CASPR and CASCR. However, management would prefer to let the pilots now under way go forward and draw on that experience before deciding on new staff guidance on CPPRs.

Conclusions

Management welcomes the AROE and finds its analysis timely for its work on the next step for the results agenda. Management has introduced RBCASs and CASCRs, so the key elements for

“strengthening country program management” are in place. However, management would emphasize the complexity of the changes necessary to significantly improve management for results—in client countries, in the Bank, and across the donor community. As the IEG notes, industrial country governments have struggled with

this agenda. These difficulties are compounded in developing countries, where the requisite data are in short supply and the requisite institutions are far less advanced. Success will require time, money, and persistence; expecting results too quickly risks undermining support for the process.

Matrix: Management Response to the 2005 Annual Report on Operations Evaluation

IEG Recommendation	Management Response
<p>1. Incorporate in the CAS a diagnosis of the country’s performance measurement capacity and recommendations for action to strengthen that capacity. The diagnosis and actions should be customized to country capacity and focused on critical information needed for the government to implement its program.</p>	<p>Management agrees that this is an issue that should be discussed in the context of RBCAS preparation. Management does not agree on the need to mandate any particular approach or standardize reporting in the RBCAS. A formal requirement applied to all countries may lead to duplication of efforts in cases where sufficient knowledge can be gained from work undertaken/planned through other operational work or by partner agencies.</p>
<p>2. Expand use of the Results-Based CAS as a tool to manage country programs and strengthen the link between the Results-Based CAS and the country program by providing country teams with guidance designed to increase the results orientation of CPPRs.</p>	<p>Management agrees with the major thrust of this recommendation. Management is mainstreaming RBCASs and CASCs to improve the management of country programs. CAS results are selected during CAS design, tracked during implementation, and evaluated at the end of the CAS period. A CAS Progress Report allows for midpoint stocktaking and, as necessary, course corrections. The results-orientation of the CAS and the CPPR is strengthened if the projects and analytic work focus on results. For both projects and analytic work, in the last 18 months Management has introduced a series of revised tools and formats to facilitate management of country programs in relation to the CAS. However, Management would wait for the results of pilot CPPR exercises before deciding on new staff guidance on CPPRs.</p>

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

Background

The *2004 Annual Report on Operations Evaluation* (AROE) reviewed the results orientation of the Bank's monitoring and evaluation (M&E) systems and how these systems contributed to the Bank's results agenda. During the Board meeting in October 2004, Directors discussed many issues raised by CODE including organizational culture and incentive framework; use of multi-year budget and key performance indicators for more results-oriented behavior; and importance of country ownership, capacity, and interest in managing for results. Management reported on progress in introducing the results-based CASs in February 2005; these CASs have now been mainstreamed and management is updating its operational guidelines. A progress report on the results agenda will be prepared in the first half of FY06.

AROE Recommendations

The 2005 AROE focused on assessing the use and usefulness of M&E information to improve results and enhance Bank effectiveness at the country level. The Report's main message was that the Bank is making progress on the results agenda, but that it will take continued effort by the Bank and time to strengthen and link up with government systems for performance measurement and management, to generate demand for this information among country partners, and to harmonize further with other development agencies. The three recommendations for management were: (i) incorporate in the CAS a diagnosis of the country's performance measurement capacity and recommended actions to build required capacity; (ii) expand the use of results-based CASs (RBCASs) as a tool to manage country programs and strengthen the link

between RBCASs and the country program by providing country teams with guidance to increase the results-orientation of Country Portfolio Performance Reports (CPPRs); and (iii) subject to CODE/Board endorsement of new procedures, establish clear lines of accountability for the follow-up and oversight of adopted recommendations and include in Management Responses specific actions to complete the implementation of accepted recommendations and agreed-upon time lines for adoption. The two recommendations for IEG were: (i) review results of the piloting of its outreach and communication strategy before the full implementation of the strategy; and (ii) subject to CODE/Board endorsement of new procedures for the handling of recommendations, institute a formal system for tracking and managing IEG recommendations.

Management Response

Management broadly supported the IEG findings. It commented on initiatives in progress, including the country program rating methodology being discussed with IEG; the fact that RBCASs were relatively new and expected to identify baselines and realistic CAS outcomes as well as identify and address capacity issues; initiatives to strengthen country capacity and demand to manage for results; improvements to the Sector Strategy Implementation Update; and M&E of global programs and partnerships. It also stressed the important role of governments and countries in managing for results. Management generally concurred with the IEG recommendations but disagreed on the need to mandate the assessment of country capacity to generate, collect, and analyze performance information and its inclusion as a CAS annex. IEG noted that it had not proposed a mandatory annex, but that CAS

preparation involved an assessment of a country's performance management capacity. Management also proposed that any new staff guidance of CPPRs should be considered based on the results of pilot CPPRs. Management also suggested that IEG recommendations could be more precise and prioritized.

Overall Conclusions and Next Steps

The Committee welcomed the 2005 AROE. It noted the progress made in the Bank's results agenda and stressed the need to continue its efforts. There was broad support for IEG's recommendations and EDs encouraged management to implement them pragmatically. They endorsed management's views that assessment of a country's M&E and statistical capacities should not be mandatory and required as a separate CAS annex; they agreed that new guidelines for the CPPRs should be based on the results of the ongoing pilot. Members appreciated the initiatives to improve IEG effectiveness, particularly the dialogue between IEG and management. In this regard, a member stressed the importance of upfront IEG and management discussion and understanding of key definitions, scope and methodology of each evaluation. Some members encouraged efforts in enhancing IEG outreach and communications.

The main issues raised during the meeting were the following:

Challenges in Measuring and Managing Performance at the Country Level. Members appreciated the challenges, including the issues of attribution of outcome to Bank efforts and limited country capacity to collect and analyze data. They agreed with IEG on the need for selectivity in choosing performance measures and indicators. On harmonization of performance measurement systems, one member considered that the AROE appropriately emphasized its importance, while another felt that government capacity and not lack of interest was the explanation for the slow development of a common system. A member noted the efforts to align organizational culture and incentives to manage for results, while another urged continued efforts to address these issues, which the 2004 AROE highlighted. Another member also encouraged

strengthening of M&E of sector strategies' outcomes, and another member emphasized the importance of M&E of trust fund activities. There was a suggestion for AROE to return to the issue of results focus of the Bank's M&E in a few years. IEG stated that time was required to build a common M&E system relevant to the country's needs; it noted that nine out of ten country directors have an ongoing dialogue with governments on M&E systems issues, but only one-third had indicated that they had set up a common system for this purpose. IEG suggested that the prospects for establishing such a common system appeared stronger in PRSP countries than in middle-income countries

Role and Accountability of the Government and the Bank.

Members concurred with the management's view on the joint responsibility of the Bank and governments to manage for results. One member highlighted the issue of a disconnect between project and country program outcomes, as well as between country program and overall country performance, which needed to be addressed. Several members noted the need to: (i) elaborate on contributions of the country program to overall country performance; (ii) take into account the depth and size of Bank lending and analytical and advisory activities (AAA) relative to overall assistance to the country; and (iii) consider AAA and policy dialogue in strengthening the country's M&E capacities and ownership. A member stressed the Bank's important catalytic and coordinating role to promote results management and appropriate systems at the country level. Management responded that the Bank and donor community had in the last 50 years focused on project-based systems that created disincentives to strengthening country capacities. It stressed the importance of focusing on the quality of Bank inputs and on the country and its capacities and systems, which are central to achieving real impact. Management stated that it was attempting to address the disconnect between project and country program outcomes through the Results-Based CAS, which closely links project to country outcomes. The DGE suggested that this disconnect may be attributed to factors such as relevance, relative

size, and priority of projects. Management agreed with the DGE's observations on challenges in considering links between country program outcomes and overall country performance: (i) the Bank may have a satisfactory country program that may be relatively large according to Bank standards, but as a share of the borrower's GDP may be small, and other factors may affect country performance; and (ii) the Bank's share of AAA may be relatively small but it may have a large leveraging effect. One member suggested an econometric exercise to look at the correlation between a country's economic rate of growth and Bank activity, both lending and AAA. In response to some questions about the different performance of Bank's Statistical Capacity Building Program (STATCAP) and the Global Trust Fund for Statistical Capacity Building, management suggested that the number of entrants for the Global Trust Fund had been higher due to DEC efforts and to the fact that this support is provided on a grant basis; for the STATCAP, countries have specific financing obligations for which some countries are still reluctant to borrow.

Results Based CAS (RBCAS) as a Tool to Measure and Manage Results. Members appreciated the efforts to pilot the RBCAS, and looked forward to the OPCS paper on Progress on the Results Agenda and on a Results Reporting System. A member reiterated the AROE findings on the need to further improve baseline data, and concrete and measurable outcomes to evaluate the Bank's development effectiveness. Another member said that performance indicators in CASs should be aligned with the PRSP results framework. Questions were raised on how country-level results can be effectively measured with flexible CASs and in middle-income countries. IEG responded that flexible and middle-income CASs were reviewed around CAS outcomes, as agreed with management. Members supported the incorporation of an assessment of the country's capacity to collect and analyze data as well as actions to address capacity issues in the CAS, but cautioned against a standardized and mandatory approach. IEG agreed that a formal requirement to assess a country's M&E and statistical capacity as a mandatory annex

may be unnecessary. At the same time, it stressed that country's M&E and statistical capacity issues should be featured prominently in CAS preparation. There was support for increased results orientation of CPPRs and agreement that the new guidelines should be based on conclusions of the ongoing pilot. IEG also viewed that it would be premature to issue formal requirements for CPPRs, but strongly encouraged systematic learning across the CPPR pilots and adoption of emerging practices. A member supported the plan for Country Assistance Evaluations (CAEs) to assess how the Bank's program supported country's capacity to adopt a result framework. In response to a member's question, IEG clarified that it has reduced the number of CAEs in its work program to accommodate the increase in CASCR reviews.

More Attention to Results of Analytical and Advisory Activities (AAA). IEG supported members' comments about the importance of AAA as key inputs to CAS, and the need to consider their impact and results, as well as to disseminate them. Some members proposed incorporating AAA in the CPPR. Management stated that closer attention was being paid to results of AAA in CAS Completion Reports, and AAA assessments are done by QAG.

Improving IEG Effectiveness. Some members stressed the importance of IEG and management discussions upfront to clarify definitions, scope, and methodology before each evaluation is undertaken; the DGE assured that IEG will address this issue, perhaps through more systematic dialogue around methodology and the scope of activities to be covered by an evaluation. Some members also supported the recommendation to enhance outreach and communication of IEG products, and called for shorter and more "reader friendly" presentations and summaries. In this connection, one member urged synchronized disclosure of IEG reports, Management Responses, and "Green Sheets." IEG stated that its Board-approved disclosure policy requires that if an evaluation report is disclosed, the final Management Response and the Chair's summary of the CODE or Subcommittee discussion

are also disclosed. It also said that it welcomed simultaneous disclosure, without undue delay, of the main IEG report, and was working with management to avoid large time gaps between CODE discussions and the disclosure of various documents. One member said that IEG's self-recommendations should be verifiable with clear actions and timelines in the AROE. He also asked about the status of revising IEG guidelines to enhance the results-orientation of M&E. IEG responded that it was introducing separate sections in PPARs and ICRs that will discuss implementation and use of M&E. It added that its assessments were increasingly incorporating M&E issues. Others commented on the IEG client survey, noting that there was a call for making findings more operationally relevant. A member noted that some data in the survey did not seem to be comparable to those in the 2004 AROE and suggested IEG might wish to revise

them. IEG stated that 45 percent of respondents had read, participated in a discussion of, or attended an event on an evaluation for which they were surveyed, which signaled a strong increase from survey results reported in the 2004 AROE and were an indication of IEG's increased emphasis on dissemination.

Precision and Number of IEG Recommendations to Management. CODE broadly shared management's view that IEG recommendations should be more precise, focused on a few critical areas, and reduced in number. However, a member cautioned against excessive precision, while another member asked for more clarity on management's expectation of the level of precision. It was acknowledged that management did not have to accept and implement all IEG recommendations, and had discretion to decide on how to implement them.

Chander Mohan Vasudev, Chairman

ENDNOTES

Chapter 2

1. Twenty-six country directors were surveyed and responses were received from 16 (62%).

2. Of the 24 CASs reviewed, 6 were Results-Based CAS pilots; 13 adopted a results approach, although they were not formal pilots; and 5 followed the traditional format in place at the time but were included in the review because they contained sufficient information to merit an assessment. Four other CASs also were prepared during this period following the traditional format, but did not contain sufficient information for evaluation. The full review is in Appendix A.

3. Performance measures are rated Level 0 if only a very few were well defined, baselined, or targeted.

4. For example, “Increase living standard of working population measured as increase in real salaries” is a well-defined measure because it indicates specifically what evidence should be used to determine the desired effect. But it lacks both a baseline and defined target that can be used to determine whether the objective for this measure was met.

5. The Country Director’s Portal is expected to allow country directors easy access to country information such as project and AAA deliverables, portfolio performance, country budget allocation, and utilization and administrative matters, such as deadlines and staff travel, which are available in different systems.

6. The stronger focus on results monitoring and evaluation in the Development Policy Loans (DPLs) may create an impetus for stronger monitoring in middle-income countries. In one middle-income country, the country director expected to develop a system to monitor results through a series of single-tranche DPLs.

7. When asked to describe difficulties they see in establishing a common system or process with the government, one replied: “It takes time and effort, and requires sustained government attention as well as Bank dedication to the task.”

8. A program was prepared for Nigeria but was incorporated into the Economic Reform and Governance Project.

9. For example, the CAS mentioned that data for key indicators may be nonexistent and need to be developed and went on to mention how the Bank and the government would collaborate to develop the necessary databases.

10. These guidelines are under revision to reflect the mainstreaming of the Results-Based CAS approach.

11. The country directors were asked to rate which they considered important for managing the Bank’s country program from a list of five possible uses of information: (1) Learning—developing alternative strategies/approaches based on success or failure; (2) Business planning—allocation and adjusting human and financial resources; (3) Modifying operations—redesigning projects, revising strategies; (4) Communications—describing Bank country performance; and (5) Accountability—assigning responsibility for success or failure for Bank operations. The results of the survey are summarized in Appendix B.

12. The list consists of strengthening the IDA14 results measurement framework, the MDGs, and the Implementation Follow Up-7 (IFU-7) indicators. The Strategic Performance Contract is discussed in more detail in Chapter 3.

Chapter 3

1. The sectors are investment climate/finance, trade, education for all, hiv/aids, child and maternal health, water supply and sanitation, environmental sustainability.

2. The indicators that will be tracked are: the quality at entry of IDA-supported operations; project outcomes for the IDA portfolio; the share of Implementation Completion Reports with satisfactory data on project outcomes; and the share of Project Status Reports with satisfactory baseline data for outcome monitoring.

3. DEC has taken the lead in improving the quality of the data for measuring progress on key indicators, including by compiling statistical capacity scores for all countries that will be monitored by KPIs.

Chapter 4

1. The surveyed audience for IEG's GPP report consisted of those that had been consulted during the evaluation, and for IEG's M&E reports alumni of IPDET, and therefore considered interested in development evaluation issues.

Appendix A

1. 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 that provide 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, only that the process chosen produces the desired characteristics.

2. Performance indicators need to be: defined, baselined, targeted, specific, aligned, timely, measurable, attainable, and realistic. The evaluators did not have sufficient country knowledge to make an assessment on whether the indicators were specific, timely, measurable, realistic, and attainable. The remaining criteria (defined, baselined, targeted, and aligned) are more technical in nature and could be fairly assessed by evaluators with expertise in results-based management, regardless of country knowledge.

Appendix B

1. Thirty-two countries submitted CAS or Country Partnership Strategies to the Board during this period. Given multiple country assignments, four country directors submitted 2 CASs and one country director submitted 3 CASs. In order to facilitate the response, these country directors were asked to submit a response for only one country of their choice.

2. Responses as a percentage of the total number of questionnaires sent out.

Appendix D

1. Statistical significance was tested at 95 percent confidence.

2. Statistical significance was tested at 95 percent confidence.

REFERENCES

- African Development Bank, Asian Development Bank, Inter-American Development Bank, European Bank for Reconstruction and Development, the World Bank, and the Development Assistance Committee of the OECD. 2004. "Managing for Development Results." Second International Roundtable Joint Marrakech Memorandum, March.
- Binnendijk, A. 2001. "Results-Based Management in the Development Co-operation Agencies." Executive Summary. Paris, OECD/DAC.
- CDF Evaluation Management Group. 2003. *Toward Country-Led Development: A Multi-Partner Evaluation of the Comprehensive Development Framework (Synthesis Report)*. Washington, DC: World Bank
- Development Committee. 2002. "Better Measuring, Monitoring and Managing for Development Results." DC2002-0019. Washington, DC.
- . 2000. "Supporting Country Development: World Bank Role and Instruments in Low- and Middle-Income Countries." DC2000-19. Washington, DC.
- GAO (U.S. General Accounting Office). 2004. *Results Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results*. GAO- 4-38. Washington, DC: U.S. Government Printing Office.
- Hatry, H., E. Morley, S. Rossman, and J. Wholey. 2003. *How Federal Programs Use Outcome Information: Opportunities for Federal Managers*. NAPA/IBM Center for the Business of Government. Managing for Results Series. Washington, DC: Center for the Business of Government.
- IDA (International Development Association). 2004. "Measuring Results: Improving National Statistics in IDA Countries." Washington, DC.
- IEG (Independent Evaluation Group, World Bank; formerly OED). 2005a. *Country Assistance Evaluation (CAE) Retrospective: An OED Self-Evaluation*. IEG Study Series. Washington, DC: World Bank.
- . 2005b. *The Effectiveness of World Bank Support for Community-Based and -Driven Development*. IEG Study Series. Washington, DC: World Bank.
- . 2005c. *Committing to Results: Improving the Effectiveness of HIV/AIDS Assistance*. IEG Study Series. Washington, DC: World Bank.
- . 2005d. "Outreach and Communications Strategy," March 8, 2005. IEGKE. Washington, DC.
- . 2004a. *The Poverty Reduction Strategy Initiative: An Independent Evaluation of the World Bank's Support Through 2003*. IEG Study Series. Washington, DC: World Bank.
- . 2004b. *Evaluation Capacity Development: OED Self-Evaluation*. June. Washington, DC.
- . 2004c. *2004 Annual Report on Operations Evaluation*. Washington, DC.
- . 2004d. *Addressing the Challenges of Globalization: An Independent Evaluation of the World Bank's Approach to Global Programs*. IEG Study Series. Washington, DC: World Bank.
- . 2004e. *The CGIAR at 31: An Independent Evaluation of the Consultative Group on International Agricultural Research*. IEG Study Series. Washington, DC: World Bank.
- . 2003. *2003 Annual Report on Operations Evaluation*. Washington, DC.
- . 2002. *The World Bank's Approach to Global Programs*. IEG Study Series. Washington, DC: World Bank.
- OED (Operations Evaluation Department, World Bank). See IEG.

- World Bank, 2005a. "Results Focus in Country Assistance Strategies: A Stocktaking of Results-Based CASs." Operational Policy and Country Services. Washington, DC.
- . 2005b. "Improving the Bank's Analytic and Advisory Services." Operational Policy and Country Services. Washington, DC.
- . 2005c. "Annual Report on Portfolio Performance: Fiscal Year 2004." Quality Assurance Group. Washington, DC.
- . 2005d. "Quality of Supervision in FY03-04 (QSA6)—A QAG Assessment." Quality Assurance Group. Washington, DC.
- . 2005e. *Global Monitoring Report 2005*. Washington, DC: World Bank.
- . 2004a. "Implementation of the Agenda on Managing for Results: Progress Report and Annexes." Operational Policy and Country Services. Washington, DC.
- . 2004b. "Country Assistance Strategy (CAS)/CAS Progress Report Guidelines." Operational Policy and Country Services. Washington, DC.
- . 2004c. "Reform of World Bank-Administered Trust Funds: The Way Forward." Report No. 29626. Washington, DC.
- . 2004d. "Methods for Monitoring Achievements Made toward IDA Results Indicators: An External Panel Review." Washington, DC.
- . 2003a. "Better Measuring, Monitoring, and Managing for Development Results: Implementation Action Plan." Operational Policy and Country Services. 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



ISBN 0-8213-6521-5

