

CHALLENGES IN
MONITORING
AND EVALUATION.
AN OPPORTUNITY
TO INSTITUTIONALIZE
M&E SYSTEMS



FIFTH CONFERENCE OF THE LATIN AMERICA AND THE CARIBBEAN MONITORING AND EVALUATION (M&E) NETWORK

CHALLENGES IN MONITORING AND EVALUATION:
AN OPPORTUNITY TO INSTITUTIONALIZE M&E SYSTEMS

June 2010

Gladys Lopez Acevedo

Katia Rivera

Lycia Lima

Helena Hwang

Editors

Document of the World Bank and Inter-American Development Bank

The content of the papers on this report are the responsibility of the authors and do not necessarily represent the views of the World Bank or the Inter-American Development Bank (IDB). The World Bank and the IDB do not guarantee the accuracy of the information included in this publication and accept no responsibility whatsoever for any consequence of their use.



The World Bank



© 2010 by The International Bank for Reconstruction and Development / The World Bank, 1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.
Internet: www.worldbank.org
E-mail: yiconferencia@worldbank.org

All Rights Reserved
Printing and Manufactured in Washington, D.C.
First Printing: June 2010

The findings, interpretations, and conclusions expressed in this book are entirely those of the authors and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent.

The World Bank does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use. The boundaries, colors, denominations, and other information shown on any map in this volume do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

Rights and Permissions

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

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

The World Bank
Edition: 1st
Cover: Larry Mayorga

Lopez-Acevedo, Gladys (Trust Fund TF096261-SFLAC)

Contents

Abbreviations and Acronyms	vii
Acknowledgments	xiii
Foreword	xv
Welcome Address	xix
Conference Overview	xxi
Session 1: Challenges to Evidence-Based Decision Making: The Role of Monitoring and evaluation and Other Factors	1
Monitoring and Evaluation System in the Minas Gerais State Government: Aspects of Management	1
Transparency Index of Public Entities: Experience in Evaluating Risks of Corruption in Government Administration at National, Departmental, and Municipal Levels	7
Challenges for Evidence-Based Decision Making: The Experience of Chile and Chief Features of Its Design, Implementation, and Use of Information	13
Session 2: Institutional Arrangements for Monitoring and Evaluation Systems: International Experiences	25
Monitoring and Evaluation in the Government of Canada: Lessons Learned from 30 Years of Monitoring and Evaluation Development	25
The Monitoring and Evaluation System in Sri Lanka: Experiences, Challenges, and the Way Forward	33
Institutional Arrangements for International Monitoring and Evaluation Systems	47
Institutionalizing Monitoring and Evaluation by the Public Service Commission of South Africa: A Continuing Journey	50
Session 3: Institutional Arrangements for Monitoring and Evaluation Systems in Latin America	57
Challenges and Opportunities for the Institutionalization of Regional Institutional Agreements for Monitoring and Evaluation Systems in Latin America: The Experience of Costa Rica	57
U.S. Monitoring and Evaluation Trends: Using Performance Information to Improve Results	61
Institutional Arrangements for M&E Systems in Latin America: The Case of Mexico: Performance Evaluation System	66
The Creation of an Evaluation System of Higher Middle Education in Mexico	75
Latin American Monitoring and Evaluation Systems and Their Governance	82
Special Session: National System for Performance Management and Evaluation (Sinergia)	89
National Management and Results Evaluation System Achievements, Challenges, and Perspectives	89
Session 4: Monitoring and Evaluation Systems: Capacity Building for Monitoring and Evaluation and Academia-Government-Civil Society Partnerships	97
Conceptual Framework for Developing Evaluation Capacities Building on Good Practice in Evaluation and Capacity Development	97

The Experience of the Latin American and Caribbean Institute for Economic and Social Planning in Evaluation Training	109
Strengthening Monitoring and Evaluation Capacities through the Design and Implementation of National Results-Based Monitoring and Evaluation Systems: Lessons Learned in Africa, Asia, and Latin America	113
Session 5: Institutional Arrangements and related Policies to ensure the quality, access, and use of performance information	123
Challenges in Monitoring and Evaluation: An Opportunity to Consolidate the M&E Systems	123
Performance Information: How Web 2.0 Can Help	127
Institutional Arrangements and Policies of Monitoring and Evaluation Systems to Ensure Quality, Access, and Use of Information	133
Select Presentations from the Fourth Meeting of the Latin American and Caribbean Monitoring and Evaluation Network	139
Institutionalizing Impact Evaluation: Lessons from Latin America and Other Regions	139
Building a Monitoring and Evaluation System Based on Results for Social Development	144
Impact Assessment of Regional Development and Peace Programs in Colombia	151
Progress and Challenges of Monitoring and Evaluation Systems in Latin America and the Caribbean	156
The Brazilian Monitoring and Evaluation Network: A Report on the Creation and Development Process	164
References and Additional Bibliography	173
Figures, Tables, and Boxes	
Figure 1: M&E System in the Public Policy Cycle	2
Figure 2: Strategic Map, Minas Gerais State Government	3
Figure 3: An Overview of the Canadian M&E system	28
Figure 4: Organization and Implementation of M&E in a Government Department	29
Figure 5: MfDR in Sri Lanka	34
Figure 6: Whole-of-Government Approach	34
Figure 7: Snapshot of the Agency Results Framework	36
Figure 8: Agency Performance Measurement Framework	37
Figure 9: Data Capture at Source	42
Figure 10: MfDR and the Paris Declaration	43
Figure 11: Performance Evaluation System, First Stage	67
Figure 12: Results-Based Management Strategy	68
Figure 13: Performance Evaluation System, Second Stage	71
Figure 14: Budget Programming Structure	73
Figure 15: National Baccalaureate System	76
Figure 16: Milestones of Reform, 1991–2009	90
Figure 17: Results Road Map	91
Figure 18: National Development Plan Goals	92
Figure 19: Annual Evaluations, 2006–10	92
Figure 20: Goals Management System	94
Figure 21: Evaluation Principles	98
Figure 22: Levels of Capacity Development	99
Figure 23: Five Dimensions of Governance	124
Figure 24: Structure of the Government M&E System in India	128
Figure 25: Linking Outputs, Impacts, and Outcomes	129
Figure 26: Different Measurement Techniques	132

Figure 27: Three Pathways for the Successful Institutionalization of IE	141
Figure 28: Results-Based M&E System, 2004–07	146
Figure 29: Results-Based M&E System, 2008–12	149
Figure 30: RDPP Change Theory	153
Figure 31: Management Cycle Pillars.	158
Figure 32: MfDR Score by Pillar and Country Group.	159
Figure 33: Brazilian M&E Web Site	169
Figure 34: Growth of Membership in the Brazilian M&E Network.	169
Figure 35: Dashboard of the Brazilian M&E Network.	170
Table 1: Logical Framework Matrix	15
Table 2: Working Design for Evaluation of Programs and Institutions (ex post), 2000–09	16
Table 3: Uses and Actors in Ex Post Evaluations	19
Table 4: The Evolution of M&E in Canada.	26
Table 5: Drivers for M&E	29
Table 6: Key Features That Reinforce Credibility and Quality of the M&E System	30
Table 7: Implementing the M&E System	32
Table 8: Building M&E Capacity	32
Table 9: Changes Introduced through NCESP 3294.	91
Table 10: Channels of Communication.	95
Table 11: Evaluation Principles and Levels of Capacity	102
Table 12: Evaluation Capabilities.	103
Table 13: Tips to Reinforce an Enabling Environment for Evaluation	106
Table 14: Tips to Develop an Institutional Framework for Evaluation	107
Table 15: Tips for Stakeholders to Develop Skills and Knowledge of Evaluators	108
Table 16: Training Evaluation Levels	116
Table 17: Governance Indicator.	124
Table 18: Modification of Indicator.	125
Table 19: Measurement Techniques.	130
Table 20: Annual External Evaluations, 2004–07	146
Table 21: Country Ranking by Degree of MfDR Development.	158
Table 22: MfDR Pillars by Country Group	159
Table 23: M&E Indicators by Country Groups.	161
Box 1: Basic guidelines for collection, review, and appraisal of evaluation data.	10
Box 2: Structural independence	100
Box 3: Selecting the head of evaluation	107
Box 4: Definitions	109
Box 5: Comparison of Web 2.0 and Government 2.0	131
Box 6: Juana's story	154
Box 7: Alternative responses to the social dilemma	154
Box 8: Information dissemination	162
Box 9: PRODEV survey results.	163

Abbreviations and Acronyms

AAPSCOMS	Association of African Public Service Commissions (South Africa)
ADB	Asian Development Bank
AEP	Annual Evaluation Program
AFREA	African Evaluation Association
AIM	Africa Impact Evaluation Initiative
ANIF	<i>Asociación Nacional de Instituciones Financieras</i> (Association of National Financial Institutions)
AOPs	Annual Operational Plans
ARF	Agency Results Framework
BNB	<i>Banco do Nordeste do Brasil</i>
CAPAM	Commonwealth Association for Public Administration
CCCI	Internal Control Coordination Commission
CEE	Centre of Excellence for Evaluation
CENEVAL	<i>Centro Nacional de Evaluación para la Educación Superior</i> (National Center for the Evaluation of Higher Education)
CERCAPAZ	Cooperación entre Estado y Sociedad Civil para el Desarrollo de la Paz
CES	Canadian Evaluation Society
CGU	Office of the General Comptroller
CLAD	Latin American Center for Development (<i>Centro Latinoamericano de Administración para el Desarrollo</i>)
CMA	Comprehensive Management Analyses
CONEVAL	<i>Consejo Nacional de Evaluación de la Política de Desarrollo Social</i> (National Evaluation Council on Social Development Policy)
CONPES	<i>Consejo Nacional de Política Económica y Social</i> (National Economic and Social Policy Council)
CREAM	Clear, Relevant, Economic, Adequate, and Monitorable
CTMA	Technical Monitoring and Evaluation Chamber (<i>Câmara Técnica de Monitoramento e Avaliação</i>)
DEPP	<i>Departamento de Evaluación de Políticas Públicas</i> (Public Policy Evaluation Department)
DFABM	Department of Foreign Aid and Budget Monitoring
DfID	U.K. Department of International Development
DNP	<i>Departamento Nacional de Planeación</i> (National Planning Department)
DPR	Departmental Performance Report
ECB	Evaluation Capacity Building
ECD	Evaluation Capacity Development

ECLAC	Economic Commission for Latin America and the Caribbean
EDMS	Electronic Documentation Management Systems
ePMS	Electronic Project Monitoring System
ERO	Education Review Office(New Zealand)
EIS	Evaluation Information System
EMS	<i>Educación Media Superior</i> (Higher Middle Education)
ENLACE	<i>Evaluación Nacional del Logro Académico en Centros Escolares</i> (National Evaluation of Academic Achievement in Academic Centers)
FEB	Federal Expenditure Budget
FLACSO	<i>Facultad Latinoamericana de Ciencias Sociales</i> (Latin American Faculty of Social Sciences)
FMAP	Financial Management, Auditing, and Procurement
FNE	Northeast Financing Constitutional Fund
GAO	U.S. Government Accountability Office
GPRA	Government Performance and Results Act (United States)
GTZ	German Cooperation Agency
HME	Higher Middle Education
HMEES	Higher Middle Education Evaluation System
HMESP	HME Scholarship Programs
IAL	Incentives-Alignment [program]
ICRME	Intersectoral Committee for Results-Based Management and Evaluation
ICT	Information Communication Technology
IDB	Inter-American Development Bank
IDEA	Institute for Development in Economics and Administration
IE	Impact Evaluation
IEG	Independent Evaluation Group
3IE	International Initiative for Impact Evaluation
IFAD	International Fund for Agricultural Development
ILPES	Latin American and Caribbean Institute for Economic and Social Planning
INEE	<i>Instituto Nacional de Evaluación de la Educación</i> (National Institute of Evaluation of Education)
IPDET	International Programme for Development Education Training
IRHME	Integrated Reform of Higher Middle Education
ISMGA	Information System for Monitoring Government Action
JPF	João Pinheiro Foundation
JPR	Joint Progress Review
KPI	Key Performance Indicators
LFA	Logical Framework Analysis
LFM	Logical Framework Matrix
MAF	Management Accountability Framework

MAPSESI	<i>Matriz Anual de Programación, Seguimiento y Evaluación Sectorial e Institucional</i> (Annual Institutional and Sectoral Programming, Follow-up, and Evaluation Matrix)
MBC	Competitive Brazil Movement (<i>Movimento Brasil Competitivo</i>)
MDS	Ministry of Social Development and Hunger
M&E	Monitoring and Evaluation
M&E-MR	Monitoring and Evaluation-Managing (Management) for Results
MDGs	Millennium Development Goals
MEC	Monitoring and Evaluation Commission
MEU	Monitoring and Evaluation Unit
MfDR	Managing (Management) for Development Results
MR	Managing (Management) for Results
MIDEPLAN	<i>Ministerio de Planificación Nacional y Política Económica</i> (National Planning and Economic Policy Ministry)
MIP	Management Improvement Program
MPFA	Modernization of Public Financial Administration
MPI	Ministry of Plan Implementation
MPOG	Ministry of Planning, Budgeting, and Management
MRI	Matrix of Results-Based Indicators
MRRS	Management Resources and Results Structure
MTEF	Medium-Term Expenditure Framework
NACH	National Anti-Corruption Hotline (South Africa)
NAO	U.K. National Audit Office
NBS	National Baccalaureate System
NCESP	National Council of Economic and Social Policy
NDP	National Development Plan
NDS	National Development Strategy
NES	National Evaluation System
NGO	Nongovernmental Organization
NIEE	National Institute of Evaluation of Education
NIST	National Institute of Standards and Technology (Canada)
NMRS	National Management and Results System
NONIE	Network of Networks for Impact Evaluation
NOR	National Operations Room
NPD	National Planning Department
NPHI	National Public Health Institute
NPS	New Programmatic Structure
NREGS	National Rural Employment Guarantee Scheme (India)
NRP	National Reform Program
ODA	Official Development Assistance

OECD/DAC	Organisation for Economic Co-operation and Development/Development Assistance Committee
OED	Operations Evaluation Department
OMB	Office of Management and Budget
PART	Program Assessment Rating Tool
PES	Performance Evaluation System
PET	PRODEV Evaluation Tool
PIE	project implementation unit
PMDI	Integrated Minas Development Plan (<i>Plano Mineiro de Desenvolvimento Integrado</i>)
PPAG	Multiyear Government Action Plan (<i>Plano Plurianual de Ação Governamental</i>)
PPE	Public Policy Evaluation
PPX	Performance and Planning Exchange
PRODEV	Program to Implement the External Pillar of the Medium Term Action Plan for Development Effectiveness
PROGRESA	Education, Health, and Nutrition Program (<i>Programa de Educación, Salud y Alimentación</i>)
PROMAP	<i>Programa de Modernización de la Administración Pública</i> (Program for the Modernization of Federal Public Administration)
PSC	Public Service Commission
RBB	Results-Based Budget(ing)
RBB-PES	Results-Based Budgeting-Performance Evaluation System
RBBS	Results-Based Budget System
RBM	Results-Based Management
RBMES	Results-Based Monitoring and Evaluation System
RBMS	Results-Based Management Strategy
RDPPs	Regional Development and Peace Programs (Colombia)
RF	Results Framework
RMAF	Results-based Management and Accountability Framework
SAMEA	South African Monitoring and Evaluation Association
SEADE	Sistema Estadual de Análise de Dados
SEDESOL	Secretariat of Social Development (<i>Secretaría de Desarrollo Social</i>)
SEEMS	<i>Sistema de Evaluación de la Educación Media Superior</i>
SEMS	<i>Subsecretaría de Educación Media Superior</i> (Under Secretariat of Higher Middle Education)
SEP	Secretaría de Educación Pública (Secretariat of Public Education)
SEPLAG	Minas Gerais State Planning and Management Secretariat
SFC	Federal Internal Control Bureau
SFGG/SISMEG	System for the Follow-Up of Government Goals
SFP	Ministry of Public Administration (<i>Secretaría de la Función Pública</i>)
SGPGM	System of Government Programming and Goal Management
SHCP	Ministry of Finance and Public Credit (<i>Secretaría de Hacienda y Crédito Público</i>)

SINERGIA	National System for Performance Management and Evaluation (<i>Sistema Nacional de Evaluación de Gestión y Resultados</i>)
SLEVA	Sri Lanka Evaluation Association
SMART	Specific, Measurable, Achievable, Relevant, and Time-bound
SMS	School Management Service
SNCG	Secretariat of the National Comptroller General
SOPS	State of the Public Service (South Africa)
TBS	Treasury Board Secretariat
TCU	Federal Government Auditing Court
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNESCO	United Nations Educational, Scientific and Cultural Organization
UPEPE	<i>Unidad de Planeación y Evaluación de Política Educativa</i> (Planning and Evaluation of Educational Policy Unit)

Acknowledgments

THE PAPERS IN THESE PROCEEDINGS WERE presented at the Fifth Annual Conference of the Latin American and the Caribbean Monitoring and Evaluation Network, which took place in Bogotá, Colombia, on November 5 and 6, 2009. The conference topic was “Challenges in Monitoring and Evaluation: An Opportunity to Institutionalize M&E Systems.” The event was hosted by the Colombian National Planning Department (*Departamento Nacional de Planeación*, DNP), with the support of the government of Spain.

This publication includes three presentations from the Fourth Annual Conference, which took place in Belo Horizonte, Brazil, on November 25, 2008.

The Fifth Conference was organized jointly by the World Bank’s Latin America and the Caribbean Region, the Inter-American Development Bank, and the National Planning Department of Colombia. Guidance for the organization of the Conference was provided by Louise J. Cord,

Sector Manager for Poverty and Gender; and Roberto García López, PRODEV Coordinator for the Inter-American Development Bank (IDB). The Advisory Committee for the Fifth Conference included Gladys Lopez Acevedo (LC-SPP), Fernando Rojas (LCSPS), Manuel Fernando Castro (IEGCS), Diego Dorado (DNP), Heroldo Laguzzi (IDB), and Katia Rivera (IDB). The team acknowledges the support from Indu John-Abraham (LCSP), Irene Clavijo (LCSP), Lucy Bravo (LCSP), Pedro Arizti (LCSP), Henry Forero (CITPO), Jaime Saavedra (PRMP), and Elizabeth Tello (DNP).

We thank our colleagues from the World Bank and the Inter-American Development Bank who traveled and contributed to the conference and the speakers and participants whose willingness to learn and share their experiences played an important role in the success of this event. The presentations were submitted by participants at the conference and later prepared for publication by Gladys Lopez Acevedo, Helena Hwang, Lycia Lima, and Katia Rivera.

Foreword

EFFECTIVE POLICY MAKING REQUIRES INFORMATION on whether governments are doing things right and whether they achieve the results intended. Strong monitoring and evaluation (M&E) systems provide the means to compile and integrate this valuable information into the policy cycle, thus providing the basis for sound governance and accountable public policies. The Latin American and Caribbean (LAC) Monitoring and Evaluation Network was established by the World Bank and the Inter-American Development Bank (IDB) (Program to Implement the External Pillar of the Medium Term Action Plan for Development Effectiveness [PRODEV]) to promote and strengthen the use of M&E systems in Latin America and the Caribbean. Since 2005, the M&E Network has promoted knowledge and a forum for discussion by practitioners in key aspects of the institutionalization of M&E systems at national and subnational level.

The Fifth Annual Conference of the Network took place in Bogota, Colombia, from November 5–6, 2009. The Conference gathered more than 350 practitioners, government officials, academia, and international organizations (IDB, World Bank, the Food and Agriculture Organization, Comisión Económica de América Latina, RELAC) from more than 30 countries around the world, and an additional 300 people watched the event by Webcast. The evaluations from the conference indicated a positive feedback from participants, with 93 percent of the respondents ranking the conference as excellent or good. Support from the Government of Spain is greatly acknowledged for the organization of the Conference.

The 2009 Conference was hosted by the Colombian National Planning Department (DNP). The partnership with DNP was extremely fruitful for all sides. Country ownership of the event was demonstrated by high Colombian participation, about 150 professionals and high government officials from DNP, the Colombian Ministry of Finance, and the presidency. The Colombian government also took this opportunity to launch a national chapter of the M&E Network.

We thank World Bank and IADB colleagues who traveled and contributed to the conference. There was an unprecedented level of cooperation within the Bank to make this year's conference an international success. The World Bank-Independent Evaluation Group (IEG) and the IADB-OVE were active partners in the organization of the conference, providing inputs in terms of technical content and other resources. Also, a large number of participants from several sectors and regions (South Asia, Sub-Saharan Africa) attended the Conference. We are extremely positive about this broad based and multiregional engagement. We hope it will continue in the future and that the annual conference is increasingly as a key knowledge event for M&E.

A virtual social networking platform was also launched at the V Conference to enable the region's M&E Network participants to share information on good practices, events, evaluation methodologies, and other resources for M&E. The World Bank and the IADB have played a catalyzing role in setting the platform, but the goal is for M&E Network members also to contribute to the dynamics of the Web site. Already more than 600 members have signed up in the social network. The social network and the confer-

ence Web site featured 13 interviews of the leading speakers from the Conference—Brazil, Canada, Chile, Colombia, India, Italy, Mexico, New Zealand, Peru, South Africa, Spain, Sri Lanka, and the United States. The interview provide a comparative overview of the main features of the M&E frameworks developed in these countries.

Structure of the publication and of the conference

The objective of the Fifth M&E Conference was to discuss challenges in institutionalizing M&E systems and using M&E information to support planning and budgeting decisions as well as to enhance government transparency and accountability.

The structure of the present publication follows the structure of the Fifth M&E Conference agenda, which included seven sessions. The first session was devoted to the challenges facing evidence-based decision making: the role of M&E. The analysis looked at experiences from Minas Gerais, Brazil, Peru, the Mexican National Evaluation Council on Social Development Policy (CONEVAL), Mexico, and Chile. The goal was to show experiences in which M&E has influenced resource allocation and the modification, strengthening, or elimination of policies or programs.

The second session focused on institutional arrangements for M&E systems at the international level, and enabled participants to know, examine, and identify the advantages and difficulties of different institutional arrangements used in government management in Canada, Sri Lanka, Spain, and South Africa.

The third session addressed the institutional M&E arrangements in Latin America and discussed experiences in the region, considering the characteristics, advantages, and disadvantages of M&E systems depending on whether they fall under or outside the executive, and how central M&E systems relate to and complement their sectoral counterparts. Presentations included Costa Rica, the United States, Mexico, and a comprehensive look at the region.

The special session discussed the achievements and challenges of the Colombian National System of Evaluation of Results-based Management (SINERGIA) in its 15 years of operation. The analysis considered its overall effects and the way it has influenced the formulation and implementation of policies and/or programs in Colombia.

The fifth session analyzed the development of M&E capacities and the alliances between the government, academia, and civil society. Various presentations created the right atmosphere for a fruitful exchange of ideas.

The sixth session dealt with the institutional arrangements and policies of M&E systems to ensure quality, access, and use of information. The core portion examined ways to obtain reliable, clear, and timely information.

The seventh session of the conference focused on the exchange of information about the experiences and challenges that surround the creation of a national chapter of the M&E Network. Brazil presented their progress in establishing a national network. The lessons and recommendations were useful for Colombia which launched their national M&E chapter during the conference.

Emerging messages from the conference

The varied presentations highlighted a tremendous growth in recent years in both the supply of and demand for quality M&E initiatives. This growth is spurred on by increased demand for public sector effectiveness and transparency, by the increased need for information frameworks and systems to increase program quality, and on the supply side by the growing availability of M&E technologies and instruments. Several countries have created high-level agencies devoted to M&E (South Africa, Mexico, Colombia), and have strengthened their legal and regulatory frameworks to require regular M&E and to ensure the public dissemination of information (Mexico, Peru, Spain, India, New Zealand, Chile, and Australia) and adopted innovative M&E tools to strengthen their budgeting and planning exercises. The Brazilian state of Minas Gerais is using benchmarking tools to better plan and monitor their development plan. Mexico introduced a Menu of Evaluations two years ago, which consists of seven different types of evaluations (strategic, impact, consistency and results, indicators, policy, and process), which has allowed a greater use of M&E information by Congress, the Ministry of Finance, and the Presidency.

The growth in supply and demand presents tremendous opportunities, but also underscores some critical challenges and priorities for the future.

First, and perhaps most important, is the need to more consistently link M&E results to policy planning and budget processes and thereby to ensure that the increased in-

vestments in this area are sustained and used. This will require institutional incentives as well as enhanced use of methods, such as business intelligence, benchmarking, target setting, and quality data to ensure that the M&E information is integrated into the decision-making processes.

Second, it will be important to continue developing capacity for M&E. Training is needed both for the users of M&E information (Congress, government officials, legislation), as well as for its producers (National Statistics Office, academia, and government officials). There are some recent efforts to increase capacity through Regional Knowledge Centers (an initiative led by IEG) and to develop and strengthen national M&E chapters (an initiative led by LAC).

Third, to continue encouraging the growth of M&E systems and their innovative uses for policy making, several speakers noted the importance of creating more carrots and fewer sticks for implementing agencies to do M&E. For example, the United States recently launched a new initiative under President Obama to create a fund to encourage programs to undertake evaluations and to implement the recommendations from improvement plans. Efforts such as these should be expanded.

Fourth, while laws on access to public information have been introduced in some countries in the region, most of the best practice remains outside the region and there is a clear need to strengthen regulatory frameworks for public information access.

Finally, several speakers discussed the challenges of using M&E systems in weak institutional or low-income countries or regional entities (states, provinces) where core building blocks for public sector effectiveness and transparent governance are lacking (effective budget systems, planning mechanisms, core data, and so on). Efforts to understand how to promote data development and design useful M&E systems for fragile environments should be supported.

Next steps

The network has been actively supporting three emerging national M&E networks. The Brazilian Chapter of the Latin American M&E Network already has a large number of members (<http://redebrasileirademea.ning.com/>) and is organizing a second annual meeting to be hosted by the Banco do Nordeste in June 2010. Discussions are under way to create a sectoral subnetwork in Brazil around health M&E. Colombia launched their national network and first meeting. Mexico is in the process of setting up one. The initial focus of these networks is on sharing knowledge and promoting learning and capacity building on M&E among subnational entities. The network is supporting these national chapters and providing training in various M&E tool.

We hope the enclosed publication serves to continue fostering knowledge in M&E in the LAC Region and even beyond. Five years have passed since our first meeting held in May 2005 and judged by the growing interest of participants and the quality of the presentations, we can all be pleased with the results obtained thus far. The Sixth M&E Conference will take place in Mexico City on August 25–27. This event will be hosted by CONEVAL, a regional leader in the area of M&E and poverty monitoring.

As suggested at last year's conference in Bogota, Colombia, the theme for this year's conference will be using M&E results in the public policy cycle. The conference will highlight specific sectoral policy uses for M&E results, as well as the role of civil society and parliaments in producing and using M&E results. Two parallel pre-conference workshops also will offer participants the opportunity to learn more about CONEVAL's structure and activities, as well as a series of seminars of good practice M&E techniques and tools. For the first time this year, the agenda reflects a regionwide call for proposals among the network members and beyond to identify many of the conference presenters. Hereby, we extend an invitation to continue working and partnering with countries toward the institutionalization of M&E.

Louise Cord

Sector Manager

Poverty Reduction and Gender Unit

Latin America

World Bank

Roberto Garcia Lopez

PRODEV Coordinator

Inter-American Development Bank

Gladys Lopez Acevedo

Senior Economist

Poverty Reduction and Equity

World Bank

Welcome Address

IT IS AN HONOR AND A PRIVILEGE FOR US TO HAVE YOU here. I hope you enjoy both the region and the city, and take new perspectives and news of Colombia back to your countries. I would also like to thank the Inter-American Development Bank and the World Bank for their efforts to bring us together once a year and keep us up to date on what is happening in other parts of the world.

This type of event is a fantastic opportunity to learn about the good and bad aspects of others countries' experiences. In Colombia, we have learned a great deal from our neighbors and we are obviously delighted to be able to share our experiences and challenges, and the problems we have faced in this and other areas. We have achieved enormous progress in the area of monitoring and evaluation, which is crucial to public policies, not only for the departments that implement them but also to ensure that this information is available to society. The National Planning Department (NPD) is continuously uploading information to its Web site so that citizens can track and evaluate the results of government programs, an essential aspect of modern democracy.

We are also working to develop forms of verification to guarantee that programs are evaluated and that this information is useful to legislators and, above all, policy designers. Colombia faces the challenge of having a president whose management model is based on results,

data, and figures, for which the latest indicators are always available.

Since we are in the last year of this government (which has been in power for seven years), it is essential for the NPD to evaluate a broad range of public programs. This process will contribute to maintaining and adjusting the programs that are underway, regardless of changes in government. To ensure continuity and adjust and correct problems or flaws, it is important to demonstrate both technically and objectively how each program works. One of this administration's main priorities is to leave the incoming administration the products, data, and information to improve planning and further incorporate the newly acquired knowledge.

I am not an expert on monitoring and evaluation, so I will leave the technical discussion up to you but I would like to take this opportunity to welcome you to Bogotá. I am sure that the meetings will be productive and enjoyable for all of you. Finally, I would like to thank Diego Dorado and his work team and all of you for gathering this enormous range of representatives from all over the world here today.

Thank you

Esteban Piedrahita Uribe
Director General of National
Planning Department, Colombia

Conference Overview

THIS IS THE FIFTH YEAR THAT AN INCREASINGLY large network of academics, public officials, and experts have met to discuss, share experiences, and reflect on the challenges and opportunities in this vast world of monitoring and evaluation (M&E). The first two meetings were held in Washington, D.C., with more than 70 participants; the next meeting was held in Lima; and then Belo Horizonte; and now, Bogotá. The first meetings were dominated by presentations and agendas defined by international entities. Since we left Lima, we have taken steps to make this an activity with ownership and participation by the communities interested in these topics.

This year, the network is taking yet another step. After successfully joining efforts with the João Pinheiro Foundation (JFP) and the Minas Gerais state government in Brazil, we now have included the Colombian National Department of Planning and Synergy as a partner, in an experience that has been very productive. It is no accident that we are in Bogotá, nor was it when we were in Belo Horizonte. It is not just that the organizers were clever enough to bring us here to enjoy a fantastic city like Bogotá. We are here because things are happening in Colombia. In Colombia, interest in M&E can be found at the different levels of public administration, in the various departments and municipalities, and in both the public and private sectors. And, despite the complex institutional challenges, the public sector has a commitment at various key points to strengthen planning, to establish goals, and to carry out M&E to provide essential feedback to the design and formulation of public action. This conviction and commitment cannot be found just anywhere.

Why are we here?

Why are we here? In one way or another, we are all involved in public endeavors. We are here to improve people's quality of life. That is the ultimate goal. To achieve this, we are attempting to urge each one of you, in Colombia and in each of your countries, to use scientific procedures to improve your decision-making processes, and to insert M&E tools into project cycles and policy cycles. How can we insert our efforts in policy cycles?

To do this, we first have to know what is important for people's well-being. Is it to reduce infant mortality? Improve citizen security? Children's malnutrition levels? Increase young people's access to Internet? It is essential that we define the results that are the most important for taking action. And our first challenge is to have reliable data that will allow us to measure whether or not we make progress toward these results.

Once we have identified our critical areas of intervention, we need to know how much we should advance. For this reason many of you are in the business of establishing quantitative goals, and we are all involved in this. It is not the same to say that our goal is to reduce poverty, as to say that we are going to reduce poverty by 10 points in five years. This definition of quantitative goals, at all levels, is a key instrument that we use, but we do not use it enough. The focus on vigorously establishing goals entered the development world through the Millennium Development Goals, but now it should be part of all planning and monitoring systems. However, and this is an interesting point, the establishment of quantitative methods is not as common as it should be—

in part, because it requires commitment and clearer accountability.

And this matter of goals is not an easy task. If we set a goal of reducing poverty by two points per year, is that a lot or a little? If we attempt to reduce infant mortality from 30 to 25 per thousand in five years, will this mean that our country's performance will be excellent or only mediocre? What we call benchmarking in English—which makes it possible to determine whether a certain performance is adequate, good or bad, by making a comparison with what has been done in other countries or regions—is something that we do very infrequently. Setting goals, and setting them in an informed way, is complex, but it is the first critical step in M&E systems.

Now, if we know what the goals are, we have a better idea of what we should be monitoring. But to move forward with these goals, we need to implement policies that work and that will be useful in reaching these goals. The problem is that we always have lots of programs and it is not always clear how they will contribute to achieving our established goals. The link between the school breakfast program and nutrition, or the link between the occupational training program and the employability of young people, are two examples of processes, of functions that are not totally clear. We are using logical frameworks and some other tools, and we have made some advances along these lines, but we still have a long way to go.

In order to identify the policies that will allow us to make these goals a reality, or to know whether the current set of policies are working, we establish evaluation systems, evaluations of different types that allow us to have a better idea of what works and what does not.

Does monitoring and evaluation matter?

And this takes me to the last point, which is feedback. Does the work of M&E that you are carrying out matter? Is it useful for anything? It is useful if the decision-making process is going smoothly. Succeeding in making program evaluations and the information generated have an impact on decisions regarding ways to expand, perfect, or close programs is complicated and is the challenge faced by many of us who are here. A challenge for many of us is to establish forums for making evaluation results known, and to establish institutional mechanisms through which finance ministries and congresses can request and use this information.

How to make progress in inserting M&E in the policy cycle is a complex challenge, but that is why we are here.

Monitoring and evaluation is increasingly important

Policy discussions in Colombia and Latin America in general have been gradually changing in recent years. There is a clear consensus on the need for macro stability, low inflation, and openness to international trade. But there is also consensus in many of our countries that huge gaps and major inequalities persist between different groups of the population in terms of access to services, access to opportunities, and differences in access to good nutrition for children, to a good education for young people, and to good jobs for adults.

The development of increasingly more efficient public policies is indispensable to closing these gaps in opportunities. That is why government efficiency is at the center of the discussion. And this points to the need to ensure that the methodological tools and institutional mechanisms that will be discussed here will contribute to the process of designing and formulating policies. Only to the degree that we are able to move forward in this task of ensuring that all the accumulated information generated by M&E systems has an impact on policies will we know that we are doing a good job.

But it is difficult to move away from speaking about the global crisis we are facing (which possibly we have recently been moving out of, although it has undoubtedly halted or at least decelerated what we were gaining in terms of growth and poverty reduction). And this raises an additional challenge beyond the structural challenges in public policy and fiscal policy. These challenges vary among Latin American countries and throughout the world.

In some countries—unfortunately only a few—it is fiscally possible to increase spending. And M&E tools can assist us in making this spending more effective, to ensure that funds are being used correctly and on the right things.

In other contexts, we are already seeing the need to reduce spending. But how can we reduce spending, given the needs of our populations, given the increasingly greater need to provide opportunities for all—a task in which the state has a critical role? Spending cuts require the use of a scalpel, not a sledgehammer or a guillotine. And this can be done only if we know what works and what does not. “A

scalpel and not a sledgehammer,” President Obama said in his campaign. The M&E tools that we will be talking about here are precisely that scalpel.

Without information, we have nothing (capacity building)

At the World Bank, we consider contributing to capacity building at all levels to be one of our main missions. Better M&E systems will make it possible for policy design and implementation to be based on scientific evidence of what works and what does not.

Although all of us who are here today agree on many basic principles, determining the appropriate M&E system for each specific country depends on local conditions. It would be fantastic to have a perfect model of the organization of an M&E system that could be transplanted from one country to another. But that is not possible, and it is not feasible or appropriate. We know a single route is not the answer.

That is why the central role of the Inter-American Development Bank (IDB) and the World Bank is to help create and share knowledge to support the process of educating the critical masses of people who can help to build the best system in each country. In countries like Brazil, Chile, Colombia, and Mexico, advances have been made in this process, but we must keep going, and the road is even longer in other countries. And this is the reason we have contributed to forming national M&E networks, such as the Brazilian network that was created in a meeting like this one a year ago, and about which we are going to hear tomorrow, and like the one that is beginning to take shape

under the leadership of the *Departamento Nacional de Planeación* (National Planning Department, DNP) in Colombia precisely as we are meeting here. And, that is why we are going to be making a call to the IDB and the World Bank to support the formation of national networks in other countries.

And we are going to support the development of platforms that facilitate exchanges of knowledge and information between members of this Latin American network and members of national networks, with a much more aggressive use of the Internet and of new means to promote exchanges among all of you. As an example, every single World Bank employee can be watching this event right now, in any World Bank office anywhere in the world.

I want to close by expressing my appreciation to the IDB, to Roberto García López and Mario Marcel personally, for their constant spirit of collaboration and for enthusiastically maintaining this collaboration with the World Bank for five years now. And, second, I want to thank the DNP and Diego Dorado personally, as well as Director Esteban Piedrahita, for their support and trust, and for their hospitality in this marvelous city of Bogotá.

I also wish to thank the Colombian government, in general, and all of you who are fervently dedicated to this task of improving policies in your countries. I want to thank you for allowing us the opportunity to assist in this task.

Jaime Saavedra Chanduvi

Sector Manager, Poverty Reduction and Equity
World Bank

Session 1: Challenges to Evidence-Based Decision Making: The Role of Monitoring and evaluation and Other Factors

Monitoring and Evaluation System in the Minas Gerais State Government: Aspects of Management

Introduction

This paper describes the monitoring and evaluation (M&E) system implemented in the state of Minas Gerais in 2007, based on the Management Shock Project of 2003–06. The M&E system in Minas Gerais consists of (1) a planning process (as an important element when determining priorities), (2) institutional alignment instruments (such as results agreements, strategic management, and results committees), (3) benchmarks and strategic project goals and their final results (outputs and outcome indicators), and (4) the feedback process (an effort to allow for ongoing learning and reinvention of the modus operandi in public policy).

Foundations of the Minas Gerais model: Management shock 2003–06

The idea that Minas Gerais needed a management shock emerged from diagnostic work by the Minas Gerais Development Bank (*Banco do Desenvolvimento do Minas Gerais*, BDMG), which concluded that “any lasting reform of the state apparatus implies a ‘management shock’ and a revamping of the social insurance system” (BDMG 2002, 219).

The expression “management shock” was first introduced during the governorship of Aécio Neves (2003–06). From the beginning of his term in office, Governor Neves set out to implement the Management Shock Project, in a context “very similar to that of most other states in the federation, namely a very serious fiscal situation, including

a substantial budget deficit that had existed since 1996,” in which “the state apparatus was being systematically ruined” (Anastasia 2006, 13). In describing the background to management shock, Anastasia (2006, 15) emphasizes how the Management Shock Project became “the main guiding principle underlying the Minas Gerais government program.” In other words, the project enjoyed support and the political will to implement it. This, in large measure, proved essential to the successful implementation of the government’s measures.

In 2003, the Minas Gerais state government introduced results-based management. The Neves administration introduced (1) the linking of planning and budgeting, (2) a number of strategic projects (intensively monitored by a team of project management experts from the GERAES project office), and (3) results agreements and performance evaluations.

The Minas Gerais planning process included “double planning—a strategy based on a dual approach, with short-term strategies for the first year, while simultaneously working on a medium and long-term strategic plan” (Guimarães and Almeida 2006, 44), and the linkage of planning and budgeting to ensure that the strategy is viable in the short run (Oliveira, Almeida, and Guimarães 2006, 67). This two-pronged approach to planning can be described as follows:

The state strategy rested on two pillars: fiscal balance (current strategy) and a development agenda (future strategy) that justified before society the initial adjustments of the reduction and rationalization of spending, administrative reforms and increase in tax revenues. (Guimarães and Almeida 2006, 22)

Another important initial aspect of the management shock was the creation in 2004 of the GERAES Project Office, attached to the Minas Gerais State Planning and Management Secretariat (SEPLAG). The GERAES project office began its monitoring activities by preparing 31 strategic projects (4 percent of budget expenditures for 2004). Currently, the portfolio has 57 projects (11.5 percent of budget expenditures for 2008). Managers of strategic projects meet each month to present a status report. The report describes the physical and financial situation of these projects.

In her summary of what was accomplished during the first generation of shock management, Renata Vilhena (2006) highlighted two dimensions of the policy: on the one hand, a fiscal equilibrium that ensured the budget's effectiveness as a management tool; on the other, modernization and innovation in state public management by means of instruments that assisted in alignment and personal development of staff and thus of institutions. Looking forward, Vilhena also stresses the embryonic idea behind the second generation of shock management:

The great challenge we now face is to consolidate and institutionalize the process of transformation, in order to ensure it becomes irreversible. This must be done by going beyond the idea of fiscal responsibility and including social and managerial responsibility. In the ultimate analysis, this means a commitment to development results and accountability vis-à-vis society. (Vilhena 2006, 354)

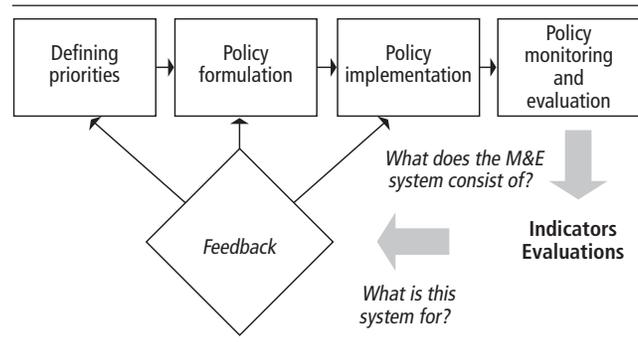
The foundation was thus laid to build a management model geared toward the achievement of outcomes that benefit society, known as the State for Results Program (*Programa Estado para Resultados*). Such a concept needed a strong M&E system, capable of generating reliable information.

Monitoring and evaluation system

Figure 1 illustrates the system's logic and the two questions that an M&E system should ask. An M&E system consists of a set of government performance indicators that are verified by a series of evaluations of the government's public policies. Based on these inputs, results-based management is made possible—meaning that the information collected

FIGURE 1:

M&E System in the Public Policy Cycle



Source: Adapted by the author from Barkenbus (1998).

is used to focus on public management. The creation of an M&E system must not be understood as an end in itself. Rather, it must be linked to the public policy planning and management process. To originate public value, therefore, the M&E system must be capable of bringing forth information that buttresses the decision-making process leading to the attainment of government objectives. It is essential, of course, that this information be used in practice by decision makers.

It is pertinent, then, to ask “What does the M&E system consist of?” Put otherwise, “what are the system’s main components?” The structure of the system leads to the second question, namely, “What is this system for?” The answer is generic: to obtain feedback from the public policy cycle.

Planning and setting priorities

The planning process in Minas Gerais can be differentiated from other Brazilian states. According to Article XX of the State Constitution, the Integrated Minas Development Plan (*Plano Mineiro de Desenvolvimento Integrado*, PM-DI) is a mandatory document crucial to strategic planning. The capacity and importance that a particular administration ascribes to government’s planning function stand between the formal obligation and the concept of real strategic planning. Starting in 2003, this long-term development plan took on a real role in demarcating government action and, in 2007, with the development plan review having taken place, the document’s function as a goal-setting tool that reflects the results the government strives to attain was strengthened.

The 2007–23 development plan includes the results areas, or “target areas,” on which the government’s best efforts and resources are focused as they work to achieve the desired transformations and results. This introduces the important idea of final results in public administration. A total of 13 such results areas were created (see Figure 2).

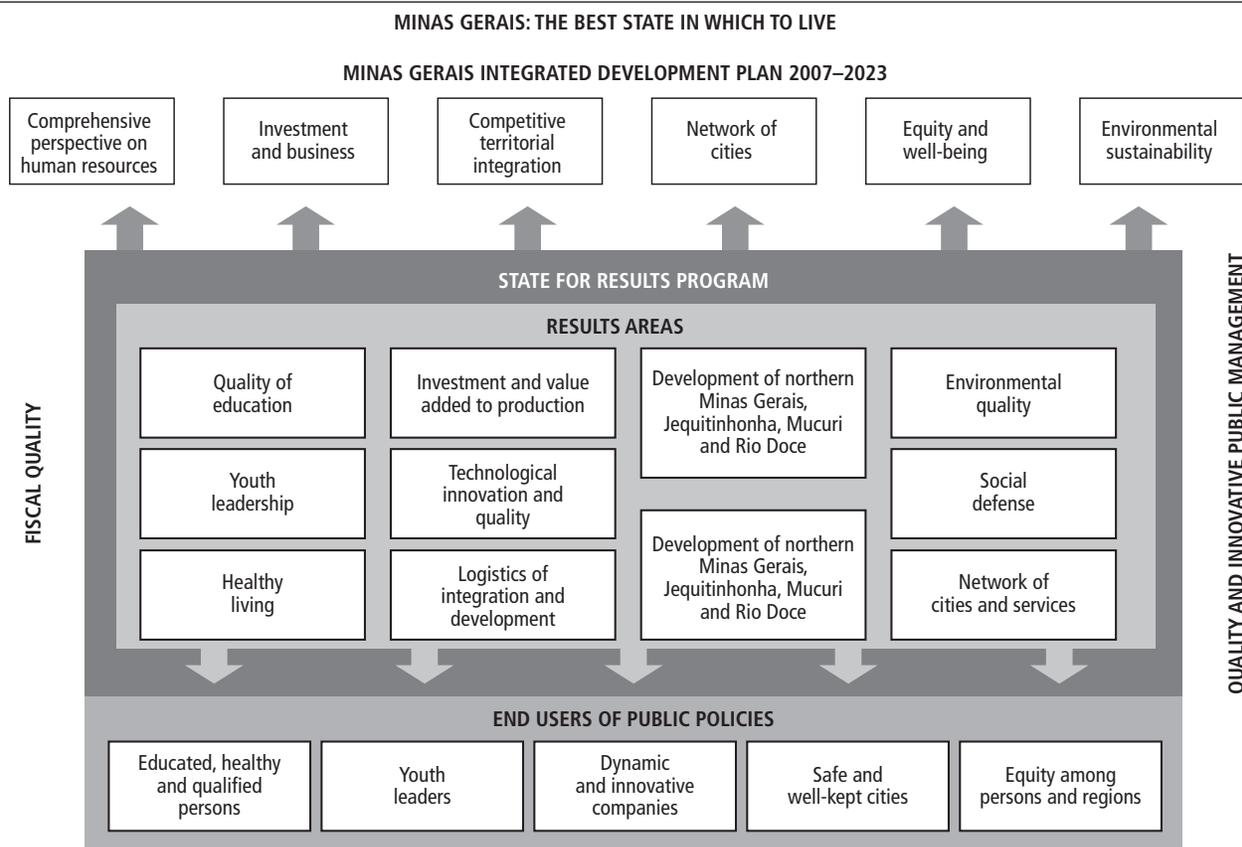
Aligned to the development plan, the Multiyear Government Action Plan (*Plano Plurianual de Ação Governamental*, PPAG) is bolstered by the linkage between planning and budgeting. It is presented as a document that determines the portfolio of structuring projects as strategic government initiatives. Taken together, these two documents signal the strategic objectives to be pursued (the development plan) and how it is thought these might be achieved (the action plan). Put another way, an implicit causal chain underlies each of the results areas. A fundamental point in this process is that it is not led autonomously by SEPLAG. Quite to the contrary, during the

entire procedure of creating and determining the strategy, discussions take place and morally binding agreements are entered into with sectoral agencies, thus lending legitimacy to this implicitly causal model.

Once top management at sectoral agencies agreed to the priorities upon which to act, an important meeting took place in February 2007, which resulted in a series of morally binding agreements. The governor presided over the event, accompanied by top-level officials of the state executive branch. “Commitment Notebooks” were distributed and agreements were reached by each secretariat to attain certain goals and projects.

These initial commitments were fundamental to aligning shared objectives. Once these objectives had been determined, it followed a process of design and detailing of strategic projects that made up the new strategic portfolio (57 projects) under the leadership of the GERAES Project Office along with the respective project managers.

FIGURE 2:

Strategic Map, Minas Gerais State Government

Source: Integrated Minas Development Plan 2007–23.

With the guidelines and policy design defined, undergirded by this implicit causal chain, implementation of these public policies was ready to begin. A basic premise of the M&E system in Minas Gerais is that of supervised autonomy, in which the planning unit signs the objectives and supports the sectoral agencies in laying down the path by which to reach the desired goals. However, it is up to each one of the sectors to find the best possible solutions for implementing the projects. Thus, implementation in the M&E system devolves on incentives for implementation. These are detailed in the next section.

Incentives for implementation: Results agreements and results committees

Based on the premise that the best road to results-based management is to generate an adequate set of incentives to implement the strategy, two instruments stand out: results agreements and results committees. Results agreements are management partnership contracts entered into by the governor and the state secretariats (first-tier agreements), and between state secretariats and their work teams (second-tier agreements). Keeping in mind the strategic vision and desire to strengthen the importance of targeted actions, as well as the notion of supervised autonomy, the M&E system spends more energy on the first tier, and leaves the respective sectoral agencies in charge of second-tier agreements.

This instrument, in turn, is an important tool in the effort to align the institutions with the government strategy, insofar as it brings together the main results and deliveries to be pursued by the different sectors under one contract. These documents reflect the final results sought (indicators for outputs and impacts with goals), the strategic projects (100 percent implementation of frameworks and goals planned), the sectoral agenda (set of outputs to be achieved), and expenditure streamlining indicators (savings goals with costing of activities and means of implementation, as well as quality indicators in budget execution at the agencies).

Each secretariat has a contract with a different correlation of weights among these four themes, depending on the features of each system. The Education Secretariat, for example, directs 70 percent of the contract to reaching the final results, with the remaining 30 percent going to the other themes. Other secretariats, however, show greater

balance. The Health Secretariat spends 30 percent on final results, 30 percent on structuring projects, 30 percent on the sectoral agenda, and 10 percent on expenditure streamlining indicators.

A major positive incentive links the logic underlying the results agreements. The outcome of the evaluations (first and second tier) determines who receives the productivity prize of up to one month's wages. The prize is linked to the proper implementation of works and the attainment of results by the state secretariats and their work teams. This instrument, created in 2004, became more relevant to the M&E system starting in 2007, when it was linked to the strategy indicated by the budgetary process and the four themes.

A fundamental role of the State for Results Program (created in 2007 in close partnership with SEPLAG) is to align the partnership contracts with the strategy, and more important, to carry out evaluations that harmonize and balance the different sectors. Taking as a point of departure the principle that some areas are more strategic than others, and the fact that there are different levels of difficulty in achieving the desired results, the State for Results Program endows the evaluation system in these contracts with simple and transparent rules that lead to a solid and lasting system.

Two rules drawn up by the State for Results Program together with SEPLAG illustrate this spirit:

- Create a bonus for those secretariats with the four largest budgets and a discount for the six secretariats with the smallest budgets (Decree 44.873/08). This makes it easier to balance these differing degrees of difficulty, given that more time is spent and greater audacity is called for when setting goals for large systems (transport, education, health, and social defense) compared with the smaller systems (culture, sports, advocacy, audits, agrarian reform, and tourism).
- Draw up permanent yet flexible rules to award bonuses for meeting the goals and adjust goals that turn out to be unreachable. This is necessary because of cases in which goals are agreed on for year $t + 1$, without evaluating the results for year t . This flexibility strengthens these contracts (Decree 45.191/09).

In addition to the alignment generated by means of the positive incentive contained in the results agreement and the subsequent awarding of a prize for productivity, the results committee is another important instrument. The State for Results Program's committee meetings, presided over by the vicegovernor, deal mainly with strategic projects by results area. Therefore, they are attended by those in charge of such projects (managers and top-level administrators at the responsible agencies). The objective is to systematize the vicegovernor's administration of government activities. At each meeting, an operational-strategic plan of action with a focus on policy implementation is drawn up. These meetings set the pace and keep the government's strategic planning going, thus mitigating the risk that operational routines end up dominating the agenda.

These committees, then, are concerned with implementing the strategy, insofar as "concrete experience provides clear evidence that, to be effective, it is not enough that the government's strategy be well-formulated. It is also important to watch after its introduction and implementation with the same care and even greater intensity" (Porto and Silveira 2006, 67).

These meetings involve a significant moral incentive, given expectations and the vicegovernor's accountability regarding the implementation of the planned strategy. Furthermore, such forums permit participants to analyze difficulties, risks, and opportunities when implementing projects, as well as generating support for sector agencies undertaking certain actions.

In brief, in the sphere of implementation, two important instruments are intended to encourage and in some measure monitor and evaluate the execution of public policy. The next section describes two important tools of the M&E system: the final results (a theme selected in the results agreement) and the development of evaluation of these public policies.

Results-based management and M&E

One of the main innovations introduced to the 2007–2023 PMDI deals with the creation of a set of final results with short-, medium-, and long-term goals in the 13 results areas. During 2007 and 2008, the State for Results Program launched an effort to strengthen the M&E system and establish a framework for the results-based management

model (to be evaluated annually through the results agreements). After two rounds of evaluation of the results agreements (2007 and 2008), and having reached consensus on a set of 104 indicators for the 2009 agreements, this M&E tool is fully implemented. The 2009 Indicators Notebook was a benchmark for this process (Guimarães and Campos 2009).

The evaluation logic verifies final results and represents an important benchmark in the system's construction. This approach goes hand in hand with the premise that sector agencies must be aligned with the stated objectives and seek the best possible path to reach their goals.

One successful example of this logic is the final result "citizen's fear of falling victim to violence". After the Secretariat of Social Defense and the Secretariat of Government (charged with communications) agreed on this indicator, the related goals (such as a reduction in the proportion of the population that feared falling victim to an act of violence) generated activities to disseminate information on preventive actions and the struggle against violence already underway. This joint plan of action is the public management model, in which all actors coordinate activities to attain the agreed-on results.

The idea of creating a harmonious M&E system includes a variety of approaches and results that coexist in equilibrium. In addition, the notion of a model in which strategic goals are correlated to more operational execution methods is strengthened (strategic projects and sectoral agendas). This allows for an evaluation that not only verifies the effort made by sector agencies, but also strives to be fair, taking into account the externalities assumed for some goals (a reduction in the homicide rate, for example).

Although monitoring of final results was implemented, the challenge of improving the maturity and quality of these indicators persists. In some results areas—quality education, for example—a set of results exists that is rather solid in terms of measures taken; in others, such as the network of cities and services, the indicators necessary for gauging some of the wished-for dimensions do not yet exist.

One innovation still underway concerns the creation of benchmarks for these indicators, as well as the development of methods capable of suggesting goals that, while challenging, are at the same time feasible to attain.

Another tool in the Minas Gerais M&E system is the impact evaluation of strategic projects. In 2007, when the new project portfolio was conceived, it was taken into account that it would be necessary to develop evaluations capable of testing the implicit causal models. Thus, the state government's strategic partners (the Minas Gerais Federal University and the João Pinheiro Foundation) were asked to evaluate nine social projects, with the aim of learning if, in fact, the actions implemented were achieving the intended impacts. This project evaluation culture is still embryonic in Minas Gerais, and the first results of these evaluations will be available in late 2009 and mid-2010.

The third wave in the consolidation of this M&E system will introduce public policy evaluation tools in their various stages (targeting, efficiency, process, impact, and so on). One part will be introduced toward the end of the current administration's period (2007–10), by means of "executive evaluations." These evaluations will look more closely at the elements still lacking before project planning processes become definitively rooted in a culture of evaluation.

As it stands, the evaluations have yet to entirely fulfill their role of feeding back into the public policy cycle. What does exist, however, is a clear understanding of the need to complete this M&E cycle. The final section of this paper traces some of the paths being taken to strengthen the feedback process.

Feedback: The current paths

It is up to the sectoral agencies to find the best ways to reach the agreed-on objectives. When these have been determined, implementation can be monitored and supported. Thus, the feedback process must come mainly from the sectoral agencies. In the light of indicator performance and the outcome of the evaluations, policy design and implementation may be reviewed, if necessary.

The planning process cannot be rigid but rather should remain open to changes in frameworks and goals. As long as a plausible rationale exists, the attitude toward such initiatives is open-minded, mainly during the annual planning reviews.

Inputs for this process of reflection usually have been offered at the meetings of the results committees, including, even more important, the delivery of strategic reports for discussion purposes. For example, at the May 4, 2009,

meeting, the Results Committee in the Quality Education Area brought up for discussion the results of the 2008 proficiency exam in Math and Portuguese taken by students in public schools. The evolution of pertinent indicators was described, along with the performance of the indicators in relation to short- and medium-term goals.

In brief, a space was created for the State Education Secretariat to explain the efforts being made in the early years of schooling, and to signal problems faced in middle school that explain why the indicators have stagnated. The Education Secretariat stated that, in addition to measures already underway, it will take corrective actions, such as reviewing the curriculum, that are thought to overburden students at this stage of the learning process.

This example hopefully will elucidate what is expected in terms of results-based management and making sectoral agencies responsible for developing new strategies. In this context, the role of the State for Results Program is to monitor and deepen some of the evaluations currently underway. In this particular case, these refer to the strategic project for vocational training at middle schools, whose goal is to improve education among middle school students.

Nevertheless, other activities also are being carried out by sectoral agencies as they seek to identify corrective action plans. This is the case, for instance, of the Transport Secretariat. Upon realizing that the proposed final result ("number of accidents per 100,000 cars") was overly ambitious, a workshop was held in September 2009 at which a Plan of Action was drawn up to help improve results.

Other examples of even more systematic interventions include the "accompaniment cell" for the Pedagogical Intervention Program. This small working group focuses on analyzing the results of the PROALFAexam, which is a reading and math proficiency exam administered to students in the third year of primary school. Results range between 0 and 800 points. A score above 500 indicates adequate literacy skills. The goal of the accompaniment cell is to help the management team improve its interventions.

Despite the foregoing, feedback as a systematic activity based on evaluations and final results has not yet been fully implemented in the state government. Part of the difficulty resides in the fact that indicators for some of the re-

sults areas are still immature; another hurdle is the incipient degree of consolidation of an evaluation culture in Minas Gerais.

Notwithstanding these challenges, some good examples already demonstrate how the introduction of a culture of results-based management has had a positive effect on the identification of good practices. It can be observed how the building of a well-balanced, harmonious M&E system, bolstered with appropriate incentives, can lead to a gradual change of culture in the public sector as it moves toward an effective results-based management.

Tadeu Barreto Guimarães
Eder Campos

Transparency Index of Public Entities: Experience in Evaluating Risks of Corruption in Government Administration at National, Departmental, and Municipal Levels

Background

Since 2002, *Corporación Transparencia por Colombia* has led the Government Transparency Index (formerly the Government Integrity Index), which is a regular, systematic study that aims to strengthen public sector institutions and encourage measures to control the risk of corruption in the public sector. It is the first exercise to quantify risks of corruption in state institutions conducted by an independent organization that monitors the risks of corruption from a civilian standpoint, and strengthens the definition of public policies in combating corruption.

With financial support from the Inter-American Development Bank, the Dutch Embassy, the German cooperation agency GTZ (CERCAPAZ Program), the British Embassy, U.S. Agency for International Development, the European Community, and its own resources, *Transparencia* has produced and published five editions of the National Transparency Index, four editions of the Departmental Transparency Index, and three editions of the Municipal Transparency Index.

Initially, the Corporation focused its efforts on producing the index of national public entities. The first edition evaluated 88 entities in 2002. The most recent edition evaluated 158 bodies of the three branches of government, au-

tonomous entities, and the electoral body for the 2007 and 2008 terms.

In 2005, the index created a special methodology for evaluating special state entities, such as state-owned industrial and commercial companies and public-private entities. In 2003, the Departmental Transparency Index began to evaluate government departments and agencies (Ministries of Finance, Health, Education, Public Works and Infrastructure, Traffic and Transportation, Comptrollerships, and Assemblies of the 32 departments). Lotteries and liquor stores also were monitored. The last measurement, of 2005 and 2006, suspended the evaluation of the Departmental Assemblies because the evaluation tool did not show other kinds of risks that do affect these entities. Importantly, in the case of national- and departmental-level indexes, *Transparencia por Colombia* defined the coverage of the indexes and took into account the characteristics of entities in terms of their functional importance and budget-handling levels.

In 2005, the Colombian Federation of Municipalities, the National Planning Council, and the *Corporación Transparencia por Colombia* formed an alliance with citizens to jointly launch the Municipal Transparency Index. The alliance called on the municipalities, with the exception of large cities, to participate in this measuring exercise and obtained answers from 146 municipalities that voluntarily joined the process. In addition to measuring transparency and institutionalism, this index pays particular attention to accountability to citizens and promoting civic participation in municipal administration. The first measurement of firms was conducted in 2007, and evaluated the 2005 and 2006 periods. The risk of corruption of 252 municipalities was evaluated with regard to centralized administrative management.

The fight against corruption as a starting point

Like poverty, corruption often is perceived to have no permanent solution; its impact on society might possibly be “reduced,” but it will never disappear. Many people, however, say that not only is it possible to combat corruption, but also it is necessary. Ethical, moral, and political forces demand fundamental changes in public administration and management processes. It is of the utmost priority for government and citizens to take action against corruption because in a country with high poverty levels, the people

should have access to public goods and services—to which they have full right.

Risks of corruption, and transparency of government entities

One of the core values of modern government management is democracy as a basis for exercising power. Democracy as the genesis of power and its rules constitute the political and ethical guide that organizes the relations between civil society and state. The rules of democracy to remember include consensus, controlled power, accountability, legality, and access to information, among others. All these rules are aimed at generating a space of trust in the relationship of social and political actors, trust that guarantees the existence of society as a whole, although not in the hypothetical idyllic picture of harmony, but in a context of diversity and pluralism in which tolerance and respect make coexistence possible. Trust in diversity is based on the assumption that all the actors who live in society develop their role within the framework of law, while the state as the lead actor represents the institutional framework that ensures the integrated management of public affairs, where collective interest prevails over private interests.

Viewed in this light, transparency appears as a medium that allows the different members of society to know what the government is doing. Its every action is of public interest, which is why the government has the utmost responsibility to provide information about its actions, and why it must also be willing to submit such information to public scrutiny. Such scrutiny may result in society's appreciation or sanction, depending on how well the power holders respond to citizens' demands. Thus, transparency in governance is one of the favorable conditions (coupled with institution building, political skill to manage social demands, a high level of trust, and the existence of social, political, and economic dialogue scenarios) that allow for public deliberations. These conditions are crucial for governance.

Corruption is one of the most serious problems affecting the democratic exercise of power. It deforms democracy, especially if it becomes embedded in the social, political, and economic fabric of society. Not only is corruption a phenomenon that speaks for the actions of individuals, but it also constitutes a nefarious feature of the "architecture" of government in all periods. The Government Transparency Index focuses on the architectural characteristics of public

institutions (and the practices of the individuals who manage those institutions—civil servants) to identify the conditions and practices that can create a risk of corruption in government. The analytical category of risk of corruption is understood by the index program as the possible occurrence of acts of corruption in public institutions because of the existence of certain institutional conditions and certain practices of employees that the evaluations, developed by *Transparencia por Colombia* in 2002, have shown to be susceptible to corruption.¹

Risk is a concept under construction that stresses the existence of an intermediate state between security and destruction. Pinpointing this state alerts people to the "perception of risk" and defines their thinking and actions. Risks of corruption in public institutions do not exist until an inquiry, investigation, or evaluation brings them to the attention of public servants and citizens and, most important, when this new information generates actions to counter the risk. Our primary purpose in evaluating risk levels of the entities is to promote institutional changes as a result of changes in the ideas and attitudes of public servants, specifically in relation to management of the public. Obviously the index is a civilian contribution. It is one among several government-modernization and democratization initiatives to emerge from the government, from teaching institutions, other civilian organizations, and international cooperation.

The Transparency Index evaluation is based on the assumption that corruption risks are reduced when management is transparent and abides by the established rules and standards, and when local authorities promote civic participation and social control. From this perspective and as a starting point for the evaluation, the following are risk scenarios:

- Monopoly of information and decisions, which affect the population in general
- A high degree of discretionary behavior by civil servants
- Poor development of institutional processes and procedures
- Absence of or weakness in checks and balances and controls of power, and institutional controls—internal and external—and control by citizens
- Zero or near-zero transparency of civil servants' actions

In these risk scenarios, public affairs stops being public and becomes private due to the existence of management processes and governance that are shrouded in secrecy. For this reason, the evaluation focuses first and foremost on the search for greater transparency in the activities of government personnel. That is to say, government action should be visible, governed by clearly established rules and procedures, and subject to public scrutiny. The formula is simple: the greater the transparency of public entities' processes, the lower the risk of corruption and therefore the lower the occurrence of acts of corruption.²

Measurement structure

Index results are not indicative of the level of corruption in an institution or sector, but rather of the corruption risks that entity faces based on its administrative management. Consequently, these results provide an input that enables entities to identify the areas and determine prompt actions to counter the corruption risks identified. The index identifies corruption risks in different core factors to be measured, each of which is composed of multiple variables and indicators. The factors, in the case of national and departmental indexes are transparency, punishment, and institutionalization.

Transparency

The indicators under this factor evaluate permanent exposure of information about administrative management of the entity. They evaluate the quality of the information and ease of access. Thus, they show whether management renders the procedures by which it operates transparently, and whether its results are submitted to outside scrutiny by the general public or private stakeholders.

Punishment

The indicators included in this factor examine judgments and sanctions—disciplinary and fiscal—related to behavior considered corrupt by our legislation. Going by these indicators, judgments against officers of a particular entity, issued by the supervisory bodies in cases of fiscal or disciplinary liability, are seen as a warning that the entity's internal processes are not sufficiently shielded or protected from significant risks of corruption. The indicators also note advance action by internal disciplinary control offices, but from a different perspective, in this case a greater response

from those offices in allegations against staff is valued as an important internal control mechanism.

Institutionalization

The indicators that assess compliance with legal standards, guidelines and policies, and procedures established to achieve the institutional mission are based on the premise that compliance with such standards, guidelines, and procedures limits the discretion of officials, politicians, and individuals in the exercise of public functions, and, thereby, reduces the risk of corruption.

In addition to these three factors, the National Transparency Index evaluates the corporate governance factor for government-owned industrial and commercial enterprises and for mixed economy partnerships. Here, an evaluation is made of the existence of formal principles and management processes based on transparency, corporate responsibility, and respect for the rights of shareholders and stakeholders.

The Municipal Index incorporates the civic participation factor, whereby an evaluation is made of all government initiatives that promote citizens' involvement in public affairs to influence decisions (on services, definition of priorities, spending, and so on) and exercise control over the performance of government authorities.

Purpose and scope of measurement

In terms of knowledge production, the index has established itself as the first independent instrument of civil society that periodically measures the performance of public entities in the prevention and control of their risks of corruption, and it enjoys great credibility in public opinion and in the institutions.

The design of this tool has made some progress at the national and departmental level, and has made notable progress at the municipal level. Today, the Transparency Index's objective indicators cover current issues in the anticorruption agenda. They include openness of information on government activities and decisions, accountability to citizens by governments, and publication of recruitment and selection and evaluation processes for public employment, among others.

The frequency of measurements has converted the index into a benchmark for institutions to intensify their institutional analysis, while allowing them to eventually compare

BOX 1:

Basic guidelines for collection, review, and appraisal of evaluation data

- The index uses objective information obtained directly from the monitored entities and secondary government sources and control bodies. The entities appoint within their organization an official spokesperson in the data collection process, to centralize requests for information from the entity, and delivery responses.
- Information from the entities is obtained through two processes. In the first, the entities complete a questionnaire designed by *Transparencia*. In the second, they review information published on-line and in citizen service lines.
- *Transparencia* obtains information directly from other public sector sources, which is required to construct the indicators that form the index.
- The information is consolidated in a single database and processed to calculate the indicators and the index for each monitored entity, with formulas defined in the methodological design.
- Data validity filters are applied. The calculations made and checked within *Transparencia* are sent out to the entities as preliminary results for their information prior to publication.
- All of the entities' doubts about processing and calculation are clarified, and any observations about possible misinterpretation of any data will be heard to incorporate the corresponding adjustments.
- The results are published.

Source: *Transparencia*, www.transparenciacolombia.org.co.

themselves with the rest of the world and with their own progress over time. The index is a yardstick for public opinion about the evolution of institutions in controlling corruption risks. As a tool that promotes improvement processes in the management of public entities, the index seeks to capture their interest in identifying corruption risks and to encourage the design of improvement strategies. Over the years, the index's results have become a reference tool for institutions to improve each indicator, raise performance in the graded factors, and grow more self-confident because they are monitoring their corruption risks and the effectiveness of the institutional steps to minimize them. Below are some of the elements that indicate this:

- The measurements made in 2005, 2006, and 2007 highlight a trend of improved scores for most of the entities monitored at the national and departmental level, particularly in aspects related to the "public exposé" of information concerning the management of the institutions and the increased public hearings of accountability to citizens. In the 2007 measurement, 75 percent of the entities held a public hearing of accountability.
- The issue of transparency to fight corruption has been incorporated into the agenda of a large number of entities. The proof is that 70 percent of assessed entities deliver information in a timely manner.
- The research team has met with more than 180 entities nationwide and more than 1,200 employees, becoming a tool of consultation and management for public officials. Entities have received warnings about their institutional weaknesses and have developed improvement plans guided by the results of the index.
- At the national level, and in some organizations at the departmental and municipal levels, the index has been connected to the strategic plans of the entities. It has been integrated as an indicator of quality management systems and institutional development in relation to transparency.
- Nationally speaking, the index has marked improvement processes in such areas as information published on Web sites, announcement of contractual processes, and accountability to citizens.
- The entities that are sources of information have contributed significantly to the process of strength-

ening the index. The management teams of the General Comptrollership of the Republic, the Attorney General's Office, and the Accountant General's Office have shown a timely response in providing their information. Likewise, the Attorney General's Office and General Comptrollership have been working toward improving their information systems and fulfilling their monitoring and control functions.

- The National Transparency Index has become one of the targets of the Information System for Monitoring Government Action (ISMGA). The National Vice Presidency, headed by Dr. Francisco Santos, has proposed for the next measurement that national entities should aim to obtain at least 90 points in the index. The analyses and recommendations of the national and municipal indexes, delivered in December in two publications, have resulted in an interinstitutional meeting and a meeting with civilian stakeholders on issues of participation and consolidation of reports from local authorities. Both meetings were coordinated by the Vice Presidency and the Presidential Program to Fight Corruption.
- In response to the index's poor results, the Presidential Efficiency, Modernization, and Anti-corruption Program proposed that corporations subscribe to transparency pacts, seeking to engage these entities via voluntary agreements on issues like recruitment, meritocracy, and information systems to ensure their integrity. It will also take the results of the Municipal Transparency Index and include them as a substantial part of the Transparency Pacts, which this entity endorses with the municipalities.

In public policy, the index provides a benchmark to analyze the impact of public policy to combat corruption in the management of the entities and to generate recommendations for public policies aimed at improving institutional control and transparency. The variables covered by the index are related mostly to the policies pursued and the current regulations aimed at reducing the risks of corruption, so that their periodic measurement reveals, at least partially, that the changes produced in the management of the entities are explained by the attention paid to such policies and standards.

The Transparency Index enjoys credibility with public entities and the government. In fact, following advance

coordination work with *Transparencia por Colombia*, the Vice Presidency, the National Department of Planning, and the National Bureau of Statistics have shown interest in facilitating and enhancing the availability of public information so that it will strengthen the index and other gauges of governance, and boost their use in defining public policies.

Some evidence shows certain interdependency between the index results and measures taken by the government to obtain stronger institutions and transparency. Thus, based on the results of the index and the possible adjustments and improvements in public institutions that they trigger, the government and the vice president have publicly committed to concrete measures to entities that have ranked in the ranges of high and very high risk of corruption. In addition, some departmental comptrollers have used the results obtained by some departmental entities to define strategies and policies to support their monitoring function.

In terms of public opinion, the index offers a reference to the evolution of public institutions and promotes involvement by other stakeholders in the field. To date, publications of the index have been well received because of the independent view it represents. The results of its application have been disseminated widely in the media, thereby increasing both its visibility and public pressure on institutions to participate in the measurements and to implement improvements.

The challenges of measuring

All these advancements need to be consolidated and projected to new dimensions that increase their potential impact. Reducing corruption and corruption risks is not an end in itself but a necessary step toward more effective, fair, responsible, and reliable government. Corruption is a major concern because of the effect it has on society and institutions.

The effects of corruption can spread across all sectors and areas of public and private life. Corruption is a threat to democratic institutions and democracy itself. For this reason, it is crucial to increase levels of transparency and to prevent and combat corruption; and, in this way, to help increase effectiveness and accountability and to consolidate and increase public confidence in democratic institutions and governance. In this vein, the Transparency Index will continue working in the following areas:

- Continue the work of measuring the three levels of government using a tool to detect the transformation of the corruption phenomenon and institutional development of organizations.
- Strengthen the analysis of the results obtained through the inclusion in the program of lines of qualitative analysis.
- Increase opportunities for discussion with evaluated entities in such a way as to make in-depth institutional adjustments in those entities.
- Increase opportunities for promotion and dialogue with representatives of social, political and financial organizations, and with international cooperation on the results of analysis (both quantitative and qualitative) to provide information and create public awareness on the importance of transparency in government management.
- Develop steps to promote adjustments and improvements in government policies to guide the modernization of the state and the fight against corruption.

Conclusion

In general, Colombian municipalities are on the frontlines in the fight against corruption. Their institutional capacity, however, is weaker than the central government's and inadequate to deal with the scale of the problems and citizen demands they must address. This situation proves their low degree of effectiveness and legitimacy. Recent analyses also warn about a rise in corruption at the municipal level, adding that this phenomenon has become a serious obstacle to increasing equity, prosperity, and local democracy. Given these circumstances, some voices even argue for the need to reverse the process of decentralization, which is just two decades old.

The decentralization process is not responsible for corruption, although it has highlighted this phenomenon. To the extent that decentralization has brought power and resources to municipalities, it has created an attractive opportunity for corruption, which has become a barrier to effective attention to the social demands of citizens who expect them to be solved. Undesirable or perverse effects of decentralization are explained largely by the degree of development of local political culture and institutional designs and practices in the municipalities. The truth is that within the complex context of the municipalities, in terms of

acute poverty and exclusion of large sectors of the population, coupled with the presence of an armed conflict, more transparent, efficient, capable, democratic, and reliable local governments are required that can ensure better governance.

From this perspective, the Municipal Transparency Index seeks to make Colombian towns a cornerstone of the state structure. And it is betting on improving the institutional conditions of transparency as a formula for securing the welfare and development of communities and achieving better local governance. The systematic exposure of information to teams, government, and public opinion about corruption risks associated with institutional conditions and practices within the municipalities will motivate them to improve, as has happened with national entities. In controlling risks of corruption, we expect to break one of the strongest barriers that prevent local governments from achieving greater levels of efficiency in addressing citizens' demands. And by analyzing a significant number of municipalities of varying size (population and fiscal) and in different regions, it will be possible to deduce concrete policy recommendations for public transparency and the democratic and institutional development of Colombian municipalities. To encourage these changes, a medium-term intervention is needed to keep up the level of continuous analyses and propaganda, institutional improvement, and favorable public policies to form a virtuous circle of transparency, effectiveness, and confidence. This is the purpose of the Municipal Transparency Index.

The road to the consolidation of the state as honest, capable, and legitimate is long. Measurement of corruption risks implemented since 2002 can achieve the greatest gains in the promotion of internal institutional adjustments and advocacy on issues of transparency and risk control of corruption. Measurement also has received the attention of government agents and public authorities.

In this sense, the gains and improved overall performance of the index should not divert attention from areas in which adjustments are still necessary. Decisive action is needed in the formulation of public policies against corruption and against those that compromise administrative processes. Although much progress has been made in a number of national institutions, much remains to be done regarding the management of recruitment and human resources and on issues such as transparency of management processes, and adjustments to planning, evaluation, and control

processes. Institutional arrangements regarding meritocracy, the wage gap, assessment of human resources, and training are still weak. Moreover, the continuing existence of institutions with a high risk of corruption is mainly a result of undue influence in the selection processes and appointment of officials. All this has a negative impact on the efficiency of these institutions and citizens' confidence in them.

We need to learn more and more about the institutional conditions and practices of public servants that explain the management of entities and account for society's high or low confidence in government action. Corruption can occur in different ways while blending into the management processes to the point of becoming invisible. For this reason, *Transparencia por Colombia's* work over the last seven years has generated important information, methods, and tools, which must be collected and used in the fight against corruption.

Given its independent nature, thoroughness, and objectivity, the index has proved to be a valuable input to promote institutional change, guide public policy, and mobilize citizens. Certainly, progress has been made in understanding the phenomenon of corruption from the results of the index, but this is a complex and dynamic phenomenon, which changes constantly and imposes new challenges. Among these, an important ongoing challenge is to explore innovative ways to detect new expressions and manifestations of corruption and to encourage institutions and governments to act on them.

Additionally, although the media has widely publicized information produced by the index, it has yet to exploit a wider stage with a view to raising public awareness of the problem and ensuring that social organizations are armed with data to support their social vigilance of public management.

Elisabeth Ungar Bleier

Executive Director, Corporación Transparencia por Colombia

Challenges for Evidence-Based Decision Making: The Experience of Chile and Chief Features of Its Design, Implementation, and Use of Information

Introduction

This paper presents some of the challenges faced by the development of monitoring and evaluation (M&E) in opera-

tions in Chile. The process of resolving those challenges has identified key features of good M&E design, implementation, and use of information. The paper focuses on the instrument of program evaluations and institutions, based on the important role the latter have played in supporting decision-making processes. The paper provides a brief description of the definitions adopted to provide data on the performance of programs and public institutions. It presents alternative methodologies and their main features, and concentrates on the functions that the evaluated institutions develop. It also describes the type of external assessment and the process through which evaluations are conducted and used, and presents the differences compared with the commencement period of 1997–99. The paper further describes ways in which the data yielded by the evaluations are used in different decision-making processes, and the people involved. Finally, the paper discusses considerations and lessons learned when building a program evaluation system and ensuring its function and development over time.

Definitions adopted to acquire information on the performance of public programs and institutions

Chile has been no exception to the desire in many countries to improve the quality of public spending. In the early 1990s, the Finance Ministry began developing Management Control Systems and Budgets by Results, a key part of M&E of public programs and institutions. The M&E developed in Chile has involved a group of definitions on the performance of programs and institutions, and has yielded the following characteristics.

Existence of different instruments

Chile's M&E consists of instruments that measure performance and evaluation through different instruments or methodologies. This use of different instruments is based on experience; a single instrument or a single methodology proved to be insufficient to obtain the required information and explain the functioning and results of the many public actions.

Combination of instruments

The instruments are based on methodologies and formats of different degrees of complexity, thematic scope, and requirements. This takes different technical skills, generates different kinds of performance information, with different

costs and application times, and enables different uses. The M&E installation has used the budget process as both a basis and an incentive. Its development takes into consideration budget control, mainly derived from its annual cycle, where well-defined, regulated, and repetitive stages allow instruments and processes to be integrated and tied to specific work targets, and generate cumulative learning experiences.

Performance data

The use of performance data in decision-making processes for resource allocation also considers other data categories, such as political-programmatic priorities and constraints imposed by fiscal policy. That is, the link between performance data and funds-allocation decisions ensures that the information is integrated into a decision-making process that is both technical and political.

Political will

Political will refers to the institutions leading the process. Although it is true that the initiative to develop tools for M&E emerged in the 1990s, it has received basic support from the budgetary and fiscal authority throughout its development and operation. It was promoted by technical areas of the Budget Office of the Ministry of Finance—at a second level of authority, not from central bodies such as the Presidency or Congress as in other experiences.

Public information

Since the first tools were installed, it was established that the information they generated would be public. This principle has been observed mainly by forwarding data reports to Congress and publishing them in the Budget Office's Web site.⁴

Gradual development

The M&E instruments were introduced gradually, starting in 1993. Despite the fact that their processes and effects in the first stage evolved slowly, and in some aspects even receded, the gradual incorporation of methodologies has enabled the persons concerned to learn, correct, perfect, socialize, and generate a critical mass base for a greater sustainability of the system.

Design of institutions

Design of institutions means operatives and leadership from the Budget Office of the Ministry of Finance. The

M&E is performed by institutional players with roles and responsibilities. This takes the form of assignment of people in charge in the institutions or public services covered by the instruments. The existence of an interministerial committee responsible for making decisions for key aspects regarding the application of evaluations, formed by the Ministry of Planning, General Presidential Ministry, and Finance Ministry's Budget Office. This includes the definition of technical, operative, and financial handling by the Finance Ministry's Budget Office.

Supply of the M&E instrument: The importance of evaluation

The Chilean M&E is composed of performance-indicator instruments and evaluations. The indicators are measuring tools that use formulas, which must contain the necessary variables. For example, an indicator covering a program for school meals must use data regarding student population served by the program and information about the population eligible to receive the benefit, that is, the potential beneficiary population.⁵ The ratio between both variables will measure the program's coverage.

Despite the importance of the measurement indicators, they do not explain the reasons for whether or not a certain performance level is obtained. This limits their use in decision-making processes, making it necessary to complement the information with other data of a partial, qualitative nature, or taken from other performance data-collecting tools. In the latter case, evaluations have proved valuable in Chile to complement the shortcomings detected, generating a mutual support between the two instruments.

In 1997, the government introduced various program evaluation lines, in effect, to obtain more complete information about the performance levels experienced by public actions and to better understand the reasons or causes for those levels. Initially the evaluation of public programs was incorporated in a line of flash or desktop evaluations, called government program evaluations. Later the impact evaluation line was introduced, which was followed by the institutional evaluation called comprehensive spending evaluations and, more recently, new programs evaluation design.

With the application of these different evaluation lines, this tool has become a key piece of M&E. In effect, given the different types of information it generates, it has allowed for a more intensive use in the decision-making pro-

cesses; it has facilitated the job of informing outside players and has produced a base for following up any changes that the programs and institutions have had to introduce as a result of the findings and recommendations of evaluations. Additionally, the evaluation process has been a learning space for the organizational units evaluated and the Budget Office. This has helped develop skills. Below are brief descriptions of the different lines of evaluation.

Flash or desktop evaluations

Government program evaluations are based on the basic concepts of the logical framework matrix, supplemented with the review of a set of topics. The evaluations use available information (mainly simple polls), existing or background data that can be quickly systematized and transformed into relevant data, interviews with key players, and, in some cases, the application of other simple qualitative methods. They are conducted by a panel of three experts. The panel includes an expert in evaluation, an expert in the programmatic sector or area to be evaluated, and an expert in management and administration of public programs. The evaluations last between four and six months. The evaluation begins with the development of a logical framework matrix (mainly simple polls), which includes different levels of program objectives through identification of its end and purpose (overall objective), the specific objectives of each one of its components, its main activities, and their corresponding performance indicators. Table 1 is an example of a logical framework matrix.

The evaluation can begin to be developed after the matrix has been constructed, updated, or perfected, and used to gather data on the following topics.

Justification

Justification is the stage in which an analysis is made of the diagnosis or problem that triggered the program and that the program hopes to solve (or help solve).

Design

The evaluation of the design consists of analyzing the organization and consistency of the program, and the components and activities. Design evaluation considers the program's initial justification and how these diagnostic elements could have evolved. It uses the logical framework matrix, specifically the vertical logic analysis.

TABLE 1:

Logical Framework Matrix

Description of Objective	Indicators/ Information	Means of Verification
End:		
Purpose:		
Components:		
Activities:		

Organization and management

Evaluation of organization and management consists of the analysis of key institutional aspects within which the program operates, key work processes, coordination mechanisms within the institution and between institutions, if applicable, and instruments or mechanisms available for operations and reaching objectives.⁶

Results

The results or performance evaluation focuses primarily on assessing the effectiveness, efficiency, and economy of the program, observing its processes, products, and intermediate and final results, as appropriate and feasible, depending on the available information. The evaluation of the results or performance is made by applying the horizontal logic of the matrix.

Sustainability and continuity

Based on analysis of the design, organization and management, and outcome or program performance, the evaluation incorporates judgments about the program's sustainability in terms of institutional capacities and whether or not it should continue considering the evolution of initial diagnosis and results achieved.

Finally, the evaluation must present conclusions and recommendations, identifying weaknesses and shortcomings that affect program performance, while consistently suggesting actions to improve performance, including institutional and organizational, programmatic, or managerial (operational).

Impact assessments

The flash or desktop assessment methodology makes use mainly of available information. As a result, the information delivered in some cases is incomplete. In most cases,

it does not specify how it meets the needs of the program. This prompted the introduction in 2001 of the impact assessments line. The line of impact assessment methodology is designed to measure the final results or net effects of a program on the beneficiaries, that is, the efficiency in the final stage of the value chain. It uses differences regarding baselines (initial diagnosis), compares these differences with control groups to isolate external effects, and also uses other quantitative techniques to obtain precise measurements. It requires high technical skills and fieldwork. It is more expensive, lasts between 12 and 18 months, and is conducted by consultants or universities.

Although the gist of impact assessments is to measure the effects of the programs, in the Chilean experience, this focus has been supplemented with the review and evaluation of specific complementary topics. These topics provide other categories of information that help understand the

observed results. In effect, depending on the needs for information, the following topics, described above, are added: justification, design, organization and management, results, and sustainability and continuity. Finally, as in the line of flash assessments, the evaluation findings and recommendations must be submitted.

Comprehensive spending evaluation

The comprehensive spending evaluation line began in 2002 and is an evaluation of an entire organization or agency. It mainly uses available information, existing background data that can be systematized during the evaluation period and transformed into relevant information, and interviews with key actors and application of simple qualitative studies. It aims to evaluate all the policies, programs, and procedures arranged into the following topics or categories of information (see Table 2).⁷

TABLE 2:

Working Design for Evaluation of Programs and Institutions (ex post), 2000–09

Stage	Actors
1. Selection of programs or institutions. The evaluation process begins with the selection of programs or organizations (or agencies), depending on whether the evaluation will focus on programs or include expense.	Selection is made with the participation of the National Congress, based on a proposal submitted by the Finance Ministry's Budget Office. The Ministry of Planning and the Presidential Secretariat participate in preparing the proposal.
2. Selection of consultants or consulting firms. After the selection of programs and institutions comes the selection of consultants or consulting firms (including universities) through public bids.	Consultants or consulting firms are selected through public tenders held by the Budget Office, supported by selection committees that include representatives from the Planning Ministry and Presidential Secretariat.
3. Preparation of basic information. At the same time that the evaluators are selected, the commencement of the evaluation process is announced to the ministers and organizational units to be evaluated, and they are asked to prepare the basic information for conducting the evaluation.	The information is prepared by the organizational units that are to be evaluated and coordinated by the Budget Office.
4. Mention of critical aspects of the unit to be evaluated. Remarks are made to the evaluators about those critical areas of programs or institutions, so that they can be reviewed carefully in the evaluation process.	One or more meetings are held, which are attended by the evaluators, the persons responsible for the program, or the institution being evaluated and the Budget Office. Meetings may also be held with the Congress for the same purpose.
5. Training and preparation of frameworks. In the case of flash or desktop evaluations, before their commencement, both the evaluators and the professionals who work in the programs' responsible units are trained in evaluation methodology and logical framework basics. For Impact Assessments and Comprehensive Spending, the basic methodological elements to be used by the evaluations are socialized, and in the latter instance, the preliminary evaluation framework is prepared.	Training is given by the Budget Office, and initial preparation of the logical frameworks is done by the evaluators along with the evaluated unit, under the coordination of the Budget Office.

(continued on next page)

TABLE 2:

Working Design for Evaluation of Programs and Institutions (ex post), 2000–09

Stage	Actors
6. Evaluation. Evaluation begins with available background data as a base. The execution stage of the evaluation requires progress reports and a closing report to be submitted. These are commented on by the institution under evaluation and the Budget Office.	The evaluation is made by outside evaluators. The technical and administrative counterpart is the Budget Office. The organizational units evaluated participate by submitting information, attending meetings, and checking reports.
7. Analysis of evaluations. As soon as the evaluation is finished, the evaluated organization analyzes the final report and issues an Institutional Response to the evaluation. The report also is reviewed in the Budget Office, which pinpoints the changes needed and the possible effects they would have on the budget.	The analyses are made on the one hand by the evaluated organizational units, and on the other, by the Budget Office. Participants in the Budget Office's analysis are budget experts, experts in charge of evaluations, and the budget director.
8. Forwarding to Congress and presentation on Web site. The final evaluations report, an executive report prepared by the Budget Office and based on the final report, along with final comments from the evaluated institution (institutional response), are officially sent to the National Congress.	The budget director sends the evaluation reports to Congress, and the Budget Office posts the data on the Web site.
9. Forwarding to Congress with draft budget. Incorporation of the data added in the public finances report and executive summaries in attachments to the draft budget.	The budget director sends the Public Finances Reports to Congress.
10. Formalization of commitments. Based on the evaluators' recommendations, commitments are defined between the evaluated institution and the Budget Office.	Commitments are established between the Budget Office and the evaluated institution (ministry).
11. Follow-up. A follow-up of institutional commitments in keeping with recommendations is made every six months.	The Budget Office follows up on improvement commitments, based on the data sent by the evaluated organizations.
12. Report of fulfillment of commitments and analysis of follow-up data. The Budget Office analyzes data tracking commitments in the Evaluation of Budget Execution stage.	The Budget Office analyzes the follow-up data.
13. Forwarding report on fulfillment of institutional commitments to Congress. This is sent twice a year, as part of the Public Finances Report and Report on Evaluation of Financial Management by the Government, which also updates information about completed evaluations.	The budget director sends the government financial management evaluation reports to Congress.

Source: Author.

Institutional design

Institutional design considers evaluating consistency among government priorities, mission, strategic goals, strategic products, and their beneficiaries. Similarly, based on the foregoing, it evaluates the coherence of institutional structure and division of responsibilities between work units that make up the body or public agency.

Institutional management

The aspects evaluated in institutional management relate to institutional capacity (professional, technological, organizational) and management mechanisms or procedures applicable and relevant to the organization, such as (1) mechanisms for coordination and allocation of responsibilities; (2) allocation mechanisms, funds transfer, payment procedures, and audits; and (3) accountability and transparency

in the use of resources, activities, M&E tools, and targeting criteria, or selection of beneficiaries.

Performance and use of resources

Performance and use of resources consists of evaluating the performance of products and strategic subproducts,⁸ and linking this information to the information about the resources assigned to providing them. Finally, as in the line of flash evaluations, the evaluation findings and recommendations must be submitted. Comprehensive spending evaluations last about 14 months. Considering the various activities that need to be undertaken, it was decided that they should be carried out by consulting firms or universities selected through a public bidding process.

Evaluation of new programs

The new program evaluation line was introduced in 2009 and is developing the design to include evaluation of new public programs from the beginning to ensure the availability of the basic elements and information that will enable the effects to be measured with the utmost rigor once the program is up and running. The evaluation design considers the baseline identification with regard to the variables that will be affected by intervention and construction of the control group and uses the best possible methodologies to apply to each program. An external agency experienced in quantitative techniques and data collection and management designs the ex ante evaluations. The development of this evaluation line has the technical support of an international advisory panel established for this purpose and composed of academics with extensive experience in the area of impact assessments. This panel meets annually in Chile and reviews the progress of the designs concluding with a detailed report of comments and recommendations.⁹

Principal features of operational design

Outside aspect of the evaluations

Chile opted for an external evaluation process for the organizational unit responsible for the program or institution being evaluated, and that also is external to the ministry to which the program or institution belongs. Indeed, evaluations are coordinated from the Ministry of Finance through the Budget Office, and are executed by external

consultants organized into panels of experts, universities, or private consultants. In all cases, the selection is made through public tenders.

This process does not mean that the institutions evaluated do not participate but rather that the form of this participation should allow all facts and data available to be gathered, while safeguarding the objectivity of the process. In other words, they provide information, participate in technical meetings, and comment on reports, but they do not play the role of technical or administrative counterpart.

Operational process and actors

The evaluations are conducted according to a working design that orders the different work stages of the evaluation cycle, and incorporates the use of information. Table 3 shows these different stages, in sequence, and the actors involved in each one.

The first phase of the Chilean experience did not consider all the stages that characterize the current process, nor were they developed with sufficient technical quality. This jeopardized the quality of information delivered and its use. The lessons learned, however, from this first period (1993–2002) allowed for a number of changes to be introduced. Moreover, during the last 10 years, the operative process has undergone a continuous improvement.¹⁰

Process for selecting programs or institutions for evaluation

The decision of “what to evaluate” should consider the demand for information. In Chile, as in most countries in the region, the demand for appraisals to learn more precisely about government actions and performance exceeds the resources available. As a result, every year only a limited number of assessments are made, and it is necessary to select those programs and institutions to which each of the available methodologies will be applied. In 2009, for example, 20 quick or desktop evaluations, eight impact evaluations, five comprehensive expenditure evaluations, and five new program evaluations were conducted.

The choice of what to evaluate arises from a proposal prepared by the Budget Office and presented to Congress for its final definition within the context of discussions

TABLE 3:

Uses and Actors in Ex Post Evaluations

Uses	Actors
<p>Review of evaluation results in the Finance Ministry's Budget Office defining guidelines to establish pledges for improvement and identifying budget impact.</p> <p><i>Period:</i></p> <ul style="list-style-type: none"> GPE evaluations before the start of the budget formulation Evaluations on impact and comprehensive spending as soon as each is completed <p>Incorporation of data from the evaluations into meetings of the Technical Commissions as information on which to base decisions.</p> <p><i>Period:</i> Draft budget formulation process</p> <p>Summary report of the evaluations and compliance with commitments for better dissemination of the public finance report and its submission to Congress.</p> <p><i>Period:</i> Presentation of Draft Budget Law to Congress</p> <p>Presentation of evaluations data to Congress in a document complementary to the Draft Budget Law.</p> <p><i>Period:</i> Discussion of Draft Budget Law in Congress</p> <p>Preparation of pledges to improve evaluated programs and institutions.</p> <p><i>Period:</i> As soon as the evaluations are completed and their results have been analyzed</p> <p>Overview of evaluation results and compliance with commitments published in the "Evaluation of Government Financial Management" and submitted to Congress.</p> <p><i>Period:</i> May to June each year, before commencement of budget formulation</p>	<p>Budget Office:</p> <ul style="list-style-type: none"> Management Control Division's Evaluation Unit Head of Management Control Division Head and analysts of budgetary sector to which the evaluated program or institution belongs Budget director <p>Budget Office:</p> <ul style="list-style-type: none"> Head of budget sector Sector analysts <p>Ministry or institution:</p> <ul style="list-style-type: none"> Undersecretary Services chiefs Other institutional headquarters and analysts <p>Budget Office:</p> <ul style="list-style-type: none"> Management Control Division Director <p>Congress: Mixed Budget Commissions</p> <p>Budget Office:</p> <ul style="list-style-type: none"> Assistant director Head of budget sector <p>Ministry or institution:</p> <ul style="list-style-type: none"> Undersecretary Services chiefs Other headquarters and institutional analysts <p>Budget Office:</p> <ul style="list-style-type: none"> Management Control Division Evaluated ministry or institution <p>Budget Office:</p> <ul style="list-style-type: none"> Management Control Division Budget director

Source: Author.

about draft budget laws. The final selection is established in an agreement protocol signed every year by the Ministry of Finance and members of the Special Mixed Budget Commission after the Budget Law has been passed.¹¹ That proposal, which represents the relationship between supply of and demand for performance information, takes into account the following aspects:

- Preliminary performance data of programs or institutions that the Budget Office holds in keeping with its functions, both regarding budget and management control (in addition to performance

and evaluation indicators, the current management control system includes instruments such as the management improvement programs and the integrated management inventory, which also gather information).

- Status of programs or institutions as to whether they have been evaluated in the context of the evaluation program run by the budget office, or by other evaluation processes, and the year when that happened.
- Time the program has been in existence, and its priority.

- Suggestions from the General Presidential Ministry-Secretariat and the Planning Ministry, both members of the Interministerial Evaluation Program Committee.
- Suggestions put forward by the special mixed budgets commission and congressional budget sub-commissions during the processing of the draft budget law in Congress.

Ways to use the information: Incorporation into the budget cycle

In terms of resource allocation, the use of performance data that emerge from the evaluations has not translated into an immediate mechanical step to allocate more or fewer resources depending on the results of the evaluation during the period, but rather the data have been integrated into more comprehensive analysis processes. One way to use the evaluations is to integrate the data at different stages of the budget cycle. A review of evaluation results in the Budget Office implied the following:

- Guidelines are defined for making commitments to improve evaluated programs or institutions, classifying programs into categories according to the effects or implications from the evaluations.
- Budget impacts of component adjustments are identified if they correspond to findings of the evaluation, or extensions if warranted by those findings, if the financial resources are available and if they represent a high priority (high relative importance in terms of established policies).
- Data from the evaluations are incorporated into the technical commissions' meetings during the budget formulation process, to analyze and lay the groundwork for decisions among the Budget Office, the ministers, and their institutions.¹²
- Evaluation-derived data, together with the draft Budget Law and also the data summarized in the public finances report, are presented to Congress.
- The evaluation data in the Budget Office's evaluation of budget execution are presented for analysis along with financial performance data, and summarized data are incorporated into the annual publication "Evaluation of Government Financial Management."¹³

Pledges to improve programs and institutions

The evaluations also can be used for the commitments made to improve the programs and institutions evaluated. These commitments are made on the basis of an analysis of the assessment's findings. These commitments promise that actions will be undertaken in the coming semesters, and they are formalized by the Budget Office. Subsequently, this office monitors compliance with promised actions and specific aspects of the changes introduced. The commitments concentrate mainly on improving design and management, and imply that the evaluations will be used for purposes other than those directly related to resource allocation. Table 3 shows the different uses of evaluations and the actors involved.

Main challenges

The development of evaluations as part of the M&E and the growth of this line over time has allowed important insights that, introduced in a timely manner, have granted sustainability. New challenges and risks, however, are even more complex than those faced in the early stages. Some of these are described below:

- Optimize the M&E cycle for the purposes of evaluation. Advance the link between evaluation, monitoring incorporation of changes or improvements and ex post measurement of performance or effects. The evaluation aims to improve the quality of spending, which means actual achievements in the purposes of the programs and their contribution to policy goals. The commitment to improve arising from evaluations and compliance therewith is essential to improving program results. Achieving better results is the intended effect of the implementation of evaluations. So how does one know whether these effects have occurred? A first step to get an approximation of these effects is to observe measurements by comparing performance indicators over time. This line of work, seemingly simple, is a major challenge for the links between evaluations, indexes, and reevaluations.
- Integrate the evaluation function. Move toward more coordination with other elements of the evaluation function, such as investment appraisal, evaluation of government policies or goals, and so on. This raises a challenge to harmonizing technical,

operational, and institutional facets, particularly because these areas have different levels of development, have different features, and are located in different institutions. It entails risks as well since the integration of functions and institutional relocations may affect current achievements, especially regarding the use of evaluations.

- Consider the growth in number and complexity of the evaluations and care about their quality, timeliness, and use. How many more evaluations can we make without affecting their quality or the timeliness of the information available (availability of evaluation results to use for decision making)? In Chile, the number of evaluations has increased in the last three years. This makes it necessary to review the ratio between growth and the quality-timeliness-use trilogy, before continuing to increase evaluations.
- Disseminate the information. Despite all the propaganda efforts made, in the form of publications and information available on the Web site, press conferences, and other tried-and-true outlets, greater transmission of the evaluation results and their effects still is needed. This information is need for both the actors involved in the execution and the beneficiaries, with the latter requiring a different form and language.
- Conduct self-evaluation, goal assessment, and continuous improvement. Considering the progress achieved and the size of Chile's Evaluation Program, it is a challenge to maintain this regular evaluation, either in the form of self-evaluation of process and uses, or external evaluations in the form of meta-evaluation as carried out by World Bank in 2004–05.

Lessons and recommendations

The key lessons learned and recommendations regarding the development of evaluations arising from the Chilean experience, the experience of other countries in which the author has been involved, the review of documentation, and participation in workshops and international seminars in the region are the following:

- Consider the purpose of the evaluation, methodology, and subject covered. It is recommended that

the evaluation consider definitions on three key interrelated aspects: (1) the purpose of the evaluation, (2) the methodology to be used, and (3) the focuses on which each evaluation will concentrate.

- The object of an evaluation may be a program, a component of a program, a set of programs with a common goal, or even an organization. In this regard, the Chilean case shows that it is possible to initiate an evaluation system of between five and eight individual programs within two years, gradually increasing the number, and always prioritizing the country's greatest needs on the question of performance data.
- Having identified the object of evaluation, it is then a question of deciding what methodology will be used to perform the evaluation. In general, this selection must consider the resources available to make the evaluations; organizational, financial, and professional skills; the runtime required for each program; former evaluations available, if any; and the date when the information is required. Chile began with flash or desktop evaluations using a simple methodology that allowed the development of skills that later made it possible to implement more complex methodologies.
- After selecting the object and the methodology, the most important subjects to be addressed by the evaluation can be specified. To this end, it is advisable to gather the opinions of people responsible for the programs and the actors who would be using the information (technical professionals, authorities, parliamentarians).
- Consider progression. Evaluations should be developed gradually, as far as selection of the object of evaluation, methodology, and focus are concerned. They should be accompanied by the corresponding operation processes. It is possible to start program evaluations using a simple methodology, and subsequently, depending on progress, move toward more complex methodologies. The number of evaluations to be performed each year may be increased in keeping with the resources and skills available.

This pattern requires progress in constructing internal task forces to ensure strong partners, while at the same time

constructing an evaluators' market. In this area, a team of between five and seven people should be trained to perform the technical leadership role. These people then would train other professionals or supervisors (cascade training), prepare technical guides, train evaluators to act as the technical counterpart of the evaluations, and learn from the experience to perfect the system in the following cycles.

- Consider the availability of information and limitations of the evaluation process. The availability of information is central to evaluation processes. Whatever the methodology, it must use information. Failure to do so will limit the evaluation process. The true scope of the evaluation in terms of implementation of methodological elements therefore depends on available data, in addition to the time allotted for the evaluation process and technical competence of the evaluators and counterpart. Consequently, it is recommended that special attention be paid to the generation and collection of information as part of a regular work practice.
- In many instances the evaluation will face a situation in which the units responsible for the program will not have complete or reliable information on crucial aspects of performance. That could hinder the identification of the program's baseline and prevent any changes from being measured, or it could prevent core performance-gauging indicators from being measured. In these cases the focus should be on evaluating other aspects, such as program design, basic management processes, and the efficiency level of products (outputs).
- Develop a mechanism for following up the recommendations. The design of an evaluation system should not end with the delivery of the final report, or after completing the analysis of its findings. Along this line, a phase should be incorporated to facilitate, promote, and encourage compliance with the objective of "improving" after an evaluation, introducing changes into the "object evaluated" based on the findings and recommendations that the evaluations indicate. Indeed, the recommendations should be seen as an opportunity to discuss and commit to changes by the institutions or organizational units responsible for the program evaluat-

ed. Additionally, these improvement commitments must be monitored every six months to one year to verify the incorporation of agreed-on measures.

- Consider the institutional and basic operating model. For evaluations to fulfill their purpose of providing feedback for decision-making processes, basics such as institutionalization and implementation management should be considered. In fact, the development of evaluations and their use do not happen automatically. They require new practices and skills that involve a variety of technical and political actors, as well as a large number of technical and administrative processes. In view of these factors, it is recommended that an institutional organization be created to perform evaluations and use any information that may be obtained from them. In this regard we suggest considering institutions and organizations that play a part in defining, assigning, and reporting on the various functions and responsibilities; and in designing the stages, procedures, and annual timetable of the operation process, including the phases of analysis of information for integration into the decision-making processes, as well as for preparation and monitoring of commitments.

Marcela Guzmán S.

*International Consultant, former
Chief of Budgeting Division Office, Chile*

Session 1 Notes

- ¹ The institutional conditions are the rules, functions, skills, actors, and systems that structure administrative and government action. The practices of authorities and public agents are the behaviors with regard to institutional conditions.
- ² For more information on the variables and indices, see www.transparenciacolombia.org.co.
- ³ See the Web site, www.dipres.cl/control.
- ⁴ The measurement unit of the indicator is expressed as a percentage, according to the following formula: No. of students attending the program/No. of students eligible to receive the benefit * 100.
- ⁵ A program consists of a number of activities arranged into components or lines of work to supply goods or services, whose purpose is a common objective.

- ⁶ Characteristics of the program's responsible organizational unit, mainly considering its functions, its location in the structure of the organization, its own structure of subunits in the central and subnational levels, and allocation of responsibilities to each, including staffing.
- ⁷ Legal information, strategic and programmatic definitions, institutional memories and balances, prior studies, official documents, financial data, indexes, data, and so on.
- ⁸ These may refer to programs or sets of programs or actions with a common unit or purpose.
- ⁹ This report is presented to Congress in certificate form. See Budget Office (2009a), 155-161.
- ¹⁰ This improvement has been helped by self-evaluations conducted periodically by the Budget Office, and by the Impact Assessment made by the World Bank in 2004–2005 (see World Bank 2005).
- ¹¹ Document available at the Web site, *www.dipres.cl*.
- ¹² Technical commissions are meetings of the Budget Office of the Ministry of Finance and ministries and institutions whose budgets form part of the Budget Law, and take place each year when the Budget Law bill is drafted. Each ministry explains the priorities and technical aspects on which the proposed budget is based.
- ¹³ The Budget Office conducts an evaluation of financial performance and management as an internal task. The budget director and the Sub-Department of Budgets and the Management Control Division analyze background information on financial management and the performance. This analysis serves as the basis for defining exploratory budgets and budgetary frameworks for the next stage of drafting the Budget Law Bill.

Session 2: Institutional Arrangements for Monitoring and Evaluation Systems: International Experiences

Monitoring and Evaluation in the Government of Canada: Lessons Learned from 30 Years of Monitoring and Evaluation Development

Introduction

This paper provides an overview of the Canadian government's monitoring and evaluation (M&E) model and points to how the concepts have evolved and key lessons learned from more than 30 years of public sector experience.

The evolution of M&E in Canada

In Canada, the concept of M&E recognizes that evaluation has a distinct identity from monitoring. In other parts of the world, M&E may ignore or minimize evaluation. Evaluation in public sector management precedes by many years the notion of "outcome monitoring." It goes back to 1969, with the initiation of formalized and centralized evaluation practices. That initial approach to centrally led evaluation was replaced in 1977 by the first government-wide evaluation policy that established the model on which the practice of evaluation still functions in the Canadian government.

The focus on evaluation was inspired by the notion of "letting the managers manage," that is, allowing deputy ministers of federal government departments to assume greater responsibility of their departments and programs, and be accountable for the performance of those programs and the prudent use of public funds. The model is based on a strong central management board that oversees and holds deputies accountable. One of the mechanisms to achieve this is performance evaluation.

Over the past 30-plus years, however, many changes have occurred with policy, practices, and standards of evaluation. The policy, for example, has been updated on three occasions: 1991, 2001, and 2009. These changes occurred for a variety of reasons. They included (1) a recognition that the need for evaluation or performance measurement has evolved or become clearer, (2) M&E practices have matured, or (3) the demands of the system through public sector reform put emphasis on new areas—for example, the widespread introduction of results-based management (RBM) in the 1990s and 2000s.

The 1990s saw an increased move to performance monitoring and high-level reporting in an attempt to make performance information more accessible and useful to parliamentarians and parliamentary committees. A formal requirement for each government department to submit an annual departmental performance report (DPR) to parliament placed new emphasis on program managers (and not simply evaluators) to be more accountable for measuring and reporting on the performance of their programs. Associated with this was a new emphasis on the measurement of "outputs" and "outcomes" and not simply "inputs" (resources expended) when reporting on "performance." As with the evaluation function, the Treasury Board Secretariat (TBS) has played an active role in providing guidance and oversight over performance reporting and the production of DPRs across the system.

Unlike some of the other member countries of the Organisation for Economic Co-operation and Development (OECD), which significantly reduced the role of evaluation during their reform of public sector management,

Canada held the view that the two tools were complementary, and not competing, within the performance measurement tool kit.

The 2000s have seen two attempts to revitalize the evaluation function, with updated policies in 2001 and in 2009. In line with the government's modern management agenda, as articulated in *Results for Canadians*, evaluation and performance monitoring (and internal audit) are seen as key tools to ensure a results focus, responsible spending, and greater transparency and accountability across government. Table 4 highlights key elements in the evolution of M&E in Canada.

An overview of the Canadian M&E system

Many uses and users of M&E information

The intent of M&E in the Canadian system is to provide results information that will serve a variety of needs and users at different levels throughout the system. At an operational, departmental level, M&E would be expected to serve as learning tools to assist in program improvements and sound management practices. At the level of an individual department, with the deputy head accountable for good governance and performance of the department, M&E represents key management and accountability tools

available to the deputy head, and provides important inputs to strategic reviews that may be needed to make management decisions regarding program priorities and possible changes.

At a governmentwide level, the TBS is an important player in the production of M&E information and in its use in informing funding decisions about certain programs (grants and contribution programs subject to funding renewal) and governmentwide expenditure management.

Finally, in a legislative context, M&E information about government programs and operations is reported directly to parliament on an annual basis through DPRs and an overview performance report, *Canada's Performance*. The intent is to enhance the transparency and accountability of government operations with parliamentarians and Canadian citizens. More detailed discussion about the performance of government programs or operations may occur in the parliamentary committees that have the authority to seek additional information or clarification directly from the relevant departments.

Roles and responsibilities in Canada's M&E system

Within the Canadian M&E system, the delivery and use of M&E information has two key focal points: the individu-

TABLE 4:

The Evolution of M&E in Canada

Key Features	Canadian Experience
Initiation of formalized evaluation practices	1969
Governmentwide Evaluation Policy introduced	1977
Evaluation Policy revised	1991, 2001, 2009
Drivers of ongoing performance monitoring and reporting (date introduced)	<ul style="list-style-type: none"> • Departmental Performance Reports, DPR (1995) • Results-based Management and Accountability Frameworks, RMAF(2000) • Management Accountability Framework, MAF (2003) • Management Resources and Results Structure, MRRS (2006)
Key drivers of change in emphasis on governance tools	<ul style="list-style-type: none"> • Public sector reform • Fiscal restraint • Government Management Agenda: <i>Results for Canadians</i>
Current emphasis on evaluation capacity	<ul style="list-style-type: none"> • Formalized role within each department • Central management (guidance and oversight) of the function (Centre of Excellence for Evaluation, CEE)
Current emphasis on governance tools	<ul style="list-style-type: none"> • Evaluation (program and policy) • Performance monitoring • Internal audit • Strategic reviews

Source: Author.

al government departments and the TBS. The TBS plays a strong and proactive role in the practice and performance monitoring of evaluation within departments. The evaluation policy center, the Centre of Excellence for Evaluation (CEE) resides within the TBS. This unit plays a variety of roles in support of evaluation capacity building and its use, and provides operational oversight regarding standards and quality of evaluation practices and products. With the 2009 policy, TBS can now direct the conduct of governmentwide evaluation studies. TBS also provides guidance, proactive support, and oversight for performance monitoring that takes place in departments. Finally, TBS links evaluation to inform program funding decisions and the broader expenditure management system.

For individual government departments and agencies, deputy heads are required to provide dedicated resources to establish an evaluation capacity appropriate to their organizations. Although the resourcing of the evaluation function is flexible, the TBS/CEE monitors this and ensures that each department establishes the required evaluation infrastructure—that is, a senior-level evaluation committee, annual and multiyear planning for evaluation, a departmental evaluation policy reflective of the government’s policy, and the mechanisms needed for follow-through on delivery of credible evaluation products.

Part of the evaluation infrastructure in any department is the evaluation unit, led by the head of evaluation. This position ensures that the evaluation function reflects the government’s policy requirements and the deputy head’s priorities in the department’s work. The position generally reports to the deputy head or, at least, to ensure the independence of the departmental evaluation function, has direct access to the most senior official in the department.

Deputy heads of departments are accountable for establishing a corporate performance framework, called the management resources and results structure (MRRS), which links all programs of the department to the expected outcomes. This is the basis for performance monitoring and reporting, and its development is monitored closely by TBS to ensure adherence to the government’s MRRS policy. Performance monitoring is an ongoing responsibility of individual program managers, although development of the underlying program logic model and identification of appropriate performance indicators often comes about with the assistance of the evaluation specialists. In theory,

ongoing performance monitoring will provide much of the needed data for the eventual evaluation of a program; in practice, however, this does not always happen.

The national audit office, the auditor general (AG) of Canada plays an important role in providing oversight of the effectiveness of evaluation and performance reporting—that is, by examining the use of the function across the system and the “quality” of performance reporting. As an independent auditor and reporting directly to parliament, the reports of the AG raise the profile of M&E and its use in the public sector. The reports generally receive public attention and provide constructive advice to individual departments as well as to the central agency. Historically, the AG has been a strong supporter of the central agency role as a champion and facilitator of departmental evaluation.

The other key group implicated by M&E is represented by elected officials. Parliamentarians may be exposed to M&E information in various ways:

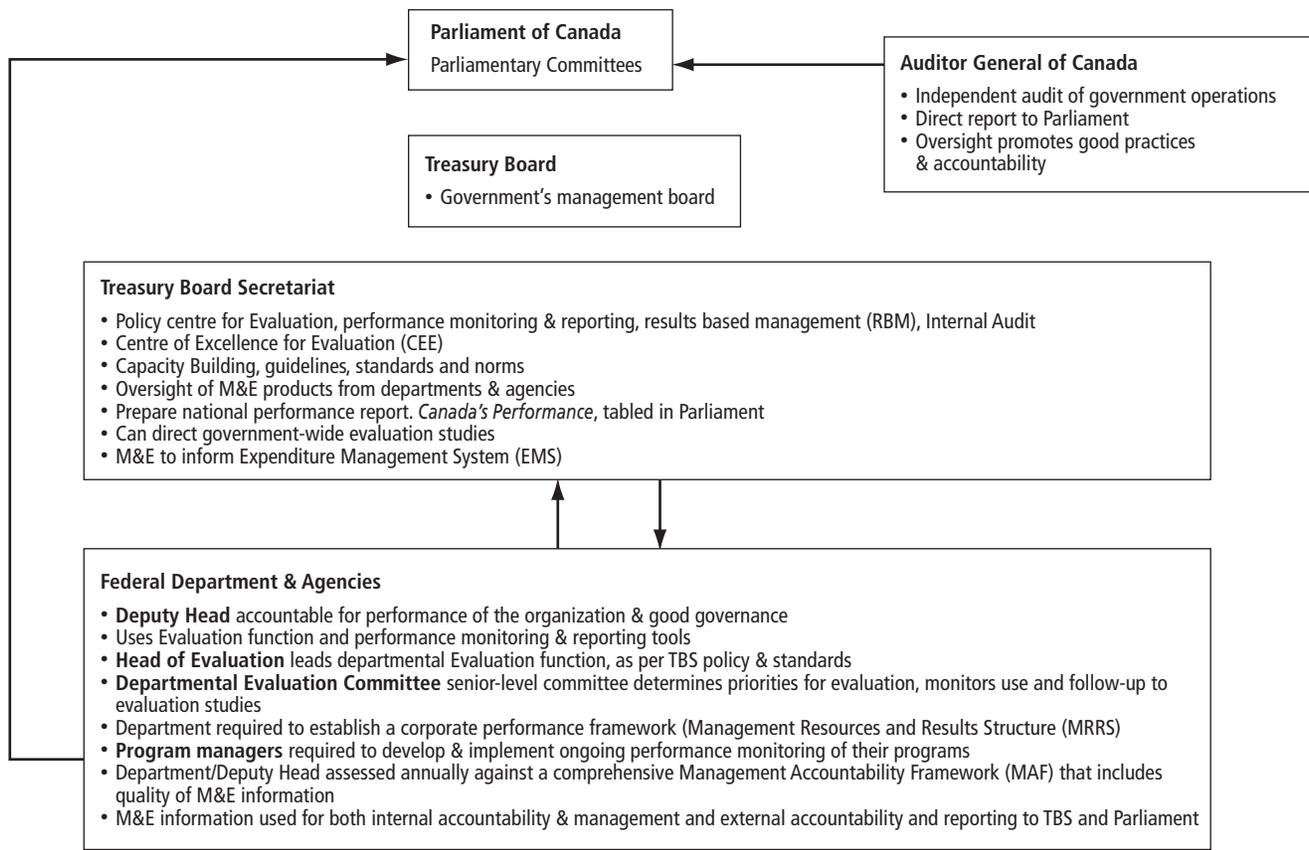
- through the annual tabling of departmental performance reports and the strategic-level report, *Canada’s Performance*;
- through more detailed discussions that may take place in the context of a parliamentary committee; or
- through discussion among TBS ministers on programs seeking funding renewal or broader-based expenditure review exercises.

Often, though, more exposure is given (through the media) to results emanating from internal and external audits of government operations. Clearly, though, the public sector reform agenda of the government (including fiscal restraint) has over the past 30-plus years had a major impact on the direction of evaluation and performance monitoring. Figure 3 provides an overview of the Canadian M&E system.

Some key features of the M&E system

The Canadian M&E system has invested heavily in both evaluation and performance monitoring as key tools to support accountability and results-based management. Performance monitoring provides managers with an ongoing assessment of how programs are performing, while

FIGURE 3:

An Overview of the Canadian M&E system

Source: Author.

evaluation provides a deeper understanding of performance and whether adjustments might be needed. Both advise deputy heads on how well their departments are functioning and inform parliamentarians on the performance of government operations.

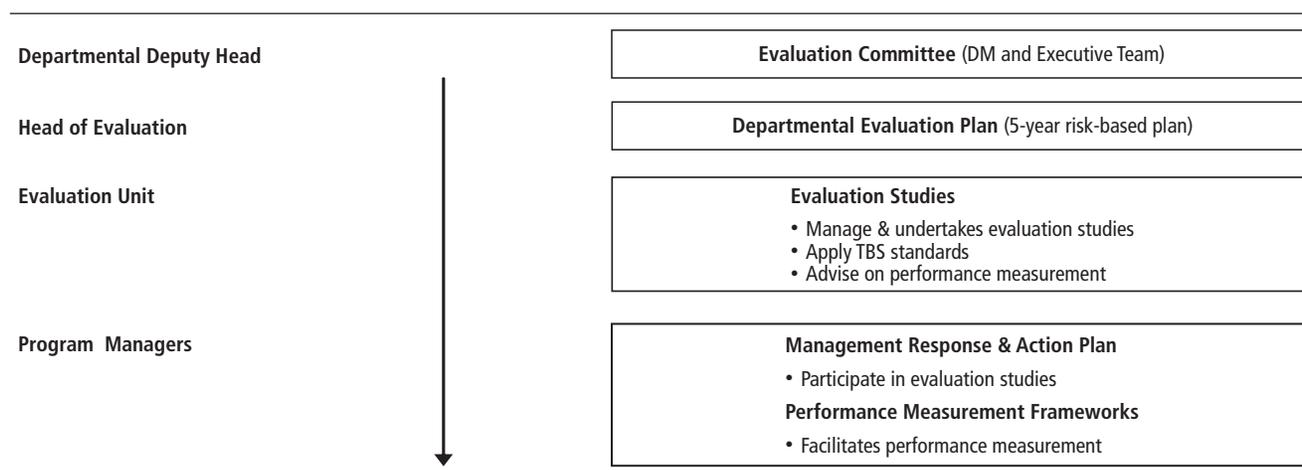
The responsibility for M&E in a government department falls on both technical staff evaluators and nontechnical officials (program and senior managers) in the department. The intent is that M&E becomes integrated into the decision making of the department and is a key driver in moving the organization toward a “results” culture.

The current state of the M&E system has evolved over time, as the central designers have recognized that the development and implementation of M&E is long term and iterative. This puts emphasis on the “process” of implementation as an important mechanism in itself in developing an “evaluation culture” or “results culture” in an organization and across the entire system.

The introduction of a number of centrally driven administrative policies that have been key drivers for M&E throughout the 1990s and 2000s generally have been phased in (that is, pilot tested in select departments) to allow for a period for learning and adjustment before their systemwide rollout.

The broad goal in investing in an M&E system has been to generate and use results information that supports the government’s management agenda from the perspective of learning and accountability in the design and delivery of government policies, programs, and services and the use of public funds. In this way, performance reporting generally aims to “tell the performance story,” rather than simply reporting on a limited set of indicators in the context of an accountability scorecard. Furthermore, evaluation informs decision making, rather than expecting that decisions will rest solely on the results of an evaluation study.

FIGURE 4:

Organization and Implementation of M&E in a Government Department

The model for evaluation in the Canadian system relies on a strong central presence in setting the policy, standards, and guidelines for the conduct of evaluation. Most evaluation studies are carried out by evaluation units that are internal to departments, although the new 2009 policy allows for centrally driven and governmentwide evaluations. To reinforce the independence of evaluators, so that they may “speak truth to power,” a number of elements in the model ensure objectivity and neutrality of the evaluation process and the reporting on results. Table 5 identifies some of the key features of the Canadian M&E system that

reinforce the credibility and quality of the practice of monitoring and evaluation.

Some lessons learned from 30 years of M&E development in Canada

Many lessons can be drawn from the 30-plus years of Canada’s experience with M&E. To help in drawing the broad implications, a set of lessons learned have been organized under three key headings: (1) drivers for M&E; (2) implementing the M&E system; and (3) key elements of capacity building.

TABLE 5:

Drivers for M&E

Building and using M&E capacity requires more than resources and technical skills—it requires a political will and sustained commitment—central leadership and a plan are important.

Think in terms of years, not months, to getting to a mature M&E system.

M&E information is not an end in itself; it needs to be linked to particular management and decision-making roles, particularly in the context of public sector reforms or government agendas.

Clarify the distinction between the “M” and the “E” and what each contributes to results-based management (RBM), and what each requires regarding capacity building.

To be effective, build a capacity to do evaluation (and gather performance information), and the capacity to use M&E information within organizations and across the system. A supply of good evaluations is not enough to maintain a robust evaluation function. There also should be a realistic demand for evaluation.

The capacity to use M&E information relies on the nature of the incentives in the system for managers to demand such information and actually use it as part of their normal operations. Incentives can take many forms, both sanctions for not complying, as well as rewards for meeting requirements.

It is important to manage the expectations for evaluation and not create unrealistic expectations about its role. While evaluation can and should inform decision making, it is generally one of many sources of relevant information. Questions about the performance of government programs generally do not have simple yes or no answers. Evaluations can and should provide useful information to inform debates about programs, their performance, and future direction.

Drivers for M&E

Developing M&E in an organization or a country requires a significant investment in time, resources (human and financial), and energy. To be successful, there ought to be a real need for M&E information. In Canada, much of this need has been driven by various public sector reform exercises and, most recently, the government's management agenda, *Results for Canadians*.

At times, the tendency is to adopt unrealistic expectations in considering what evaluation and performance monitoring can deliver on, as might be the case in satisfying, for example, a political need. While M&E should serve an important place in public sector management, the expectations about how and when M&E can be used and

what it can actually deliver on need to be tempered with reality, and generally remain modest. Table 6 provides some lessons relating to the drivers for M&E.

Implementing the M&E system

Implementation of M&E is long term and iterative, and not costless. As such, senior-level commitment and champions at both senior and operational levels are important elements to ensure sustainability through the long period of development and implementation. Over time, as M&E fatigue appears, reminder participants why M&E is important. Eventually, the goal is to move M&E beyond the point of being a special project to one in which it is a normal part of doing business and the management practices of the organization.

TABLE 6:

Key Features That Reinforce Credibility and Quality of the M&E System

Element	Description
Where is the emphasis placed: on the M, on the E, or on both?	<ul style="list-style-type: none"> Balanced emphasis and active support for <i>both</i> ongoing monitoring and ad hoc evaluation. Program managers accountable for performance monitoring; evaluation specialists within departments accountable for more probing evaluation studies. 2009 policy gives TBS authority to conduct governmentwide evaluations. Deputy head of a department or agency accountable for high organizational performance and delivering on expected results—uses both the M and the E (as well as internal audit) to support this—some flexibility in the resourcing of tools to ensure that they are appropriate to the size and needs of the organization.
How formal is the basis for the M&E requirements?	<ul style="list-style-type: none"> M&E requirements are based on administrative policies (allows more flexibility than legislation to modify and improve the policy as experience is gained over time). Identifies roles and responsibilities for all the key players and expectations for the system. Current basis relies on several centrally administered policies and procedures: evaluation policy; annual requirement to report to parliament via the departmental performance report (DPR); management resources and results structure (MRRS) policy requiring each department to develop a corporate performance framework and reporting structure; policy on transfer payments (revised 2008); program and strategic-level performance frameworks (RMAFs); annual departmental assessment by TBS through the management accountability framework (MAF).
What capacity-building efforts exist?	<ul style="list-style-type: none"> Proactive M&E capacity-building efforts from the central agency. Expectation that deputy heads will resource M&E to the level appropriate to their department/agency—central agency provides both guidance as well as monitoring and oversight. For evaluation, creation of the Centre of Excellence for Evaluation (CEE) has provided a system-wide champion to lead various capacity-building efforts (formalized training and workshops; a network/forum for heads of evaluation; and so on). Professional association (Canadian Evaluation Society, CES) actively leads training and development and, recently, an initiative to identify competencies for evaluators. For performance monitoring and reporting, TBS has issued and updated guidelines over the years, initially working with pilot departments and phasing in the introduction of reporting requirements for departments—intent was to implement in pilots, learn, adjust, and reissue guidelines on a broader basis.
Who conducts evaluation studies?	<ul style="list-style-type: none"> Internal evaluation units within government departments and agencies. New evaluation policy (2009) recognizes resource constraints for smaller departments and agencies—some special considerations, although evaluation is still a requirement. Central agency (TBS) may dictate a department to undertake a specific evaluation or initiate a governmentwide evaluation—a new requirement in 2009. Conduct of individual studies may be through use of internal evaluators, private sector consultants, or a combination.

(continued on next page)

TABLE 6:

Key Features That Reinforce Credibility and Quality of the M&E System

Element	Description
What quality control and oversight mechanisms exist?	<ul style="list-style-type: none"> • The TBS CEE monitors the planning and conduct of evaluation in individual departments, including evaluation coverage and quality of evaluations. • TBS reviews annually the incoming performance reports (DPR) from each department and agency before their tabling in parliament • Through the MAF process, TBS assesses each department/deputy head annually against a number of criteria (including the use of results and performance information). • The Auditor General of Canada (AG) conducts periodic audits of how well the evaluation policy is being implemented across government, evaluates the state of results-based performance reporting, and issues public reports to parliament on the findings and recommendations. • Formal requirement (2009) to conduct an evaluation of the government's evaluation policy every five years.
How is the evaluator's independence of maintained?	<ul style="list-style-type: none"> • Evaluation function has "direct and unencumbered access" to the deputy head of the individual department or agency. • Neutrality of both the evaluation function and the evaluator ("impartiality in behavior and process") is stressed in the evaluation policy and by the government's central policy center (TBS/CEE). • Each department has a senior-level Departmental Evaluation Committee in place that plays a variety of roles regarding evaluation planning, conduct, and follow-up. • Departments may make use of external review panels (independent experts, for example) for the planning and conduct of individual evaluation studies • Ongoing monitoring of the evaluation function and individual studies by the TBS/CEE—an operational-level oversight. • Oversight role played by the AG, an independent body reporting directly to parliament; public disclosure of information that reinforces both "independence" of the function and "transparency."
How transparent is the information?	<ul style="list-style-type: none"> • Evaluation policy stresses the importance of "transparency." Formal requirement to post all evaluation reports on the department Web site. • Annual Departmental Performance Reports (DPR) are public documents. • Annual reports of the AG are public and generally receive wide attention in the media. • Access to Information legislation requires compliance with all external requests (general public, media, and so on) for public disclosure of government documents (some exceptions). • Over time, the media has made increasing use of this access to publicize certain findings of particular audit or evaluation studies.

Table 7 offers some lessons on implementing the M&E system from the Canadian experience.

Building M&E capacity

A critical element associated with the sustainability of an M&E system relates to the adequacy of human resources with the needed skill sets. Human resources capacity development has and continues to be an ongoing issue. Capacity building is aided by the existence of well-developed professional networks. Canada have two evaluation networks. The first is within government, where heads of evaluation in departments meet on a regular basis. The second is across the professional spectrum, with the Canadian Evaluation Society (CES). CES is an active professional association representing evaluators from the public, private, and nonprofit sectors, including academics. A third association, the Performance and Planning Exchange (PPX) brings evaluators and managers together on issues of re-

sults-based management (RBM) practices. These networks provide part of the enabling environment for professional development, but it is equally important that time and resources are set aside to allow for the necessary training and development.

For its part, the TBS CEE has a capacity-building component as part of its mandate. This has resulted in the development of tools, guidelines, and an evergreen Web site. Additionally, the CEE has sponsored capacity-building workshops, conducted local and national learning events, and spearheaded national training programs for new evaluators.

The CES recently completed an exercise of identifying competencies for evaluators, as a first step in addressing the issue of ensuring a supply of trained and credible professional evaluators. Table 8 identifies a number of lessons from the Canadian experience that relate to dealing with building M&E capacity. In considering training

needs, it is important to reflect on the training and orientation to M&E required for nontechnical officials (the users of M&E information). Additionally, building capacity

needs to address an often-ignored area—that is, data development and credible databases.

TABLE 7:

Implementing the M&E System

Sufficient communication (and a forum for information sharing) is needed across organizations about the role of M&E and how it can help management to link the demand for and supply of M&E information—that is, to ensure that what gets produced is what is needed and delivered in a timely way.

A formal policy document is a useful basis for clarifying roles, responsibilities, and accountabilities of key players—deputy heads, evaluation specialists, program managers, and central agency officials.

The Centre of Excellence for Evaluation (CEE), the central agency champion for the evaluation function in government, has played a key role in the M&E system. It is the policy center for evaluation; it provides guidance; it leads and promotes capacity development; and it provides oversight to ensure quality control.

The TBS has phased in a number of requirements under the general philosophy of “try, adapt, and learn.” This lengthens the period of building the M&E system, but allows for a period of learning and an ease of adjustment when needed without a major investment in reorienting the whole system. An acceptance that it is long-term and iterative manages expectations.

In establishing internal evaluation units in departments and agencies, allowance for some flexibility is important to account for the unique circumstances associated with each organization. Deputy heads are given some flexibility in implementing the government’s evaluation policy, although all are equally accountable for the performance of their organizations.

Oversight by the national audit office is important in giving broad and public exposure of how well the M&E system is being implemented and whether adjustments are needed.

TABLE 8:

Building M&E Capacity

Building an adequate supply of human resource capacity is critical for the sustainability of the M&E system and generally is an ongoing issue. It needs to be recognized that “growing” evaluators requires far more technically oriented M&E training and development than can usually be obtained with one or two workshops.

Both formal training and on-the-job experience are important in developing evaluators. Two key competencies for evaluators are cognitive capacity and communication skills.

Program and senior managers are important audiences for less technical training on M&E and RBM. They need to have enough understanding to trust and use M&E information. This type of broad training/orientation is critically important in building a results culture within organizations.

There are no quick fixes in building an M&E system—investment in training and systems development is long term. Various options for training and development opportunities include the public sector, the private sector, universities, professional associations, job assignment, and mentoring programs.

In introducing an M&E system, champions and advocates are needed to sustain the commitment needed over the long term. Identifying good practices and learning from others can help avoid the fatigue that typically accompanies any change process, as enthusiasm starts to wane over time.

Evaluation professionals possess the necessary skill set to play a key role in providing functional advice and guidance to departmental/agency managers about the design and development of appropriate results-based performance monitoring systems. While managers should be responsible for performance measurement and monitoring per se, a recognized role for evaluators should be to provide such assistance and oversight on results measurement and monitoring.

Ongoing performance monitoring and the conduct of ad hoc evaluation studies should be positioned to be complementary functions that support one another. Within organizations, this would mean that the measurement plans and strategies to put in place ongoing monitoring systems would be implemented so that sufficient performance information is available to effectively support the evaluation of programs. In so doing, organizations would make best use of the limited resources typically available for their measurement and accountability tools.

Data quality is critical for the credibility of an M&E system, but an area often not adequately dealt with. A data development strategy (at an organizational level or a national level) should be considered as part of the long-term strategy of building a credible M&E system. The national statistics office can be an important player in assisting data development.

Performance reporting is often required at different levels—that is, providing information about a particular program, a sector, a department, or an agency, or at a national level. It is important to recognize the various audiences for performance reporting and their different needs and to tailor the level and breadth of reporting appropriate to those different levels.

Source: Author.

Conclusion

“Getting it right” is the goal of every country implementing an M&E system. While each country has its unique circumstances, the lessons from any country can be instructive in helping identify best practices. In this regard, two broad conclusions emerge from the experience of Canada with M&E. The first is that the road to getting it right is long. The second is that the journey is important in itself, for the learning that is gained along the way.

Robert Lahey

President, REL Solutions Inc, Canada

The Monitoring and Evaluation System in Sri Lanka: Experiences, Challenges, and the Way Forward

Background

Governments are challenged to respond to the urgency of citizens’ needs and to be more accountable to them. All over the world, the public is putting governments under growing pressure to show that they are providing good value for money and results. In the past three to four decades, significant budgetary resources and efforts have been deployed to accelerate development, reduce poverty, ensure equality, and improve social living standards and quality of life of the people. With fast tracking of public investment programs, Sri Lanka has been able to maintain its average annual economic growth rate at 5 percent. However, the country experienced challenges in translating this economic growth momentum into poverty reduction. The level of poverty remained at 26 percent to 22 percent over the last two decades. This situation led to the need to strengthen the government’s planning, monitoring, and evaluation system to focus on delivering of outcomes and impacts beyond the traditional output focus. The concept of Managing for Development Results (MfDR), otherwise known as results-based management (RDM), has been adopted gradually since the mid-1990s. MfDR is about public sector management, and the concept strongly emphasizes that shared vision, clear goals, and measurement of results would lead to a better future. MfDR is a change management process that emphasizes a shift in focus from inputs, activities, and outputs to out-

comes and impacts. It promotes the concept of “accountability for results.”

The Sri Lankan government fully recognizes the growing international consensus that MfDR is an essential aspect of good governance to improve development efficiency and effectiveness, transparency, accountability, and informed decision making. In the recent past, monitoring and evaluation (M&E) expanded globally and diversified in many contexts with many uses, such as decision making, organizational learning, knowledge base, program improvement, policy development, impact and outcome assessment, improved service delivery, accountability, performance audit, empowerment, and even transformation. Ambitious government systems with multiple stakeholders needs tend to achieve most of these desired uses. A good M&E system should go beyond institutional boundaries to cover national, sectoral, program, and project levels to ensure results orientation in government.

Furthermore, Sri Lanka, as a signatory to the March 2005 Paris Declaration on Aid Effectiveness, is committed to institutionalizing MfDR throughout the government. The MfDR process typically involves several phases such as articulating and agreeing on objectives, selecting Key Performance Indicators (KPIs), setting targets, monitoring performance, analyzing and reporting on results against targets, and facilitating management to take timely corrective actions. Different countries will approach MfDR differently, as it is not a one-size-fits-all model. Ultimately, it should lead to sustainable improvement in the country’s development outcomes. The government of Sri Lanka is committed to promoting the application of MfDR principles at various levels, namely, the national, sectoral, agency, and project level and at the planning, implementation, and postimplementation stages. At the planning stage, the results-oriented country-owned National Development Plan and sectoral plans are being aligned with the Medium-Term Expenditure Framework (MTEF), and the use of performance budgeting instead of line-item budgeting has increased. Line ministries are required to justify their budgets with well-defined output and outcome indicators. Today, what counts is not so much how many clinics have been built, but whether citizens’ health has improved; not how many schools have been constructed, but how many girls and boys are better educated.

Sectoral plans use outcome-based KPIs in setting targets and directions. The Department of National Budget and the Treasury have revised their budget circulars to focus on “results” to institutionalize performance budgeting systems. The government’s three-year MTEF incorporates outcome-based key performance indicators to justify public expenditures, and this helps achieve results.

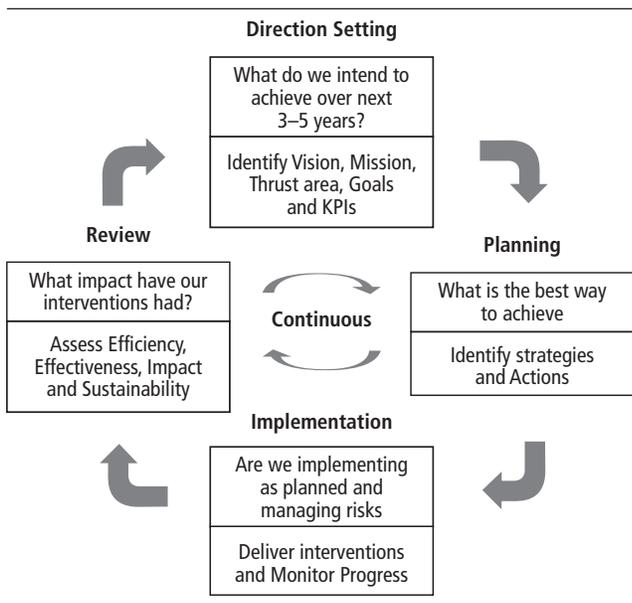
MfDR in Sri Lanka: A whole-of-government approach

The institutionalization of MfDR is seen as a major shift in focus in the M&E arrangements in Sri Lanka. MfDR includes a management cycle of setting directions, planning, implementation, and delivering and reviewing the results, which then feed back into the cycle to improve planning and to making ongoing improvement. Figure 5 represents the localized version of the MfDR process in Sri Lanka.

In the Sri Lankan context, the organizations that are managing for results must

- Have a clear vision of what they want to achieve
- Plan their work, keeping in mind the clear vision and mission

FIGURE 5:
MfDR in Sri Lanka



Source: Author.
Note: KPIs = Key Performance Indicators; MfDR = Managing for Development Results.

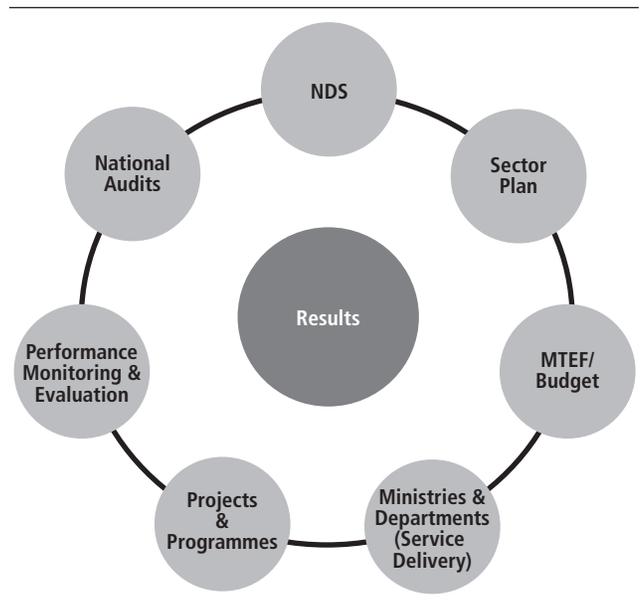
- Deliver what they planned in a manner consistent with public service ethics, values, and standards while meeting standards such as timeliness, quality, quantity, and within cost
- Track their progress by monitoring, measuring, and evaluating, learn from success and failures, and make continuous improvements

The MfDR approach adopted in Sri Lanka is more of a whole-of-government approach (see Figure 6) covering the national, sectoral, institutional, and project level. A results focus is being built into the National Development Strategy (NDS), sectoral plans, ministries, and projects.

Enabling environment

In Sri Lanka, the M&E of projects, programs, and development initiatives of the government is not a new phenomenon. Efforts to improve planning and project implementation have been a feature of development efforts since the early 1970s. One of the special features in Sri Lanka is the functioning of a separate Ministry of Plan Implementation (MPI) charged “as a National Focal Point for Monitoring and Evaluation of all government development projects and programs to ensure achievement of results and development effectiveness.” The MPI is headed by a senior cab-

FIGURE 6:
Whole-of-Government Approach



Source: Author.
Note: NDS = National Development Strategy.

inet minister and is mandated with the responsibility of monitoring and evaluating all government policies, programs, projects, and institutions. The Department of Foreign Aid and Budget Monitoring (DFABM) is the functional arm of the MPI to provide technical support and undertake the M&E functions with skills and technical know-how. This dedicated institutional arrangement places M&E more strongly in the institutional setting of the government of Sri Lanka.

One of the key features in Sri Lanka is the top-level commitment for M&E with a focus on MfDR. In the 1990s, with the technical support of the Asian Development Bank (ADB), the postevaluation system was strengthened in the MPI. As a result, the MPI conducted a number of post evaluations of projects and programs. In late 1990s, the United Nations Development Programme (UNDP) provided significant technical support way to strengthen the results-based M&E system in Sri Lanka. This enabled the government officials at the national and subnational level to understand and recognize the importance of results-focused monitoring. Also many positive factors contributed to a positive enabling environment, such as (1) political will, (2) overarching policy, (3) coordination to collect information, (4) flow of information from line ministries and projects to the MPI/DFABM, (5) strengthening of the electronic information management system in the National Operations Room (NOR) of the MPI, and (6) demand for information for decision making. However, capacity in government agencies, a large number of ministries, and the resultant coordination issues are some of the challenges that need to be addressed.

The MPI has introduced MfDR in government to track development results of the line ministries and their programs. Logical Framework Analysis (LFA) and Results Frameworks (RF) increasingly are used in planning and M&E arrangements within government. All these developments clearly indicate that the concept of “Managing for Results” set out in the Paris Declaration is being institutionalized and gradually moving forward in Sri Lanka.

The MPI and its DFABM championed the institutionalization of MfDR within government with the belief that the process has potential to ensure results through a shift in focus of public sector management from inputs and activities to outcomes and results. Many initiatives—such as the institutionalization of postevaluations, process evaluations,

and impact evaluations, the mainstreaming of the project performance management system, results-based management (RBM), performance budgeting and performance auditing, localization of the Millennium Development Goals (MDGs), and results reporting through country-based M&E systems—have been moving forward on a gradual basis since early 1990s. A strong link is evident among the MPI, the Ministry of Finance and Planning, and the President’s Office: they work closely with each other on M&E strategy. The line ministries, agencies, project management units, provincial councils, and district/divisional secretariats work closely with the MPI and provide progress reports in specified formats on a periodic basis. His excellency the president chairs the progress review meetings, which are held on a regular basis to assess progress of development projects, programs, sectors, ministries, and institutions. This high-level forum is a guiding and troubleshooting forum with top-level political commitment.

The Cabinet of Ministers made a formal decision and directed the MPI to monitor all development projects over Sri Lankan rupees (SL Rs) 50 million executed by the line ministries and submit quarterly progress reports to the Cabinet of Ministers. The MPI has developed an online, real time Web-based electronic monitoring and reporting system to collect, analyze, and report progress of all development projects over SL Rs 50 million. Progress review meetings are held by the MPI with project management units and officials of line ministries on projects that are behind schedule to address issues and ensure efficient and effective implementation. The MfDR principles increasingly are being applied in the national M&E systems and related areas. Some of the major achievements and lessons are discussed below.

Governmentwide performance measurement system and scorecards

The Sri Lankan government has introduced MfDR with the technical support of the UNDP to track performance of ministries and institutions using output and outcome indicators. This new management style holds ministries and departments and public officials accountable for results rather than efforts. At a macro perspective, ministries and departments receive budget appropriations from parliament to carry out specific mandates. These mandates have to be translated into detailed management expectations.

The results framework of the ministry or department sets out the breath, depth, and meaning of management expectations. By measuring performance against management expectations, set out in the results framework, the institution will be in a better position to objectively assess the results achievements. MfDR translates the vision and mission of public sector agencies into the practicalities of managing the organization better at every level. MfDR translates the vision and mission of public sector agencies into the practicalities of managing the organization better at every level: it defines expected results, monitors and evaluates progress toward the achievement of expected results, integrates lessons into management decisions, and reports on performance. MfDR when in operation ensures line-of-sight, which implies everyone in the organization understands the strategic vision and mission irrespective of their level and position in the organization.

A comprehensive performance measurement system was piloted in 2006–07 with four key line ministries (Education, Health, Agriculture, Highways). The selection of pilots was initially based on the interest shown in the reforms and the commitment of the leadership of the ministries to the change management initiatives. A range of activities such as awareness programs, advocacy and sensitization to policy makers and training programs were conducted with

the technical support of the UNDP and theADM. Currently, MfDR has been expanded and is operational in 35 line ministries. An MfDR core group has been established to drive this initiative with the leadership of MPI secretary. The experiences of the U.S. RBM models operated in Oregon (Oregon Benchmarks), Minnesota (Minnesota Milestones), and Virginia, as well as the Canadian models were taken into consideration in developing a localized model. The performance tracking system in Sri Lanka will be expanded to all ministries.

The MPI has established a Web-based MfDR platform in its information arm, the NOR, to maintain the agency's results framework and scorecards, or report cards, for each line ministry. The agency results framework (ARF) and scorecard are the centerpieces of the government's MfDR (see Figure 7). The ARF sets out the ministry's mission and its core business (thrust areas, goals, key performance indicators) with baselines and medium-term targets. The customized scorecards and report cards will appear on the computer screen to enable the assessment of the line ministries' performance. The scorecard employs a red-yellow-green grading system to track performance of government institutions. This "dashboard" is an early warning signal to alert when the achievement of targeted outcomes expressed by the KPIs is slipping.

FIGURE 7:

Snapshot of the Agency Results Framework

Thrust area 1 – Current and preventive health services
Strategic Objective – Provision of comprehensive Health service delivery and health actions

No.	Goal	Key Performance Indicators	Base line 2006	Targets			
				2007	2008	2009	2010
1	Reduction of infant mortality	Infant mortality rate	11.2/1000LB (2002)	11	10.8	10.6	10.4
2	Reduction of under five mortality	Under five mortality rate	4.4/1000 Under five population 1997 (AHB 2003)	4	3.8	3.6	3.4
3	Reduction of maternal mortality	Maternal mortality rate	38/1000000	36	35	33	31
4	Reduction of prevalence of underweight (malnutrition) children under five years	% of underweight children (W/Age) under five years	29.4% (2000) DHS	25%	24.5%	24%	23%
5	Reduction of incidence of low birth weight babies	% low birth weight babies	17.6 (20005) AHB	16.5	16	15.5	15

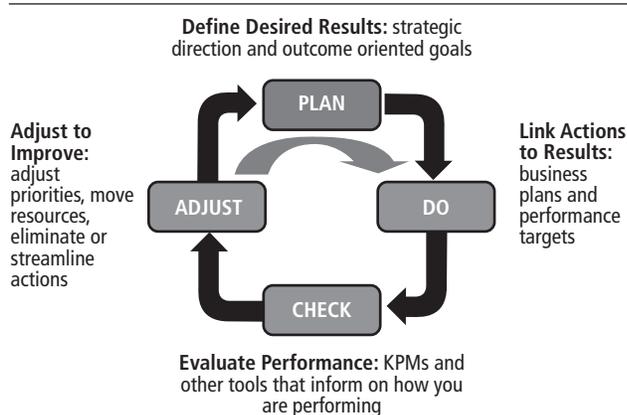
Source: Author.

A snapshot of the Ministry of Health's ARF is presented in Figure 7. Under the thrust area of curative and preventive health services, the goals and KPIs are infant mortality rate, under-five mortality rate, maternal mortality rate, percentage of underweight children, and percentage of low-birth-weight babies. The medium-term KPI target in the ARF enables a performance assessment of the Ministry of Health in each relevant thrust area. The framework is not carved in stone. It will continue to evolve as ministries and departments strengthen their management practices. The MfDR is to be used to improve results reporting to the cabinet and to the parliament. The scorecard raises a "red flag" when something is wrong. The Sri Lanka MfDR model follows the North American performance measurement framework models and is further articulated through a plan-do-check-adjust cycle, as depicted in figure 8.

- The *plan* stage defines the desired results. Performance expectations capture key dimensions of performance of interest to important stakeholders. Agreement of expectations implies a shared vision of what is to be accomplished.
- At the *do* stage, the ministries' strategies and actions are linked to the desired results.
- At the *check* stage, KPIs are used to evaluate performance.
- At the *adjust* stage, resources are adjusted and streamlined with the strategies and action to achieve desired results.

FIGURE 8:

Agency Performance Measurement Framework



Source: Author.

As in many developing countries, the availability of trained M&E personnel is a key constraint in Sri Lanka. Also, incentives need to be in place to reward success. MfDR, when fully operational, will help the government strengthen performance accountability and improve a continuous learning culture.

Readiness assessment

With the support of the ADB, a readiness assessment on MfDR was conducted in key ministries and agencies to assess the capacity gaps. Based on the readiness assessment findings, a capacity development plan has been developed to prioritize capacity-building initiatives in the selected institutions. The capacity development plan helps to address capacity gaps at the institutional level to enable the implementation of MfDR. The capacity development plan includes initiatives such as strengthening operating environment, clarifying the organizational results, aligning results to systems and processes, and upgrading management information systems to improve data collection, analysis, and reporting.

Implementation of MfDR in government

The minister of plan implementation placed MfDR in the policy arena by initiating a Note to Cabinet on MfDR, in which he outlined the importance and the steps necessary to institutionalize MfDR on a government-wide scale. The responsibilities for the institutionalization process were assigned to the MPI. The technical backstopping of the DFABM was complemented by the advisory technical support of the UNDP and the ADB. The ministry took the following steps to institutionalize MfDR governmentwide.

Launching of the process

The launching of the process involved many key sequential initiatives, including the establishment of a core group to drive the initiative, endorsement of government policy on MfDR, formulation of a governmentwide strategy, formulation of a "change agent" concept to establish ownership and leadership in line ministries and agencies to take the initiative forward, and support by advocacy and capacity development to sustain the initiative. The approach adopted by MPI in launching the institutionalization of the MfDR is discussed below.

Champion the initiative

The MPI as an apex ministry responsible for M&E, championed the institutionalization of MfDR government-wide. A core group was headed by the secretary, MPI, and included representatives of the Ministry of Finance and Planning, Department of National Planning, Department of National Budget, the AG's Department, the President's Office, the Prime Minister's Office, Department of Census and Statistics, and selected secretaries of the line ministries. The core group was to provide guidance, direction, and leadership to the initiative to mainstream the MfDR in government and to function as a think tank to formulate strategy and an action plan. The concept of core group helped to establish governmentwide commitment and ownership of the process. The core group formulated the strategy and action plan to institutionalize MfDR with the technical inputs of the DFABM of the MPI.

MfDR as a government policy

The Cabinet of Ministers approved the MfDR initiative and empowered the MPI as a lead organization to lead this initiative governmentwide. The top-level political support, which comes through the endorsement by the Cabinet of Ministers, indicates the political will and policy commitment to take this initiative forward. The MPI initiated a planned program of advocacy and sensitization at policy level. Dr. Ray Rist, advisor to the World Bank, was invited to conduct an initial orientation and sensitization program, which helped to share international experience and raised the importance of RBM/MfDR at the policy level in government. At the policy level, a number of programs were conducted to sensitize the policy makers and senior government officials on the importance of the MfDR initiative.

MfDR strategy and action plan

The strategy provided phased approaches with a pilot phase involving four line ministries and the MPI. Based on the lessons of the pilot phase, the MfDR institutionalization was expanded to cover 35 line ministries. A comprehensive strategy and action plan was developed, which included capacity development, staff training, advocacy, strengthening information systems, methodology, and reporting arrangements.

Change agent concept: Leadership at all levels of government

The MPI provided the technical support to introduce the MfDR methodology and approach. Each ministry was encouraged to establish a steering committee of senior officers who were expected to function as "change agents" to take this change management initiative forward. The change agents functioned as drivers of change and helped to facilitate the MfDR initiative at the ministry level. The change agent concept clearly shows the importance of leadership at all levels of government to take this effort forward.

*Adoption of a process approach:**Consensus building*

In each case, the ARF of each ministry was developed by that ministry through a participatory process with the MPI's technical guidance. Each ministry assumed ownership and the process approach helped to create this ownership versus a product approach. Workshops were held in the line ministries by involving all key players, and the consultations gave a sound basis for ensuring wider ownership. They helped to tap the expertise, skills, and experience on one side and the trust, support, and network on the other side.

Buy-in

The process approach and the involvement of senior government officials and other relevant stakeholders created wider ownership, which contributed as a positive factor for the buy-in ensuring sustainability of the initiative.

The ministry-level MfDR formulation workshops involved active participation of the top-, middle-, and lower-level staff of the institutions, and this enabled the operationalization of the concept of line of sight. Moreover, other stakeholders also were consulted in this process.

The ARF of the line ministry will be closely integrated into the national budget to make it a performance budget, and thus the initiative has built in factors of sustainability. Moreover, the link with the AG to measure performance through performance audit arrangements further strengthens the MfDR initiative as a national process. Expected results are mutually defined and agreed on through a consensus-building process involving all major stakeholders. This enhances the stakeholder's sense of ownership.

Country-level community of practice to facilitate peer-to-peer dialogues

The core officials of the ministries to which MfDR was introduced formed into a community of practice and shared their experiences and expertise through a face-to-face and peer dialogue. This helped to establish connections among the “islands of best practices” and helped participants learn from successes and challenges. Moreover, a quarterly newsletter on “Results Focus” helped to complement and supplement the face-to-face community of practice dialogue.

Capacity building and readiness assessment

Capacity building is fundamental to institutionalizing MfDR in government. The readiness assessment tool developed by the ADB has been used to identify capacity gaps in line ministries and to formulate a capacity development plan. The UNDP and ADB support helped formulate a capacity development plan as a complementary activity to institutionalize MfDR. Awareness creation, staff capacity building, and on-the-job training on MfDR methodologies, tools, techniques, and practices were emphasized in the implementation of capacity development plan.

Strategy to action

The MfDR approach encourages ministries and departments to understand the results chain and establish logical links between planned inputs, expected activities/outputs, and envisaged outcomes based on the theory of change. The MPI took the following key steps to operationalize MfDR in line ministries.

Articulating and agreeing on objectives

Identify clear and measurable objectives (results) aided by the results chain-logic model. This basically involves preparation of a vision statement, mission statement, thrust areas, and goals. The government’s 10-year national development framework, sector plans, and localized MDGs and mandate of the ministries are considered to be the fundamental basis to articulate and agree on line ministry objectives.

- *Selecting indicators*

Identify output and outcome-based KPIs to measure progress toward each objective. KPIs for each expectation in the framework are meant to convey the breadth and meaning of the expectation.

- *Setting targets*

Set explicit medium-term targets to provide medium term-directions and to judge performance. This is the main aspect of the process for each indicator.

- *Monitoring performance*

Monitor performance to regularly collect data on actual results.

- *Analyzing and Reporting*

Review, analyze, and report actual results as against the targets.

- *Integrating evaluation*

Evaluate to provide in-depth analysis and to provide complementary performance information not readily available from performance monitoring system.

- *Performance reporting and feedback*

Use performance information for internal management accountability, learning and decision-making process and for performance reporting to stakeholders.

The first three steps generally relate to a results-oriented planning approach, sometimes referred to as “strategic planning.” The first five steps together are usually included in the concept of performance measurement. All steps combined are essential to an effective results-based management system.

The NOR is the information platform of the MPI to support evidence-based decisions. The MfDR Web-platform was established in the NOR to facilitate and track results and performance of line ministries.

In the past, the progressive monitoring of ministries focused heavily on financial progress (budget utilization) and physical progress (activity monitoring) with little emphasis on achievement of outcomes, results, and policy objectives. However, implementation issues and interagency coordination issues were addressed through the high-level monitoring meetings held by the president. The MPI submits, on a quarterly basis, to the Cabinet of Ministers, a progress report on the capital budget of all line ministries and development projects over SL Rs 50 million. In the recent past, the emphasis was more on synthesis-type reporting rather than on the traditional way of producing large volumes of reports that resulted in information overload. The

institutionalization initiative of MfDR is a major shift in focus of the government's M&E system.

The MfDR initiative is to be complemented with performance agreements to ensure ministries and institutions manage priorities by establishing strong links between agency plans and employee performance. While the MPI is introducing MfDR to ministries and departments, the Department of Public Enterprises of the Ministry of Finance and Planning has introduced MfDR as part of its corporate plan initiative to state-owned enterprises (SOEs) and statutory bodies. Advocacy and awareness-building efforts on MfDR are in progress. Spending money is no longer sufficient evidence for results. Ministries and agencies are asked to demonstrate how their programs and strategies are effective in producing the intended results. According to Osborne and Gaebler (1992), "Perfectly executed process is a waste of time and money if it fails to achieve the outcomes desired." Performance management redirects our efforts away from "busyness" towards "effectiveness." Being busy is not the same as producing results.

***Electronic project monitoring system:
A distinctive feature in Sri Lanka***

One of the noteworthy and significant aspects is that a homegrown, user-friendly, national, Web-based online electronic project monitoring system (ePMS) has been established in the DFABM to track the implementation progress (financial/physical) and results of all development projects and programs. National ePMS is a proactive, Web-based project monitoring system that tracks whether development projects and programs are implemented according to the agreed work plan in a results-oriented manner. The system provides access to project information on a donor, sector, and ministry basis.¹⁴

The system uses early warning (traffic lights) signals that enables "problem projects" to be separated from others. The system helps to identify bottlenecks, delays, issues, and constraints in the implementation of projects and any additional needs of the executing agency. It includes results monitoring using logical framework analysis, monitoring compliance of loan covenants, tracking cash flow, reimbursable foreign aid claims, procurement progress, and major issues with pictorial proof. Flash reports on problem projects help in troubleshooting exercise and are submitted

to the Cabinet of Ministers on a quarterly basis to fulfill the reporting requirements.

The ePMS captures feedback from beneficiaries and citizens. Notable features of the system are as follows:

- Keeping track of the development project progress electronically covering financial and physical progress.
- Harmonizing results reporting through results frameworks. National ePMS is a homegrown, country-owned, country-driven system that provides the public sector with paperless monitoring capability.
- Facilitating pictorial proof to demonstrate progress and results at ground level.
- Capturing data at the source, reducing transaction costs.
- Providing alerts and early warnings on shortfalls and gaps, including time and overruns; providing red alerts on problem projects and projects behind schedule.
- Advising the higher authorities of the problems, issues, and constraints faced in the implementation of development projects and facilitating troubleshooting of problem projects and projects behind schedule.

A recently completed donor evaluation mission rated the ePMS a success story in terms of its comprehensive coverage, periodical updating, and use of information for troubleshooting. However, the current low levels of use of the system by sector ministries indicate an unexploited opportunity. The Ministry of Finance and Planning is using the system as part of its efforts to strengthen project management capacities. The MPI uses the system for troubleshooting. Also, it is necessary to enhance "data capture at source." The system allows for better overall coordination and helps to address execution problems.

The national ePMS was based on the Malaysian National Operations Room Model used by Prime Minister Dr. Mahathir Mohamed to track implementation progress of all development programs. The Sri Lankan ePMS was a homegrown model with a user-friendly facility to track development project progress. It is not intended to be used as a "policing function" but as a system that iden-

tifies projects on schedule or behind schedule or problem projects. The system depicts these projects using a traffic light color code by sector and ministry classification of projects and facilitates troubleshooting. Currently, the national ePMS includes more than 120 large and mega projects.

Key elements

The national ePMS captures many key elements that are fundamental and essential to track the progress of development projects. The key elements are the project profile, financial progress, activity monitoring, compliant of loan covenants, procurement monitoring, and major issues.

Project profile

The profile provides the basic information of the project, its name, source of funding, implementing arrangements, objectives, purpose and outputs, location, cost, timelines, and officer responsible, including other basic information.

Financial progress

This element tracks the cumulative financial progress as well as the monthly financial progress against total funds available and the annual budgetary targets. It monitors actual disbursements against targeted disbursements on a projectwide basis.

Activity monitoring

The projects outputs are broken down into components, subcomponents, activities, and subactivities using work breakdown structures. All activities have planned timelines and targets and progress is monitored against such timelines.

Logical framework analysis

The results framework of the projects shows the “program theory” with a results chain and measurement system, including KPIs and risks. The system helps to track results and major risks.

Monitoring compliance of loan covenants

This tracks the compliance of loan covenants of all donor-funded projects.

Procurement monitoring

Project management mainly involves managing contracts. Hence, procurement management is an important aspect. Procurement monitoring keeps track of the procurement progress, which is a core area in all development projects.

Monitoring major issues

The system keeps track of the major issues affecting the implementation of projects and action taken on such issues. This facilitates troubleshooting. The nine knowledge management areas identified by the International Project Management Institute are (1) scope management, (2) time management, (3) cost management, (4) quality management, (5) human resources management, (6) procurement management, (7) risks management, (8) communication management, and (9) integration management.

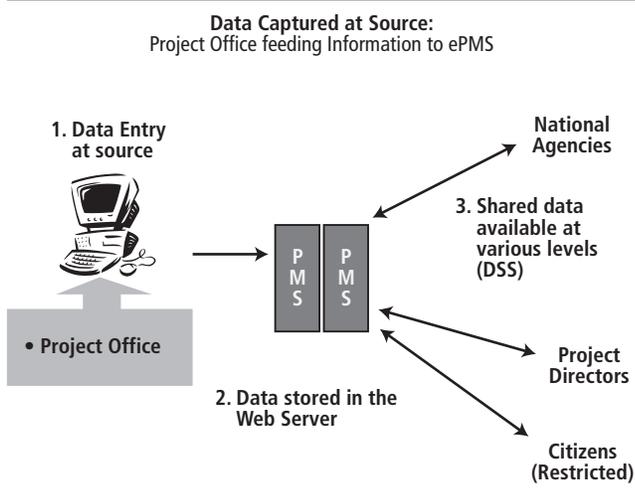
The current national ePMS of the MPI integrates all nine knowledge management areas in the system. To support the governance functions, the major functions of ePMS are as follows:

- Serve as an executive information center for central agencies such as the President’s Office, the MPI, the Ministry of Finance and Planning, and other line ministries.
- Provide information to project managers to manage projects effectively. Hence, the ePMS supports managerial functions.
- Ensures access to information to citizens and create transparency and accountability.

Figure 9 depicts the flow of capturing data from its source to its end use. The project information is captured at source from project office and stored in the Web servers and finally shared with potential users at various levels. Hence the system is a decision support system and executive information system.

The ePMS has many salient features. It helps to create online information any where, any time and enables information to be projected on to a large screen and used for discussions at meetings. The potential users of the system include apex agencies such as the President’s Office, the Presidential Secretariat, the MPI, and the Ministry of Finance and Planning, line ministries, project directors, and

FIGURE 9:

Data Capture at Source

Source: Author.

citizens (restricted access). The system's special report generating facilities include the following:

- Flash reports (higher level reports to senior policy makers on overall financial progress, physical progress, status of projects, major issues, and action taken in the implementation of projects)
- Summary reports (reports on aid utilization and budget progress, accessed by ministry, sector, donor, and project by various monitoring authorities)
- Tailor-made reports (using the crystal report facility)

Outcome-based national economic performance

Under the traditional national economic performance reporting system, high economic growth, low inflation, and unemployment were regarded as indicators for healthy economic climate and believed to result in prosperity for citizens. However, citizens are increasingly concerned about their quality of life measured in terms of quality of education, health care, safety from crimes, clean environment (safe drinking water and sanitation), and the like. Hence, traditional economic-based measurement systems failed to address such quality-of-life issues and outcomes. The Sri Lankan government has established a results reporting system for the localized MDG. The Department of Census and Statistics supports the collection of the data, which are widely used within the government. Moreover, the Central

Bank of Sri Lanka has developed the prosperity index to capture wider aspects of development.

Poverty monitoring and tracking of MDGs

The MPI also monitors the progress of national development goals. The MDGs are the best known and most ambitious global commitments to results-based development management. The agreed MDGs, targets, and indicators provide the basis for measuring progress. It collects performance indicators to measure achievement of localized MDGs and poverty with a focus on the "big picture," with the support of national statistics and surveys (such as Household Income and Expenditure Survey and Demographic Health Survey) from the Department of Census and Statistics. The government prepares an annual MDG country report to identify the gaps, poverty pockets, and regional disparities. This M&E information helps governments to allocate budgetary resources for needy areas.

National evaluation arrangements: ongoing, ex post, and impact evaluation

The DFABM undertakes ongoing, ex post, and impact evaluation of selected mega projects and disseminates evaluation findings to concerned stakeholders. The evaluation function of the MPI was strengthened through the technical assistance support of the ADB. The Organisation for Economic Co-operation and Development's Development Assistance Committee (OECD/DAC) evaluation criteria, such as the strategy's relevance, efficiency of implementation, effectiveness, impact, and sustainability, are given due consideration in undertaking the evaluation. The OECD/DAC evaluation quality standards are considered to maintain the quality of evaluations. On invitation by donors, the DFABM participates in joint evaluations with donors to create national ownership and build local capacity in evaluation. The department also undertakes diagnostic rapid assessment through field visits of problem projects and submits flash reports to the ministry's secretary to troubleshoot projects that are behind schedule. Given the human resource constraints in the ministry, it was decided to outsource the evaluation of priority mega projects. The ministry, in consultation with the relevant line ministry, decides areas to be covered in the terms of reference for evaluation. Also, the ministry identifies the information needs of the

president, the cabinet, and other stakeholders in designing the terms of reference. The findings of such evaluations are disseminated to the line ministries and project offices for necessary follow-up action.

A key problem in many countries has been the inability to access the evaluation information available on various development projects and program either completed or currently being implemented. Such evaluation lessons and findings are important and useful to improve the quality of new projects and programs, especially because it helps to avoid past mistakes and build on best practices in the formulation and designs of new projects.

Evaluation information system

Having recognized the importance of a systematic use of evaluation and feedback arrangements, the MPI has taken action to establish a Web-based evaluation information system (EIS) to ensure effective dissemination of evaluation findings, lessons learned, and synthesis of such findings. Such evaluation information will provide sectorwide synthesis to ensure more effective feedback and assist in integrating evaluation findings into the planning, budgeting, and policy-making process. Also public availability of the evaluation reports through the EIS is expected to improve public accountability and transparency.

Evaluation asks what works, what does not work, why, and in what context. The responses are important for planning and programming and contribute to development effectiveness. The EIS enables development practitioners to access evaluation information anywhere, anytime—thus empowering them to make evidence-based decisions in development. The establishment of an EIS is considered a critical milestone in the MfDR initiative as it is expected not only to improve aid effectiveness but also to promote a learning culture. The support toward the development of EIS was extended by the UNDP.

Evaluation of the implementation of the Paris Declaration on aid effectiveness

The MPI and the DFABM undertook an evaluation of the implementation of the Paris Declaration on Aid Effectiveness. The government of Sri Lanka strongly believes that the five principles of the Paris Declaration, such as national ownership, alignment, harmonization, managing for development results, and mutual accountability, are fundamen-

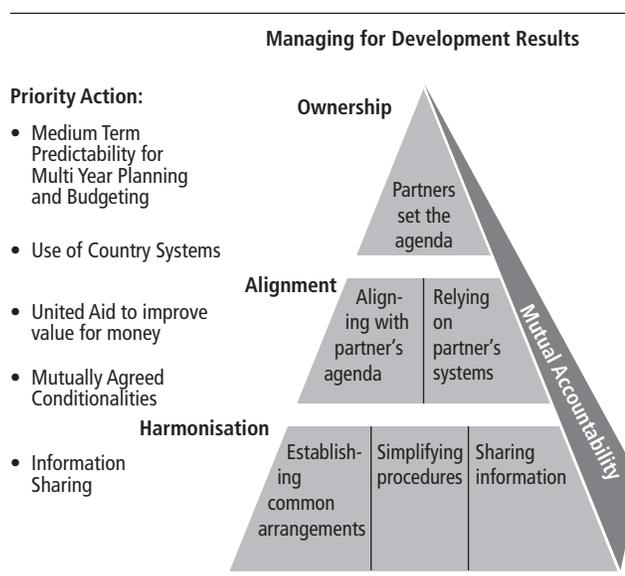
tal to improving “aid effectiveness” and development effectiveness (see Figure 10).

Hence, with UNDP’s support, the MPI undertook an independent evaluation to assess the implementation progress and results of the Paris Declaration commitments. An evaluation reference group was formed to advise, and a management group coordinated the evaluation. This evaluation identified the impediments to implementation, and helped the Sri Lankan government take suitable action to ensure effective implementation of the declaration’s principles. Moreover, the evaluation findings have fed into the global evaluation process, enabling the OECD/DAC Development Evaluation Network to undertake a synthesis of the declaration’s evaluation in eight countries and 11 donors. These findings were useful in the formulation of Accra Agenda for Action at the High Level Forum on Aid Effectiveness held in September 2008 in Ghana.

The Accra High Level Forum has endorsed the need for aid predictability, use of country systems by donors, untying of aid, and mutually agreed conditionalities instead of imposed conditionalities as fundamental action required to accelerate aid effectiveness. This implies that aid effectiveness requires radical change in behaviors of donors and development partners. Implementation of the Paris Declaration Principles and the Accra Agenda for Action is

FIGURE 10:

MfDR and the Paris Declaration



Source: Author.

fundamental to the change management process and improved development effectiveness.

Sri Lanka Evaluation Association

The Sri Lanka Evaluation Association (SLEVA), as a civil society, plays a catalyst role in advocacy, awareness creation, training, and developing standards, ethics, methodologies, and best practices to improve evaluation culture. Its membership includes academia, researchers, private sector consultants, government officials, and nongovernmental organization (NGO) representatives who are interested in evaluation. SLEVA works closely with the MPI in building M&E capacity and culture in such areas as evaluation training, sharing evaluation, best practices, organization of evaluation forums to discuss evaluation topics, international and national conferences on evaluation, and promoting community of evaluation practitioners.

Strategies to strengthen the national evaluation system

The following strategies have been identified to strengthen the MPI's National M&E system. Currently, the UNDP technical assistance project has provided assistance to support these strategies.

Policy commitment and support

- Advocate and sensitize at political and policy level on the importance of results-based M&E and ensure its acceptance and placement of M&E in key decision-making centers of the government to create local demand for M&E.
- Ensure that M&E institutions are linked to the planning, budgeting, resource allocation, and policy functions of the government and MfDR concepts are integrated in all areas of the development cycle.

Legal and budgetary support

- Develop a legal foundation to make M&E and MfDR mandatory. Use law, decree, cabinet decision, or other high level pronouncement to legitimize MfDR concepts and results-based M&E systems.
- Provide sufficient financial allocation for MfDR and M&E strengthening in the line ministries.

- Ensure an appropriate balance between “monitoring” and “evaluation.” Preferably separate evaluation from monitoring to ensure balanced resource allocation for evaluation.

Sound institutional arrangement

- Strengthen institutional arrangements to place M&E and MfDR in a strategic context.
- Establish links between evaluation exercise and performance audit exercise by encouraging partnerships between evaluation institutions and performance audit institutions (that is, the AG's Department) with regard to accountability-oriented evaluations.

Standards, ethics, and guidelines (quality of evaluations)

- Develop evaluation standards, guidelines, and ethics to ensure quality evaluations. Ensure scoping sessions are conducted to clarify the evaluative questions and to ensure that the needs of the potential users are taken into consideration and timing of evaluation is appropriate.
- Encourage national evaluation association to actively promote evaluation culture and MfDR concepts.
- Develop standards and criteria for good evaluation in collaboration with civil society such as SLEVA and undertake meta-evaluations to ensure quality evaluations.

Strengthen evaluation guidelines and systems

- Strengthen localized guidelines for systematic evaluations.

Strengthen methodologies and practices

- Make evaluation a process within the development policy and project cycle. Expand evaluation to cover projects, programs, sectors, policies, and institutions. Encourage synthesis of project evaluations to provide sectorwide learning. Promote cost-effective rapid assessment methods under time, budget, and resource constraints. Also, consistent, localized evaluation methodology and terminology are essential.

- Reexamine the approaches and tools for evaluating the multiple dimensions of development. Encourage the use of diverse or multiple methods, as well as participatory methods for lessons learning (utilization) oriented evaluations.
- Encourage more joint evaluations instead of donor-driven evaluations.

Evaluation capacity development

- Strengthen the professional evaluation capacity within the government through continuous staff training.
- Promote in-country evaluation faculty development programs in Sri Lanka Institute of Development Administration, the government arm for training, and in other universities at the graduate and postgraduate level.
- Strengthen the documentation center on evaluations, and promote the exchange of experiences and access to best practices and sharing of databases.

Strengthen the feedback arrangements

- Improve disseminations of evaluation reports through in-house workshops and seminars, customized reports, evaluation summary reports, press briefings, and post EIS.
- Establish strong feedback arrangements among evaluation, planning, decisionmaking, policy formulation, project appraisal, program management, and budgeting and resource allocation functions.
- Ensure action is taken on the recommendations in evaluation reports. Wider dissemination of evaluation information preferably should include the Parliamentary Public Accounts Committee, Parliament library and media. User-friendly evaluation synthesis or summary reports should be widely circulated.
- Stimulate the evaluation issues in the country's development dialogue and sector program assistance. M&E units must have active involvement in the planning of new programs.
- Incorporate evaluation lessons into the new project concept documents or project submission formats so that past mistakes are not repeated. Revise project submission formats to incorporate evaluations lessons of past projects.

Success factors

The ultimate success of evaluation depends on how well the planners and decisions makers utilize the valuable M&E findings and lessons to improve future programs, projects, policies, and institutions.

Sri Lanka's MfDR and M&E systems and practices have been internationally recognized as best practice approaches worthy of scaling up. Senior government officials from Afghanistan, India, the Republic of Yemen, and Uganda, as well as representatives from the Asia Pacific Community of Practice on Managing for Development Results and International Program for Development Evaluation Training (sponsored by the World Bank in partnership with Carleton University of Canada) visited Sri Lanka to study the emerging good practices of MfDR and M&E initiatives. The *OECD Sourcebook* (2008) on emerging good practice in MfDR highlights the Sri Lankan case. The Asia Pacific Community of Practice on MfDR has identified the following factors for successful institutionalization of MfDR in Sri Lanka.

- Strong support from the top and strong and sustained leadership for results oriented reforms and buy-in. Government policy on MfDR endorsed by Cabinet of Ministers and champions at national level and change agents at different levels of government to drive the initiative have been instrumental in taking this change process forward.
- The MfDR approach adopted in Sri Lanka is a whole-of-government approach covering national, sectoral, institutional, and project level. It is not just viewed as a technical tool but as a comprehensive way of thinking to achieve outcomes and impacts.
- Cascading approach, which combines a mix of top-down and bottom-up approaches and sequenced approach beginning with pilots and then expanded and mainstreamed in government, is a more pragmatic strategy.
- The MfDR reforms as complementing the existing initiatives, strategies, and general reform agenda and is part of the country systems. But "perfect" should not become the enemy of the "good" in the MfDR area, as most developed countries are still struggling to achieve a satisfactory solution.

Issues and challenges

In many countries, including Sri Lanka, the wider dissemination of M&E findings remains a problem. M&E institutions and the planning institutions seem to function in isolation and do not have an effective formalized feedback arrangement to integrate lessons into the planning and design of new projects. These institutional gaps defeat the very purpose of M&E. Therefore, it is necessary to establish strong links between M&E and policy formulation, reforms, planning, budgeting, and resource allocation functions. The Sri Lankan government has identified the need to establish strong feedback mechanism. Feedback is the weakest link in the project cycle. This issue is now being seriously addressed.

In many cases, donors and partner countries continue to be disbursement oriented. Moreover, donors tend to be using their own donor systems rather than country systems to maintain visibility and attribution. The lack of demand for MfDR, shortage of professionals, multiple results frameworks, too many indicators, lack of aid predictability, and weak statistical capacity have been identified as constraints in many developing countries. Also, MfDR, as a change management process, takes time.

Recognizing the demand side of the equation for creating local demand for evaluation with a utilization focus, the supply side of the equation, which includes skills, procedures, methodology, data systems, and manuals, also has to be addressed. The need to focus on national evaluation capacity development is equally important. Making M&E information available does not necessarily mean effective utilization. It is hard to justify the existence of an M&E system that is not effectively utilized. The Sri Lankan government was able to address some of these issues with the technical assistance support of UNDP and the ADB.

Attribution is a challenge for measuring performance in ministries and departments. In selecting the KPIs, care should be given to the attribution issues. Unlike the private sector, the public sector does not have a bottom line. Multiple stakeholders and divergent interests characterize many areas of government activities. Target setting is a challenging process. It is difficult to get consensus on goals and values and difficult to agree on KPIs.

Specifying and agreeing on expected results is not easy. The results chain is not always logical as expected. Indicators are missing for some results areas. Targets and base-

lines are not given. Setting achievable targets is not possible in the absence of baseline. The greatest problem associated with performance management is unrealistic expectations.

In addition, information overload can be a problem. Policy makers were unable to absorb the vast amount of information provided to them. Outputs are manageable for institutions, but outcomes are the results of collaborative efforts of other institutions. Agencies can more easily be held accountable to output targets but not necessarily for outcome targets.

In some countries, MfDR systems have not fully achieved the goals. Performance levels slipped partly due to a lack of resources and unrealistic expectations. Also, weak links between “agency performance” and “individual performance” are a concern. To ensure successful operation of RBM systems, incentives are needed for achievements as well as some form of penalty for slippage.

Because MfDR systems often create information overload, it is also necessary to understand who needs what information, for what purpose, and when. It should not be overly supply-driven information. Moreover, in some cases, MfDR creates fear of being held accountable for performance when cooperation and assistance from outside the organization are necessary for success.

It is necessary to look at the balance between learning and accountability. Although independent evaluation is important for ensuring objectivity, too much emphasis on accountability-focused donor-driven independent evaluation can be a potential constraint for lessons learning and feedback. Hence, the importance of lessons learning and ownership has been recognized in the government evaluation system.

Many countries have not adequately responded to the observation that ex post evaluations are more often viewed as postmortem exercises that do not contribute much to strategic decision making. It is necessary to recognize the importance of lessons learning and performance accountability. Increasingly, concurrent evaluations are encouraged for midcourse corrections. Also, the government encourages donors to undertake more joint evaluations to ensure national ownership, lessons learning, and capacity building.

There has been a generally tend to monitor rather than to evaluate. It is necessary to give equal importance to evaluations by finding the appropriate balance between the two activities. Despite the enormous methodological and technical challenges, institutionalization of evaluation is the

way to ensure results orientation in development work. Moreover, development policy and aid tend to shift from projects and programs to a sectorwide approach and, as such, an M&E approach needs to cover policies, sectors, and thematic areas on a countrywide basis. Policy evaluations, sector evaluations, and thematic evaluations are becoming equally important, and the Sri Lankan government has given much emphasis to such evaluations. The government of Sri Lanka is mindful of these aspects and views M&E from a wider country-based context.

Conclusion

Institutionalization of MfDR has been instrumental in creating a shift in focus in M&E systems and practices by expanding the focus of “results” in development. The ambitious institutionalization of MfDR in Sri Lanka has laid the foundation for the emergence of a thriving results-focused M&E community in the Sri Lankan government. Although it has a long way to go before MfDR will be assimilated into the public sector culture, the initiatives have achieved significant progress toward a government-wide shift to results-based M&E practices.

Velayuthan Sivagnanasothy

Director General, Department of Foreign Aid and Budget Monitoring, Ministry of Plan Implementation, Sri Lanka

Institutional Arrangements for International Monitoring and Evaluation Systems

Political action and the work of governments as guarantors of public interest and citizens’ welfare have gained renewed strength in addressing the new financial, economic, and social reality. Evaluation, in particular, as a mechanism for feedback on public action and accountability to citizens, will be high on the political agenda for the transformation of public administration.

This presentation is divided into five sections. The first deals with the commitment of the government of Spain to institutionalize evaluation and, consequently, to create the Government Agency for Evaluation of Public Policy and Service Quality. The second section discusses the results of the agency’s action and the basic principles that support it. The third section reflects on the Spanish reality of a decentralized state and its relationship to evaluation. The fourth section re-

fers to the work involved in the evaluation of public policies and programs in response to political reaction to the impact of economic and financial crisis on the whole population. Finally, the fifth section concludes with some ideas on the institutionalization of evaluation policies and programs.

Institutionalized evaluation of government policies in Spain

The Government Agency for Evaluation of Public Policies and Service Quality is the result of the commitment included in the electoral program the Spanish Socialist Workers’ Party presented in the 2004 general elections. In the program’s chapter on public administration guaranteeing essential public services, “the creation of a Government Evaluation Agency” was identified as a priority project. This commitment was renewed in the program in 2008, with the idea of strengthening the Agency for Evaluation and Quality.

The creation of the new institution was part of a reform program that included a set of far-reaching measures in the service of innovation and modernization of public activity, among which are the State Agencies Act, the Basic Statute for Government Employees and the law on citizens’ access to e-government.

The institutionalization of evaluation through the creation of the agency was preceded by a profound and extensive analysis conducted by a Commission of Experts, which was given the task of making the methodological and institutional proposals most recommended for starting up the agency after a review of international experiences and a diagnosis of the status of evaluation in Spain.¹⁶

The diagnosis showed the lack of a culture of evaluation in the administrative practice of public institutions and, therefore, the availability of the necessary tools to carry out evaluations of public policies. As a result, a significant shortfall was found of what might be called the design of, or “evaluator approach” to, public policy. It was evident that evaluations of public policies were not being carried out in an institutionalized or systematic manner. The lesson drawn was that accountability and transparency of public resources were not visible to the public, or, in other words, that citizens did not have access to this instrument to measure accountability in the management of public resources and assets.¹⁷

The consensus established between the political commitment expressed in the government program and

academic guidelines from the Committee of Experts was reinforced by the legislature with the addition, on July 18, of a provision to Law 28/2006 on state agencies to improve management of public services. The additional provision authorizes the government to establish the Agency for the Evaluation of Public Policies and Service Quality.

Consequently, the new agency was the result of a triple endorsement—political, academic, and legislative—and explains the government’s wish to place the administration at the service of citizens by making it more transparent, with full and fair accountability on the part of incumbents and civil servants, to enable an effective participation in public affairs by informed citizens. Government resolve should be in tune with the principles of efficiency, participation, and accountability that underlie good government.

Evaluations of public policies and programs of the General State Administration and the drive to improve the quality of management of public organizations are the two pillars of the agency’s activity. Improving the quality of government performance will in turn improve the service it provides to its citizens. Thus the action of the agency serves a dual purpose: political and socioeconomic. The political goal refers to accountability, transparency, and democratic quality, whereas the socioeconomic objective refers to the pursuit of efficiency, the rationalization of resources, decision making based on evidence, and evaluation of policy outcomes in society.

The agency is assigned at present to the first vice president and minister of the presidency. Consequently, it occupies a strategic position in the organization of the government, inasmuch as it is within the ministry that handles executive and legislative affairs jointly. This location of the agency in the executive branch confirms how important evaluation has become, as well as the government’s awareness of its dual usefulness. The government recognizes the value that evaluation provides both at the policy formulation stage, that is, on a purely legislative side, and at the policy design and implementation stage, which is an executive job.

The evaluation of public policy: Scope and practice

The evaluation and promotion of quality developed by the agency are largely determined by specific parliamentary and government mandates. This feature strengthens the legiti-

macy of the work of the institution. Thus, the agency must evaluate the following:

- Programs and public policies annually commissioned by the Council of Ministers
- The degree of implementation and success of the main measures of the national reform program
- The Spanish Strategy of Sustainable Development in collaboration with the Sustainability Monitoring in Spain (*Observatorio de la Sostenibilidad en España*)

The agency also must do the following:

- Submit an annual report to the Congress of Deputies on the activity of state agencies and their commitments to improve the quality of services provided to citizens
- Conduct an annual report on the Quality of Public Services Department (*Departamento de Calidad de los Servicios*), including citizen perception of the quality of those services

Since its inception, the agency has produced a total of 15 evaluations, which, in most cases, correspond to the government’s commitment to review the policies of the National Reform Program (NRP). Also being evaluated in 2009 are eight policies, including five for the NRP. The evaluations have focused on various issues. However, policy evaluations include social issues (increases and reductions in social security contributions, education grants, and the participation of the general state administration in the system for autonomy and care of the unit), and the promotion of economic activity (rationalization of drug expenditures, administrative procedures for startups or lines of finance for entrepreneurial activity). Certain environmental policies have been evaluated (national registry of allowances of greenhouse gases, the National Parks Network). Also, evaluation of the innovation and technology policy has risen in importance. For example, programs that promote research, development, and innovation have been evaluated. These include the *Ingenio 2010* plan to improve the technology transfer system in line with the government’s concerns.

The guiding principles that underpin the agency’s performance are (1) independence of criteria (opinion and judgment in carrying out work based on values of public ac-

countability and professional competence), (2) transparency (accountable to citizens), (3) autonomy in management and responsibility in the achievement of results, (4) quality and continuous improvement to deliver services in an innovative manner, and (5) professional ethics and public accountability, understood as the commitment of the agency staff, especially their directors, to abide by the values contained in the agency's code of professional ethics and in the rules of conduct applicable to all government employees. Strict compliance with the principles confers maximum stringency, systematization, and transparency on the activities of the evaluation of public policies. Values are formulated under these principles.

The dissemination of evaluation results, the publication of the activity reports, approval of the ethics code, the programming of training agency staff, or development of methodology are indicators of the habitual exercising of these principles.

The evaluation of public policies in a decentralized state: Cooperation with other levels of government

The evaluation of public policies, as mentioned, is a policy and management tool that allows the government to review the designs of policies and programs implemented by the administration at any time, if not the actual social need for these policies. This check is even more necessary in a government as decentralized as Spain's, where policies sometimes are designed by one political entity and implemented by another, or, they are designed with the help of different entities and implemented by one of them or by a third party—all with complex financing arrangements. The agency's activities should be configured as a staple of a cooperative government evaluation system, in which the agency's added value becomes a complement, and in which it has its own specific space.

Also, it should cooperate bilaterally with other levels of government—autonomous and local. In that sense, the agency statute not only offers the possibility of technical cooperation and joint effort with the autonomous communities, but also opens up to the participation of these entities in the institution's top management through their representation on the Governing Council. Therefore, evaluations of the results of public interventions, the quality of services, and income from the organizations are key aspects in assessing the patterns of good governance.

The value of evaluation in public action

The market-based economic system is incapable of responding to social needs derived from the principles of redistribution, equity, and social justice. This inability has become evident in different historical moments and has been exacerbated in crises that clearly show that the market alone cannot respond to emergencies and that government intervention is required to bring order and regulation. The need for public intervention is, therefore, one more sign of the imperfections of the market. Hence the need once again for the role of politics and government action to guarantee the general interest and come up with a political strategy to fit the new financial, economic, and social reality.

In Spain's case, that role has materialized in recent times in the Spanish government's commitment to renewing the economic growth model, with the strategic goal of achieving sustainability in three key dimensions:

- Economically sustainable in the medium and long term. The idea is to develop a growth model based on the known economy and innovation that will increase the levels of competitiveness of the Spanish economy and successfully confront the phenomenon of globalization.
- Socially sustainable. This objective promotes stability in employment, equal opportunities, and social union.
- Environmentally sustainable. This objective addresses the European targets on energy and climate change.¹⁸

At the same time, a management model must be developed in keeping with this political strategy, focused on revitalizing the government's spaces of action so that they play a proactive role in this design and effectively encourage activity and competitiveness of production and, furthermore, play the role of public administrators and service managers to the public's satisfaction.

This new model should gravitate toward evaluation as the hub of the new public administration. Evaluation, as a key concept of the new framework of decentralization and administrative modernization, seeks to cover all areas and reach people, organizational processes, and outcomes of the intervention. We can therefore say that the evaluation has four primary objectives:

- Accountability among managers of policies, programs and public projects
- The consideration of the actors involved in politics
- The effective and efficient management of public resources
- Analysis of the impact of public interventions

So, it is possible to identify the instrumental nature of the evaluation in two ways: policy and management. On the one hand, it enables social and political control, aimed at meeting the proposed objectives (effectiveness), and, second, it facilitates management improvement through continuous learning (mainly efficiency).

Therefore, the observance of the principles of effectiveness and efficiency becomes more important in the current context in which the evaluation has an essential role in the design and implementation of public policies. The contribution of evaluation in the current context grows and materializes on two fronts:

- Improved efficiency in the delivery of priority services and in meeting strategic objectives, which we know contain measures aimed at families, businesses, or promoting employment.
- Appreciation of compliance with the principle of efficiency in spending, which will lead to more rational distribution.

This requirement must go hand in hand with the principle of transparency and public accountability. We do not outline these two criteria independently, because in the case of evaluation, effectiveness and efficiency are not understood separately in that one needs the other to be complete. So that, “only after reaching efficacy can effectiveness be observed, since otherwise you could establish an economy of means, but without fulfilling the goals and objectives proposed in the policy design” (Bañón 2003).

Sustainability of evaluation systems

To conclude with a reflection on the sustainability of the evaluation systems, the following are some ideas that may result from this discussion:

First, the practice of evaluation contributes to transparency in public administration. Evaluation agencies, partic-

ularly Spanish ones, are a direct result of that intention and a catalyst for further efforts in that direction.

Second, the evaluation of public policies and programs is based on the idea of building a government organization that legitimizes its action depending on the performance it generates in the provision of tangible results for citizens.

Third, the independence of criterion used in the evaluation process is correlated to strict methodology, quite separate from the executive or legislative workplace where the evaluator institution is located.

Fourth, evaluation is not an end in itself, but acquires value in decision making, to replenish the beginning of the management cycle, the development of budgetary programs, and the design of public policies. It is therefore important to design procedures to incorporate the improvements early in the process and as effectively and efficiently as possible. Otherwise, evaluation itself would only increase the expenditure incurred by the policies without ensuring its effectiveness and efficiency.

Finally, this cannot be achieved without government commitment to make effective and efficient policies, and institutional technical support offered by the State Agency for Evaluation of Public Policies and Service Quality.

María Luisa CarcedoRoces

President of the Government Agency for Evaluation of Public Policies and Service Quality, Spain

Institutionalizing Monitoring and Evaluation by the Public Service Commission of South Africa: A Continuing Journey

Background

The establishment of monitoring and evaluation (M&E) in South Africa was a process that paralleled the heralding of democracy in 1994. Democracy brought about the modernization and transformation of the state and government. Before 1994, the government lacked significant M&E, as it is currently understood. Issues of transparency and accountability and the public engagement on the performance of government were largely absent. M&E in South Africa was not only about performance improvement, but also a critical strategy and tool for strengthening democracy.

The democratic government's accountability framework has been supported through different pressure points and streams of performance information placed in the public domain and the administrative and political decision-making environment. M&E information has ensured a focus on the actual performance against the plans—the promises versus the reality. In light of the prominence and value attributed to M&E, M&E institutions exert a particularly important influence on South African society. It is against this background that the institutionalization question is addressed. This paper moves from the premise that for a sustained M&E culture, central M&E institutions must engage with all sectors that have a common interest in good governance, which should include those sectors that may be critical of government. Such an approach will provide a broad-based approach to address the myriad of performance issues in a transformation context.

The Public Service Commission of South Africa

The original Public Service Commission (PSC), which was established in 1912, was fashioned after the British Civil Service Commission. By the 1980s it had burgeoned into a vast and powerful statutory body. With the new constitution, the PSC was reestablished as one of the institutions supporting constitutional democracy. It had multiple mandates, and between 1994 and 1996, it changed from an executive policy-making and implementing body to focus exclusively on investigating, monitoring, and evaluating public administration. With this new mandate, it has become the central M&E player in the country and now provides support and leadership for the development of M&E at the regional and continental level. In this way, it is different from its counterparts in most parts of the world where the focus remains on administrative issues; the PSC of South Africa has embraced and developed its mandate to become a key player in M&E.

Given the limited resources available to the PSC, against an onerous responsibility, the PSC has developed a differentiated approach to institutionalizing M&E. This involves forging social compacts, partnering with the M&E community at country and continental level, engaging the academic and research sector on its products, and generally playing a strong advocacy role for different focal areas to bring about change in the country. It thus has had not only to develop technical and managerial expertise in M&E, but

also to engage politically and with various constituencies to maximize its influence to be the “custodian of good governance,” its motto. By taking such a path, it has benefited from engaging and supporting those institutions whose primary purpose is to ensure a form of good governance, whether it is financial (the auditor general), transparency (Transparency International), and good governance initiatives of the continent (African Peer Review Mechanism), or thrusts from the Commonwealth Association for Public Administration (CAPAM), to name but a few. Its work is supportive of the Millennium Development Goals (MDGs). Finally, it joins the networks around the world seeking to measure and make public the performance of government, and ensure transparency and accountability.

This paper briefly explains the strategy used by the PSC to institutionalize M&E in the country. As mentioned, it has explicitly set about using approaches that are democratically enriching and forming partnerships to support the thrust for good governance as a key strategy. In using this embracing approach, it has managed to tap into the potential of stakeholders and role-players in the country, who possess insights and expertise that can be harnessed to help government improve. Importantly, such an approach has given the institution more credibility and respect, elements that are important when it comes to producing and presenting evaluation findings.

Advocating change to entrench good governance

The PSC has utilized its constitutional mandate effectively to craft an M&E program that addresses all areas that influence good governance and service delivery. It recognized that the democratic transition period (1994–2000) was a difficult one, as new political and administrative leadership emerged, and new policies had to be developed and implemented. The PSC's approach during this period was largely one of advocacy and support, producing guidelines and manuals to support the new government. It also raised the profile of M&E by engaging with the M&E community at the country and regional level, the most pronounced being in 2004 when it cohosted the highly successful Third African Evaluation Association (AFREA) conference in Cape Town, South Africa. This was an important event given that it helped to dispel the notion that performance and democratic culture did not exist on the continent, and many key speakers spoke about how M&E could change

the stereotypes of the continent and region, which tended to be negative. The event also brought together a community of practice, and the papers delivered across all development sectors proved that, indeed, capacity was present on the continent, as in the case of South Africa to use M&E to improve government performance and support democratic engagement. What was important about the AFREA Cape Town event was the support from many societal and academic sectors for the M&E of governments in Africa. Events such as these may be viewed as “institutionalizing moments,” in that they help propel the discipline forward. Other such events, in which the PSC engaged during this advocacy for M&E phase, were the subsequent AFREA conferences, each of which has helped to consolidate the community of practice that now exists continentally and that has increased intra-African dialogue on M&E.

In South Africa, the PSC has helped form the South African Monitoring and Evaluation Association (SAMEA) conference, and has since co-hosted two conferences (2007 and 2009). It has signed a memorandum of understanding with SAMEA to ensure a long-term relationship, which it believes is important to professionalize the discipline in the country.

The advocacy element of M&E has been necessary and important, given the historic and contextual factors in the country, which means that the good governance, with its supports and transparency and accountability imperatives, cannot be regarded as a given. The very act of M&E sometimes may be perceived as threatening to governments, as M&E talks truth to power. It is for this reason that full responsibility for M&E cannot be accorded to one or two centralized institutions; it must be a shared responsibility across society and its institutions, so that different performance perspectives can be provided, and different pressure points used to improve government performance. Only when M&E is broad-based, will citizens be empowered to hold government to account, and democracy deepened. The contribution of M&E to citizen empowerment has been an important consideration in the design of the PSC’s M&E program, and this is particularly important in a country in which political and social alienation reached a point at which most citizens disregarded government. To reverse this situation, a citizen-centered M&E component was needed, whereby citizens would be capacitated to produce and apply M&E in a useful manner.

The links that were forged with the academic and research sector come about at events when the PSC presents its annual State of the Public Service (SOPS) reports to a cross-section of society roundtables for discussion. These events have proved useful in that the SOPS offers a critique of methodology and perspective, and demonstrates once again why M&E must include the work of as diverse a grouping of persons and institutions as possible. These efforts will illustrate how a meta-evaluation has been produced and engaged with, and where the results go to bring about the desired change. Other events and strategies that are important for forging social partnerships include the following:

- Public hearings on disability where representative organizations attend and provide lived-experiences of their challenges
- Partnering with organized labor in conferences dealing with key labor relations issues
- The development and holding of the leadership of the Association of African Public Service Commissions (AAPSCOMS), an umbrella body for PSC and Civil Service Commissions in Africa
- Collaborative research with the formal research sector, for example, the evaluation of government projects on poverty
- Participation in training programs at the country, continental, and international level, such as guest lectures on M&E to graduate programs, supporting the Development Evaluation Seminars and teaching at the International Programme for Development Evaluation Training (IPDET), among others
- Presenting papers at relevant seminars, conferences and special sessions

From the above, it is apparent that the past decade has been one of charting new ground, and seeking support for the mandate of M&E from a cross-section of sectors. This has helped to get these sectors to understand the role of the PSC and also has helped the PSC to develop its own thinking and methodologies on how to conduct its work in a more effective manner. Currently, the PSC is highly respected across these sectors, and it is seen to be independent, doing its work in a transparent and credible manner, and without fear, favor, or prejudice. It is not viewed

as an extension of the executive, but as an M&E leader in the country.

M&E from process to products

The above also points to a large extent to the processes followed in securing the support of other key sectors to bolster M&E capacity in the country. This section looks at some of the processes that have been followed to ensure that the products of the PSC result in results.

National Anti-Corruption Hotline

Given the high priority of fighting corruption and graft, the PSC has set up a National Anti-Corruption Hotline (NACH), a toll-free facility operating 24 hours a day, seven days a week, in 11 official languages. The hotline receives complaints from the public on corruption and service delivery. It has proved popular with more than 7,000 complaints lodged to date and directed to departments for investigation. Its accessibility means that checks and balances within society can ensure that corruption does not go unreported. The hotline has had challenges, the main being the lack of investigative capacity within departments to follow up on cases referred to them. This challenge is being addressed through better collaboration with law enforcement agencies, but it remains a key factor as timely addressing of corruption cases is a key deterrent factor. Thus far, several officials have been dismissed because of complaints lodged, and more than South African rand (R)100 million has been recovered as a result of this facility.

Investigations

The PSC conducts its own investigations, and upon receipt of a request or complaint. Several complaints have received intensive media coverage, which helps to restore public confidence in the institutions of democracy. Tangible outcomes have arisen from the investigations.

Implementation of the transversal public service M&E system

The PSC has designed a good governance M&E system to assess departmental performance across the nine constitutional values and principles for public administration. The system, implemented since 2001, has thus far assessed more than 100 departments. For each of the principles and values, indicators are chosen and assessment takes place according to verifiable standards. Scores are awarded, and

recommendations made and tracked. The process for implementation includes a briefing session with the departments' management thorough ongoing engagement over the three-month period of assessment. The draft report is presented to the management on completion. Thereafter, the department is given 10 days to make written submissions to the PSC, which may be considered in reviewing the draft score. Once the report is finalized, the department is given six months to implement recommendations, and the implementation thereof is reported on to relevant political and administrative structures. The system is now well known and has worked by reinforcing other M&E work, as it also draws on existing evaluations, such as that of the auditor general. Through a mutually reinforcing process, the question of performance management is entrenched.

At the aggregate level, compliance results have improved slowly. The pattern has been uneven, however, and the uptake on the initial set of recommendations has in many cases been poor, resulting in repeated poor assessments for several departments. Given that the system has not changed significantly over time, it is of concern that departments continue to score poorly in some of the areas critical to good governance, such as poor public participation and inadequate decision-making processes. It may be that the turnover in leadership at the management level contributes to the limited improvements that have been seen.

Citizen M&E

The PSC has, through announced and unannounced inspections, tested the extent to which public services are provided according to the *Batho Pele* (people first) standards. These are eight standards that stipulate how government departments should operate, and what citizens can expect of them. Several surveys have been conducted, and specific investigations have resulted in reports presented to the political leadership for action. Last year, the focus was on police stations, and this year it is on clinics. These reports provide a quick and objective assessment of the facilities. By tracking the implementation of recommendations and return visits to the facilities, the PSC can see whether this form of M&E is delivering the expected results.

Evaluation of department heads

The PSC manages the secretariat for the evaluation of department heads. In this, a department head is evaluated by

a panel composed of key people from the sector relevant to the department, a peer, and others who know and understand the department. An advantage of the process is that it forces reflection and provides valuable feedback on the performance of the department heads and the departments by a well-constituted committee to the minister and the department heads. While the process is useful from a transparency and accountability perspective, further development is required so that individual performance is aligned to organizational performance. The impact of this evaluation process on accelerating and improving the management of performance within departments is not known. Questions that are being addressed currently are the arrangements for future evaluations of department heads, which is tied up with broader issues of performance contracting.

Releasing results of reports that have national importance

The PSC annually releases results from some of its systems, which have national importance. These have become key points for deliberation in the country. The most important of these annual performance releases are as follows:

- Financial disclosures of the Senior Management Service (SMS). The financial disclosures, whereby all 6,000 SMS members are expected to declare their financial status and interests, has been a measure to prevent conflict of interests, given that decision makers can abuse office by making decisions for personal gain. It ties up with the issue of fighting corruption and making the public service corruption resistant
- Financial misconduct reports. All departments are obliged to report on their cases of financial misconduct annually. This provides a status of the situation and helps to quantify graft.

Answering the brief? M&E links to the reform and modernization process

Based on the above, this section attempts to pull together some of the arguments made to answer the brief of this paper. M&E makes an important contribution to the reform and modernization process of government in South Africa. Apart from its important contribution to democracy, whereby it provides previously unknown government per-

formance information to citizens, and increases the levels of transparency and accountability of the democratic state, it also has achieved the following:

- Improved the receptiveness of government to being measured, which means that it has to review how it works and as such has instilled a performance ethos with departments. The guidelines by the National Treasury, the auditor general, and the Department of Public Service and Administration, in particular with regards to planning and reporting, has for the first time quantified what needs to be done, and with what resources. Outputs also have to be reported on.
- The results-based orientation is greater now than was the case previously, where reports such as department annual reports were largely related to activities. The current format of the annual report is an M&E document, and it allows the public to engage on actual performance versus plans, through these oversight reports.
- The process used by oversight or M&E agencies and bodies, which include parliament and its committees, have been informed by M&E reports and information. This allows for a more directed and objective engagement when accountability structures call departments to account.
- A key user of M&E in the country has been civil society and the media. This is apparent in their participation in the events of the PSC, which are diverse and indicate the broad-based interest held in government performance.

Location of the M&E function

The PSC, as a constitutional body, is outside the executive branch. This has been an advantage in that it has had the latitude to engage more broadly and is perceived as independent and impartial. A challenge is that it makes recommendations and waits for it to be implemented by the executive office. The rate of implementation has been improving, but is not satisfactory. It will soon receive powers to issue directives. One way in which the PSC has addressed the disjuncture between producing findings and realizing results is by producing reports in which all its recommendations are made public, and the action thereof reported on. This identifies where the problems are in mak-

ing the translation of findings into results. The current arrangements are thus appropriate, because they retain an arm's-length distance, but also keep avenues open for engaging the executive office, which is willing to receive PSC findings. The PSC has managed to get audience with the highest offices of the land, and as such, its location has not been an inhibiting factor.

Institutional design and its impact on M&E decision-making process

The PSC is led by a board of commissioners who are appointed by the president of the country (nine of whom are based in the nine provinces) and who may serve a maximum of two renewable terms of five years each. The board is supported by an office of public servants who are appointed according to the terms of the Public Service Act. This provides adequate capacity within the organization to engage with the various legislatures, as well as adequate capacity and skills to produce the type of networking with stakeholders mentioned. In terms of decision making, the PSC acts as one body, and although there is space for provincially specific research, the decisions of the PSC are that of a unified institution.

The institutional design has not impeded the speed of work, and provincially based commissioners are able to unblock problems at that level. The question of how findings are presented, after they have been engaged with and produced by the PSC, and acted on is being addressed on an ongoing basis by customizing approaches that seek to get the most influence in the shortest period of time.

The impact of institutional design and decision making

The design of the board supported by an office (or secretariat) has not adversely affected decision making, as over the last decade management steps have been institutionalized to address the broad scope of work undertaken, in a manner that has not compromised quality. The PSC has commissioned an assessment of its work, and the issues raised therein are being incorporated into the organization as it moves forward.²⁰

Link of PSC M&E system to sector systems

The link between the PSC and the sector systems takes place more by agreement and convention, than on a formal

basis. The PSC augments its research with that of the other government institutions. Thus, its SOPS reports are heavily referenced with work from these institutions. It draws on research undertaken by research units in the country and abroad, and attempts to avoid duplication of effort. It participates in initiatives led by the presidency from the ministry for M&E, and works with all those involved in M&E in the country. Through its memorandum of understanding with SAMEA, and as a follow-up to the SAMEA 2009 conference, it will engage more directly with the nongovernmental organization and research sector, perhaps coming up with an evaluation program that sees the existing harmonization of M&E and maximization of evaluation capacity. This remains a work in progress. There is no overt competition to the work of the PSC from the M&E sectors in the country, and this may be due to an established history of cooperation.

Sustainability of systems

The PSC is now moving to support more directly the M&E capacity-building initiatives in the country. This will take the form of producing and presenting its M&E guides, and in general providing a much-needed support for emerging M&E practitioners who have been appointed in many departments. It seeks to do this in a collaborative program with SAMEA. It believes that holding more sector- or province-specific events on M&E will be a way to build M&E capacity in the provinces. M&E currently is concentrated at the national level, which is an anomaly given that most of the departments and population are not. Through the more effective synergy between SAMEA and the PSC offices, whereby communities of practice flourish at the provincial and district level, this necessary change can occur.

Incentives for evaluation?

Creating incentives for evaluation is a difficult question, and more needs to be done. Thinking is under way on linking departmental results with some form of acknowledgment. Currently, the reports of the PSC rate departments against their peers, but this has not taken the form of awards for good performance, except citations in reports. The question of annual league tables for performance across a range of performance criteria is being considered to bring in the incentive element more directly.

Translating process into results

The efficacy of an M&E system depends on buy-in and understanding, and thus the advocacy element and engagement are needed. When the PSC brought decision makers onboard during the evaluation process, it received greater cooperation. What still needs to be achieved more fully is ownership of evaluation results, which will come about when there is a perceived benefit to M&E, and those evaluated see the direct benefits for their own operations. The challenge will remain between doing what is considered to be necessary from the top, and what is deemed important from the bottom. A flexible M&E systems approach, which caters to all of these needs, is a way to institutionalize M&E.

Indran A. Naidoo

Deputy-Director General: M&E:

Office of the Public Service Commission, South Africa

Session 2 Notes

- ¹⁴ The ePMS can be accessed through <http://www.fabm.gov.lk/>.
- ¹⁵ Ministerial Order, APU/1463/2004, of May 19, (BOE May 25).
- ¹⁶ For the full report, see http://www.aeval.es/comun/pdf/Informe_comision_expertos_esp.pdf.
- ¹⁷ For the full provision, see http://www.aeval.es/comun/pdf/ley_de_agencias.pdf.
- ¹⁸ For the complete text, see www.plane.gob.es.
- ¹⁹ For the full report, see www.psc.gov.za.
- ²⁰ See the report on the PSC's Web site, www.psc.gov.za.

Session 3: Institutional Arrangements for Monitoring and Evaluation Systems in Latin America

Challenges and Opportunities for the Institutionalization of Regional Institutional Agreements for Monitoring and Evaluation Systems in Latin America: The Experience of Costa Rica

In democratic systems, modernization of the state is not only a central concern, but also a permanent activity that seeks to strengthen public management and guarantee adequate functioning of planning, budgetary, and evaluation systems, and to enable the latter to measure the impact of public policies designed to improve the quality of life of citizens. In Costa Rica, the Planning Ministry (*Ministerio de Planificación Nacional y Política Económica*) is charged with the task of strategic planning, monitoring, and evaluation of public-management strategic actions, and to provide consultancy at the highest level for establishment of the corrective measures necessary to permit the various sectors to achieve their proposed goals in each administration.

This function takes shape in the legal framework through Law 5525 (1974) created by the ministry, which in turn creates the National Planning System, the Decree for the National Evaluation System dated 1994, and Law 8131 (2001) of the finance administration and public budgets and its regulations, a norm that—together with the creation of institutional planning units, services comptrollers, and internal audits—provide an institutional framework that strengthens ministry of planning, monitoring, and evaluation actions.

National Planning System

This system is made up of the Planning Ministry's institutional and sectoral planning units (as appropriate) and the remaining public institutions and mechanisms of coordination and consultancy, such as consultancy boards, interinstitutional committees, consultative commissions, and others. The ministry works on each administration's National Development Plan (NDP) through a participatory process in which sectors, leaders, and new authorities present their strategic development proposals for each thematic area or sector. In addition, the ministry maintains a current and prospective diagnosis that permits the evaluation of the internal and external events that can affect national development, such as the various policy proposals that require promotion or execution at the governmental level. Thus, the NDP orients the objectives of the government's actions to promoting the country's development. It establishes policies and objectives at the regional and sectoral level for public entities, ministries, and other government organs. The 2006–10 plan, in effect at present, contains the following goals and activities:

- Eight national goals: (1) combat corruption in all areas of action of the public sector, (2) reduce poverty and inequality, (3) increase the growth of the economy and employment, (4) improve the quality and increase the coverage of the educative system, (5) halt the growth rates of criminality, drug trafficking, and drug addiction, and to revert the sensation of insecurity on the part of the entire citizenry, (6) strengthen public institutions and to order state priorities, (7) recover and broaden the country's

transport infrastructure, and (8) ennoble foreign policy and restore Costa Rica's role in the world

- 84 activities in the following areas: social and the struggle against poverty, education, health, citizen safety and crime prevention, cultural, financial, productive, foreign trade, tourism, infrastructure and transportation, science and technology, employment and social security, financial supervision and monetary policy, environment, energy, telecommunications, governmental coordination, and foreign policy
- 141 strategic actions
- 492 associated goals in the following areas: social and the struggle against poverty, education, health, citizen safety, and crime prevention, cultural, financial, productive, foreign trade, tourism, infrastructure, and transportation, science and technology, employment and social security, financial supervision and monetary policy, environment, energy, and telecommunications, governmental coordination and foreign policy
- 95 public institutions with 730 linked goals

In addition, the NDP includes a Contract with the Citizens, which constitutes the commitment of the leadership to comply with the established goals. The NDP lies within the orienting framework of the institutions' annual operational plans and yearly budgets. Its purpose is to sustain the president of the republic's constitutional mandate with respect to directing, planning, coordinating, and evaluating the strategic objectives, policies, programs, and projects for the country's development. To achieve this, the Planning Ministry designed the Annual Institutional and Sectoral Programming, Follow-up and Evaluation Matrix (MAPSE-SI), which, as its name indicates, is the instrument that institutions and sectoral rectorates have at their disposal for programming and evaluating the advancement of public management in compliance with NDP-established goals.

Thus, the state's strategic planning is only one component of the mechanism that permits defining goals and the means to achieve these, and this defines institutional behavior. The remaining basic piece is the monitoring and evaluation (M&E) of public management, which, in the case of Costa Rica, is focused on the strategic actions defined in the NDP.

National Evaluation System

From its beginnings, the National Evaluation System (NES) is limited to the executive. It was designed as a self-evaluation mechanism of the government's strategic actions, with the objective of offering timely information to allow the executive branch (ministers, governors, executive presidents, managers, and even the president of the republic) to make decisions related to advancement of the NDP.

In addition, the NES establishes an evaluation model for other government programs considered strategic through the M&E of specific policies, plans, programs, and projects. Incorporation of the monitoring and follow-up system of public investment projects that are linked with NDP goals was initiated in 2009.

Law 8131 on Financial Administration and Public Budgets clearly indicates certain M&E characteristics that should be implemented each year by the executive branch and more concretely, by the Planning Ministry and the Ministry of Finance, which are amplified in their regulations. In budget formulation, accepted techniques and budgetary principles are utilized, based on functional criteria that permit evaluating compliance with annual plans and policies, as well as the incidence and economic-financial impact of execution. For this, attention should be paid to elements such as provision of services, production of goods, and general management functions and support of each organ, entity, or institution.

The Head Office of the National Budget, the Technical Secretariat of the Budgetary Authority, and the Ministry of National Planning and Economic Policy, according to their respective competencies, define the methodology and instruments utilized in the evaluation process. For the latter, the aforementioned instances will establish the mechanisms and communication and coordination that they regard as pertinent, to the effect that the requirements for information are the most uniform possible and consistent with the needs of each, without this implying duplication of functions. In all cases, the contents of the reports respond to the technical criteria and guidelines, according to their areas of competency, established by the Head Office of the National Budget, the Budgetary Authority, and the Planning Ministry.

For the information of the Ministries of Finance and National Planning and Economic Policy, the Head Office of the National Budget should prepare a report on the re-

sult of the programs executed during the corresponding economic period.

The Ministry of Finance and the Planning Ministry evaluate institutional management results to guarantee compliance with objectives and goals and the rational use of public resource. Evaluation reports are at the disposal of the comptrollership to measure compliance.

The Planning Ministry remits to the general comptroller the final report on the results of execution of the budget, compliance with NDP strategic goals, objectives, priorities, and actions, and contribution to the socioeconomic development of the country. These reports include the explanatory elements that measure the effectiveness of the programs, the cost of the services, and the efficiency of the use of public resources.

As the result of follow-up and evaluation, periodic reports are issued and presented to the hierarchies of the organs and entities comprising the System of Financial Administration of the public sector, with the aim of applying corrective measures to ensure compliance with the programmed objectives and goals, and an annual report.

The Planning Ministry's reports, as established by Law 8131, Articles 52, 55, and 56, are carried out gradually in terms of their diverse aspects and evaluation contents. This gradualism includes two dimensions: coverage areas and evaluation types. Coverage areas of the evaluation processes are national, regional, sectoral, and institutional. They are incorporated in gradual fashion according to the information available. Gradualness with respect to evaluation types are applied in the following order: (1) the evaluation process of the goals of policies, plans, programs, and institutional projects; (2) the evaluation process of the efficacy of policies, plans, programs, projects, and institutional organization; (3) the evaluation process of the efficiency of policies, plans, programs, projects, and institutional organization; (4) the evaluation process of the impact of policies, plans, programs, projects, and institutional organization; and (5) the measurement process of the costs of policies, plans, programs, projects, and institutional organization.

Guidelines

To implement each of these aspects, each year the Planning Ministry issues the Technical and Methodological Guidelines for Sectoral and Institutional Strategic Programming and Follow-up, and Sectoral Evaluation. These guidelines

“are oriented toward improving the quality of sectoral and institutional strategic programming, follow-up, and sectoral evaluation of NDP goals, National Public Investment Plan projects, and the institutional budget, generating information of a strategic nature for the evaluation processes of management, results, and accountability.”²¹

At present, a three-party commission made up of the Planning Ministry, the Ministry of Finance, and the general comptroller, discusses issues such as the guidelines. These discussions have aided in arriving at consensus and especially in achieving the placement of evaluation, not as a form of auditing or punishment, but rather, as a transparency and decision-making tool, which it is necessary to achieve better public management results.

A favorable evolution has been achieved in the search for a public administration that establishes goals and acts on the basis of the latter, although it is evident that difficulties and challenges will persist. One of these is planning in the longer term, which allows for carrying out evaluations with greater impact, in addition to creating commitments for those who assume power, not to impose solutions, but, conversely, so that those entrusted with power have goals, with each choosing the pathway to comply with these goals.

Advantages and disadvantages of NES

Taking the evaluation task established by NES as a reference, as well as the law of the Financial Administration and Public Budgets, certain advantages and disadvantages exist and function from the executive branch.

The advantages

The M&E process in the hands of the executive branch enables those who implement public policies, and the programs and projects contained therein to issue and control data that contribute to the system in a firsthand and official manner.

If there is an adequate level of communication, the advancements and delays reported allow for short-term corrections of the original plan at the strategic as well as at the operative or programmatic level.

Given the existing set of regulations, the institutions and the system as a whole are obligated to implement periodic M&E, which guarantees production of data and follow-up.

The systematic nature of follow-up and evaluation of the goals of the executive branch could facilitate implementation of evaluations on the long-term impact of policies, with consequent advantages in terms of public policy design concerning identified demands and design and according to prior implementation experiences (whether successful or not).

The disadvantages

On specific occasions, modifications to the NDP, even in its most programmatic version, could yield better results than expected, collaterally affecting the plan's strategic goals. Political agreements inside the executive power could impede timely implementation of corrective action measures or could minimize public information on advancement of the NDP.

The current institutional framework separates the design of the plan from that of the budget, and similarly separates, despite efforts, the system and the focus of the evaluation of the plan for expense liquidation reports. This has been accompanied by various practical consequences, with emphasis on the following:

- An evident lack of linking between government priorities and the endowment of systematic resources for these
- The impossibility of public implementation from a multiyear perspective of goals and budget, which displays a more strategic than programmatic character, despite the design and due to an annual budgetary focus that impedes a vision of the future
- A shallow capacity for interinstitutional coordination (Finance Ministry and Planning Ministry), which is overcome by the maintenance of public administration "ghettos" and which is due to the lack of the previous links

On the other hand, a sectoral focus of government organization has been put into practice since 2006, and sectoral links have been established, including those responsible for the planning, monitoring, and evaluation of the sector and its institutions. This institutional design, however, does not have sufficient accumulated experience to date. Thus, it is necessary that it be strengthened organically and that it be distinguished from institutional planning

units that are the most programmatic links of the public evaluation system.

Present and future challenges

Information system

An information system is needed that will speed up information gathering and exchange between the Planning Ministry public officials and its links, both institutional and sectoral. At present, work is being carried out to create a platform that the NDP has outlined in such a way that in the future, it can be accessed with a special user to achieve data incorporation in real time. The Delphos system, selected to undertake this task, is "is an administrator of strategic management indicators, commonly known as the Balanced Scorecard, which permits visualization, in very simple fashion, of the strategic state of the organization, enormously facilitating its follow-up and compliance," which is compatible with the task performed to date.²²

Culture change

Rejection of evaluations is common, because evaluation is perceived as synonymous with punishment and negative consequences, and backwardness in thought that requires modification. The new culture of evaluation should not lead to discrediting an institution or a program, but rather should perfect and use it to greater advantage. Therefore, evaluation becomes a required process for the greatest exploitation of the state's resources.

Evaluation and the project cycle

Evaluation must be established as a part of the project cycle and not an optional element, to avoid developing programs in which no evidence of the results and impacts, negative or positive, is ever produced. One practice that has been carried out in countries such as Mexico is that of allocating a percentage of the budget to the evaluation of the budget itself, ensuring that the funds are available to prove whether the direction that has been chosen is appropriate, or whether a change is required.

Short-, medium-, and long-term planning

Planning in the short, medium, and long term encourages more revealing evaluations that will make it possible to make changes in the direction the administration has tak-

en at all stages of the projects and policies. Deep reflection on the needs of the country will draw the government closer to a planning and evaluation orientation and toward integrated development.

Improved routine planning and evaluation processes will lead to a public administration that is better able to respond to goals and objectives and to employ the budget in these terms. This means a change in mentality among managers and users, and the utilization of information systems that facilitate contact between those interested in the evaluation and the planning establishment at all levels and at all centers.

System sustainability

The sustainability of the M&E system is guaranteed mainly by strengthening the sectoral planning and evaluation sector and its link with NES. This agreement toward achieving sustainability has to do with the following:

- Maintaining the present decree of sectoralization, ensuring its being in effect in the future
- Ensuring sectoral decentralization of the automated information system in favor of the Delphos M&E system
- Naming sectoral links in an organic manner that is distinct from institutional links
- Strengthening and updating NES M&E techniques and practices, orienting these toward results-based management, in which the plan determines the budget

Costa Rica is making efforts to shift toward results-based management, strengthening the culture of evaluation and its advantages; however, many efforts are pending. The way to better public management is outlined; what Costa Rica, and in general, Latin America, needs is a willingness to assume the commitment of proposing quantifiable goals that permit evaluation, which makes the perfecting of projects, policies, and programs possible, and that encourages accountability.

Carla Morales-Rojas

Vice Minister, Ministry of National Planning and Economic Policies, Costa Rica

U.S. Monitoring and Evaluation Trends: Using Performance Information to Improve Results

U.S. President Barack Obama is committed to results and is taking steps to improve on the accomplishments of his predecessors. On January 20, 2009, in his Inaugural Address, President Obama said,

The question we ask today is not whether our government is too big or too small, but whether it works—whether it helps families find jobs at a decent wage, care they can afford, a retirement that is dignified. Where the answer is yes, we intend to move forward. Where the answer is no, programs will end. And those of us who manage the public's dollars will be held to account—to spend wisely, reform bad habits, and do our business in the light of day—because only then can we restore the vital trust between a people and their government.

Over the past 50 years, the U.S. Congress and previous administrations have laid important groundwork for governmentwide performance improvement, including most recently the Government Performance and Results Act (GPRA) enacted in 1993 and the Program Assessment Rating Tool (PART) introduced by President George W. Bush in his 2003 budget.

The Government Performance and Results Act

The GPRA of 1993 requires every federal agency to set strategic and annual goals (dealing with societal outcomes, not just agency activities), measure performance, and report to Congress, the Office of Management and Budget (OMB), and the public on progress relative to selected goals. Federal agencies must identify both long-term and annual goals, collect performance data, and justify budget requests based on that data.

Under GPRA, each federal program was required for the first time to identify explicit measures and goals for judging program performance and to collect performance information annually to determine whether the program was meeting those goals. Senator William Roth, sponsor of the measure, stated at the time of its enactment that the act would bring about “accountability by federal agencies for the results they achieve when they spend tax dollars.”

Unfortunately, the implementation of this law has not met its authors' hopes. Agency plans suffer from performance measures that are meaningless, vague, too numerous, and compiled too often by people who have no direct connection with budget decisions. Agencies produce thousands of pages of performance plans every year that are largely ignored in the budget process.

The Program Assessment Rating Tool

With the creation of PART within the OMB, President Bush attempted to carry out the spirit as well as the letter of GPRA. The Bush administration created the PART to shift the focus of federal goal-setting and measurement from the agency level to the program level. In the 2003 Budget, the Bush administration rated approximately 130 federal programs on their effectiveness. Within five years virtually every federal program was rated—more than 1,000 individual programs, representing 98 percent of all federal spending.

This attempt to rate the effectiveness of all federal programs on an objective and consistent basis was far from perfect. The criteria used to rate programs were not uniform and ratings were based on limited information. Its influence on budget decisions was limited. Nonetheless, with PART, the federal budget process adopted a new tool for rating program performance that in successive years attempted to generate more and more extensive information on program performance that was more objective, more credible, and more useful than had been available ever before. In cases in which a program was rated, the results of the rating process often influenced program policy and management decisions as well as the budget request.

PART is a multipart questionnaire that was developed by career staff within OMB. The questionnaire was completed by small teams of career professionals who worked within each program being assessed at larger agencies across the government. Each team completed the survey instrument with the assistance and oversight of professional OMB career staff ("program examiners") whose regular responsibilities largely focused on oversight of budget and policy development and execution for the program.

The goal of the PART initiative was to devise a tool that was objective and easy to understand. Most important, the findings had to be credible, useful, and ideologically neutral. For instance, the first draft PART questioned whether a particular program served an appropriate federal role. Be-

cause many believed that the answer to this question would vary depending on the respondent's or the reviewer's philosophical outlook, the question was removed so that the PART would yield a more objective rating.

Building on the past: Going beyond GPRA and PART

OMB's current Deputy Director for Management Jeff Zients has stated that both GPRA and PART placed too much emphasis on producing performance reports for its own sake and too little attention on analyzing and acting on that information. Zients has stated that the federal government should use program performance information to achieve the following:

- Communicate priorities, progress, and raise issues
- Illuminate what works that should be continued and what does not work that needs attention
- Motivate the best from our workforce and our service delivery partners
- Allocate scarce resources wisely

Shelley Metzenbaum, OMB's current associate director for performance management, in a study prepared before the 2008 presidential election concluded that both GPRA and PART failed to offer a comprehensive way to see how the federal government is performing and to identify the agency goals and program targets (Metzenbaum 2009). Despite large amounts of GPRA and PART performance material, she found that it is difficult to find meaningful government performance information, for example, performance levels, performance trends, and even targets, because too little attention has been paid to communicating targets and trends and too much attention has been paid to communicating the "percentage of targets met" as the primary indicator of overall performance. Metzenbaum further asserts that too little attention was paid in these initiatives to understanding the size and characteristics of problems and why performance levels changed. She states that few agencies and programs routinely analyzed their performance and other data, as businesses do, under these initiatives to generate insights needed to improve performance. In particular, she argues that the program review process has been overly subjective, creating unfair inconsistencies and frustrating disputes about what constitutes appropriate measures, targets, and

evaluation methods. She asserts that too little attention has been paid to identifying key audiences for goals and performance data, to determining their performance information needs, and to delivering information where and when it is needed in a format the target audiences can understand. As a consequence, key audiences, including Congress, field offices, delivery partners, and others, have not received the performance information that they need. Finally, she asserts that under GPRA and PART, too much attention has been given to program review, assessment, and control and too little attention to providing expert advice and to stimulating innovation, discovery, cooperation, and assistance

Guiding principles to improve federal performance management

Metzenbaum concluded that the key performance management challenge is to use—not just produce—performance goals and measures. She enunciated four guiding principles on which to base changes to current federal performance management efforts:

- Communicate performance trends and targets, not target attainment and ratings
- Encourage performance improvement with increased diagnostic analysis, practical experiments, and knowledge sharing
- Present information in ways that meet the needs of specific audiences
- Structure accountability mechanisms to encourage and inspire, not embarrass, reprimand, or punish

Metzenbaum then outlined a series of interconnected recommendations for the president and for OMB to put these principles into practice.

Recommendations for the resident

- Clearly identify presidential priority targets. Meet at least quarterly with each responsible cabinet secretary to keep agencies focused on the targets. Appoint a chief performance officer and create a White House Performance Unit working closely with OMB to advance progress on presidential priorities.
- Run goal-focused, data-driven meetings pertaining to the priority targets. Increase analysis of perfor-

mance and other relevant data pertaining to presidential, cross-agency, agency, and program targets.

- Appoint experienced performance managers to key government management positions, and recruit experts to provide counsel to cabinet secretaries. Identify and manage cross-agency targets and measures where performance improvement needs cross-agency attention and cooperation.
- Adjust accountability expectations. The president should hold agencies accountable for the persistent application of evidence, intelligence, and effort to achieve continual performance gains.

Recommendations for the OMB

- Direct agencies and programs to communicate targets and performance trends for key indicators.
- Create a federal performance portal with easy-to-find performance targets, trends, and other related information.
- Encourage outside expertise and multiple perspectives on selection of targets, performance measures, and strategies to improve performance.
- Build communities of practice to facilitate cross-agency learning.
- Increase training in performance management practice and analysis.
- Revise, but continue PART, shifting the emphasis from program rating to performance improvement.
- Remove the centrality of “rating” and “assessment,” and shift attention to performance improvement and communication, not target attainment for its own sake.
- Eliminate barriers that complicate audience-focused delivery of performance reports and information.
- Continue the President’s Management Council as a forum for senior agency deputies to meet regularly to discuss progress toward performance and management priorities and to reduce management risks.

Recommendations for cabinet secretaries and agency heads

- Immediately review agency performance trends and update priority targets. Each cabinet secretary

and agency head should review and refine their organization's strategic and annual targets to reflect and communicate the new administration's priorities, informed by a review of past performance trends.

- Run goal-focused, data-driven meetings. Each cabinet secretary and agency head should run their own goal-focused, data-driven meetings to keep the organization focused and continually searching for opportunities for improvement.
- Identify information needs of key audiences. Cabinet secretaries and agency heads must ensure that their organizations identify key audiences for federal performance information, determine their needs, and establish priorities among the audience.
- Improve federal information presentation and dissemination capacity. Each cabinet secretary and agency head should pay increased attention to the presentation, dissemination, and use of performance information to communicate more effectively with targeted audiences and inform their priority-setting and performance improving decisions.
- Create agency Web-based performance portals. Each cabinet secretary and agency head should direct their organizations to add a performance portal on their home pages that makes it easy to find performance targets, trends, and other related information.

Recommendations for the Performance Improvement Council

On November 13, 2007, President Bush signed an executive order on improving government performance that established an interagency Performance Improvement Council composed of performance improvement officers from every major federal agency who were responsible for coordinating the performance management activities of the agency, including the following:

- Developing and improving the agency's strategic plans, annual performance plans, and annual performance reports, as well as ensuring the use of such information in agency budget justifications
- Ensuring program goals are aggressive, realistic, and accurately measured

- Regularly convening agency program management personnel to assess and improve program performance and efficiency
- Assisting the head of the agency in the development and use within the agency of performance measures in personnel performance appraisals, particularly those of program managers, to ensure real accountability for greater effectiveness

Metzenbaum recommended that the interagency Performance Improvement Council

- Lead a review of PART. The Performance Improvement Council should be directed to lead a process to propose changes to PART for subsequent action by OMB. This should include revising and renaming the process to shift the emphasis from program rating to performance improvement, fixing some the questions, and adding a few new questions.
- Consider specific revisions to PART. The Performance Improvement Council should consider the following recommendations concerning revisions to the PART process:
 - Align program targets with GPRA goals and allow agencies to define what constitutes a program
 - Revise the PART scoring system and eliminate the ratings
 - Increase reviewer perspectives
 - Reorder and revise PART questions to be served

State and local government models

OMB's performance managers are particularly impressed by successful efforts at the state and local government level to use performance information to monitor, evaluate, and manage government programs. Local governments—including New York City; Charlotte, North Carolina; Baltimore, Maryland; King County, Washington; and Austin, Texas—have demonstrated how performance information collected and analyzed with the aid of management tools like CompStat and CitiStat can be used to improve outcomes, reduce crime, increase housing starts, and drive down costs. With these tools, managers lead frequent goal-focused, data-driven staff meetings that review statistics about outcomes, costs, unwanted side effects, patterns of problems in an area, probable causes, apparent effects of

prior actions, and future plans. For example, Baltimore's Office of CitiStat is a small performance-based management group responsible for continually improving the quality of local government services of all kinds. Staff analysts conduct studies to identify areas in need of improvement. Department managers participate in regularly scheduled presentations with the mayor and other top city officials designed to maximize accountability.²³

Planning for the FY 2011 budget and performance plans

OMB intends to develop an improved performance management framework aligning high-priority performance goals, GPRA performance reporting, and many program-level PART measures. It will focus on outcomes, allow comparisons across programs and agencies, and illustrate performance trends over time. It will focus on information supporting long-term decisions about targets and organizational strategies, as well as key decision makers in the service delivery chain, including those in federal field offices, state and local partners, nonprofit grantees, and contractors. It will use new information technologies to make this more feasible, less cumbersome, and far more useful than past alignment efforts. In addition, the administration is proposing historic investments in comparative effectiveness research and evaluations, integrated with performance management efforts.

Increased emphasis on program evaluations

In June 2009, OMB Director Peter Orszag wrote in his regular blog on the OMB Web site that

[W]e have not been making the right investments to build a new foundation for economic prosperity—and we need smarter investments in education, health care, and social services. Much of the time, it is hard to say whether a program is working well or not. Many initiatives drive funds to the local level, but do not track how they are spent; others spend dollars bit by bit, so the results are hard to see. Too often rigorous evaluations do not happen. This has to change, and I am trying to put much more emphasis on evidence-based policy decisions here at OMB. Wherever possible, we should design new initiatives to build rigorous da-

ta about what works and then act on evidence that emerges—expanding the approaches that work best, fine-tuning the ones that get mixed results, and shutting down those that are failing.²⁴

Beyond investing more in evaluations, we are also designing new initiatives with evaluation standards built in using a two-tiered approach—providing more money to programs that generate results backed up by strong evidence. That is the top tier. Then, for programs, with some supportive evidence but not as much, we've said: Let us try those too, but rigorously evaluate them and see whether they work. If not, we'll redirect their funds to, more promising efforts. This two-tiered structure will provide objective criteria to inform our decisions on which models to invest in. It will also create the right incentives for the future. Organizations will know that to be considered for funding, they must provide credible evaluation results that show promise, and be ready to subject their models to analysis. As more models move into the top tier, it will create pressure on all the top-tier models to improve their effectiveness, so they continue to receive support. By instilling a culture of learning into federal programs, we can build knowledge so that spending decisions are based not only on good intentions, but also on strong evidence that carefully targeted investments will produce results.

New evaluation initiative

Most agencies review ongoing programs, but not robustly enough to provide meaningful data about outcomes and progress toward the objective. Many agencies lack the capacity to carry out a rigorous, strategic research agenda, and policy priorities are established without evidence to back them up. Programs are continued year after year, without a hard, objective look at their effectiveness.

On October 7, 2009, OMB asked agencies to volunteer for an initiative designed to strengthen rigorous, objective assessments of existing federal activities to improve results and better inform funding decisions. Ongoing program evaluation research will be published online, and an interagency task force will identify and shape evaluations of programs that cross over several agencies. Participants

will serve as demonstration projects through which we can test approaches to improve program effectiveness and efficiency, share best practices, and further improve performance. After assessing the initiative in fiscal year 2011, the administration will implement governmentwide evaluation metrics.

In addition, OMB will work with agencies to make information readily available online about all federal evaluations focused on program impacts that are planned or already under way. Along with other White House offices, OMB will establish a new interagency working group to promote stronger evaluation across the federal government. OMB will allocate a limited amount of funding for agencies that volunteer to

- Show how their fiscal year 2011 funding priorities are evidence-based or otherwise subject to rigorous evaluation
- Assess their own capacity to support evaluation and suggest pathways for strengthening that capacity
- Propose new evaluations that could improve government programs in the future
- Identify impediments to rigorous program evaluation in their statutes or regulations

The initiative initially will focus primarily on impact evaluations, that is, evaluations aimed at determining the causal effects of programs. Although the administration is committed to improving a wide range of evaluation and performance measurement activities, the initiative initially will focus primarily on social, educational, economic, and similar programs whose expenditures are aimed at improving life outcomes (such as improving health or increasing productivity) for individuals. OMB will consider on a case-by-case basis the inclusion of evaluation efforts in other spheres, such as procurement, construction, taxation, and national defense. Because drug and clinical medical evaluations independently have received extensive study, they are not included.

Conclusion

This is an exciting time for those interested in improving program evaluation in the federal government and in seeing that the results of those studies have a more direct impact on federal decision making. For the first time in mem-

ory, staff members have been told that the annual budget reviews with the director for all parts of the federal budget, not just social programs, must give prominent attention to quantitative evidence of program performance and results. As a top manager said in a meeting on the program evaluation initiative this week, this is a very “fluid and dynamic process,” and more developments are sure to come.

John R. Pfeiffer

*Program Examiner, Energy Branch,
Office of Management and Budget, United States*

Institutional Arrangements for M&E Systems in Latin America: The Case of Mexico: Performance Evaluation System

Background

In Mexico, the tradition of carrying out monitoring and evaluation (M&E) programs and projects dates back to the 1970s. The Secretariat of Finance and Public Credit and the Bank of Mexico promoted the evaluation of productive and infrastructure projects through development banks, with assistance from the World Bank and the Inter-American Development Bank (IDB). Efforts to monitor and evaluate the management of public entities were enhanced in the 1970s with the 1976–82 Administrative Reform Program, which introduced a new planning and control model that grouped the various offices and entities of the federal public administration into various sectors, organizing them into sector-based groups and establishing the fundamental elements of the planning, programming, budgeting, and evaluating cycle.

In 1982, the Mexican government created the Secretariat of the National Comptroller General (SNCG), a government office with a broad spectrum of responsibilities and functions, and especially the operation of the National System for Government Control and Evaluation. This system operated through the internal control bodies within federal public administration offices and entities, and the comptroller’s offices in state governments, as well as through public commissioners and external auditors. Although its responsibilities included verifying the efficiency, effectiveness, and operational economy of the offices and entities, it was essentially more oriented toward the auditing and

supervision of public spending. In 1986, the Federal Law on Para-State Entities went into effect, and in 1990, the Mexican government created the National Program for the Modernization of Public Enterprises.

In the 1990s, an important change directed by the Secretariat of Finance and Public Credit was implemented. While the control functions carried out by SNCG continued, a strong impulse was given to the systematization of both monitoring and evaluation before, during, and after the completion of programs and projects. The Mexican government incorporated the socioeconomic evaluation of projects so that, in the structuring of infrastructure projects, prior formulations and evaluations would be carried out to demonstrate whether or not the investments proposed were socially profitable. Development banks, particularly the National Public Works and Services Bank, were pioneers in ensuring that socioeconomic evaluations were conducted for the infrastructure projects linked to the Fund for Investment in Infrastructure and selected for financing (loans, venture capital, nonreturnable financing, or a combination of the three).

During the 1995–2000 period, the composition of public spending changed, with increasing resources used by legislative and judicial branches of government, compared with those used by the executive branch. Specifically, federal resources transferred to states and municipalities increased, and resources corresponding to the federal level decreased. This strengthened the orientation of public spending toward social development with a regional perspective. The rules of operation for the programs that grant subsidies and transfers to the population were published for the first time in 1999, with the aim of assuring their efficient, effective, fair, and transparent application, and achieving the anticipated goals and objectives, their application in the target population, and the corresponding evaluation of results.

The process of reforming the budgeting system began in 1998. A New Programmatic Structure (NPS) was established with the aim of contributing better quality information on public programs and the orientation of spending. With the establishment of performance agreements between the Secretariat of Finance and Public Credit, the Civil Service Secretariat, and semiofficial entities, institutional incentives were established for these government entities to make decisions with a focus on re-

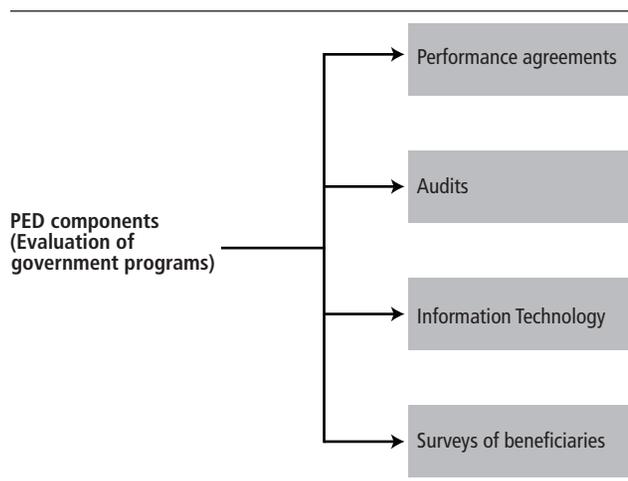
sults, within a framework of greater normative flexibility (see Figure 11).

The NPS was defined as “an instrument for planning, programming and budgeting, with the objective of improving the quality of results-based spending, through the re-composition of programmatic structure.” With the NPS, the federal government sought to achieve greater interaction among budgeting authorities in government offices and entities as well as among the units responsible for and the individuals charged with strategic planning. The budget was defined as a vehicle for establishing greater links between the objectives of public policies and those making expenditures. With these elements, the goal was to improve implementation, control, evaluation, and accountability in public spending. The NPS included new programmatic categories such as the following: function, sector-based program, institutional activity, and investment project. The mission, objective, and indicator goal were considered in the new programmatic elements.

The use of indicators to improve the achievement of strategic objectives in national planning and efficiency in the management of public spending has followed an evolutionary process that has prepared the federal public administration for more advanced phases in conducting evaluations. Between 2000 and 2006, the NPS encouraged the use of performance evaluations, urging government offices and entities to report information on their progress in reaching goals. Indicators were reported on the basis of

FIGURE 11:

Performance Evaluation System, First Stage



Source: Author.

institutional activity and the responsible unit, and consequently the monitoring of goals was carried out with a diversity of priority programs, projects, and activities that added this cost into the corresponding programmatic categories. From 2005 to 2006, federal public administration offices and entities reported information on meeting their goals on the basis of selected indicators.

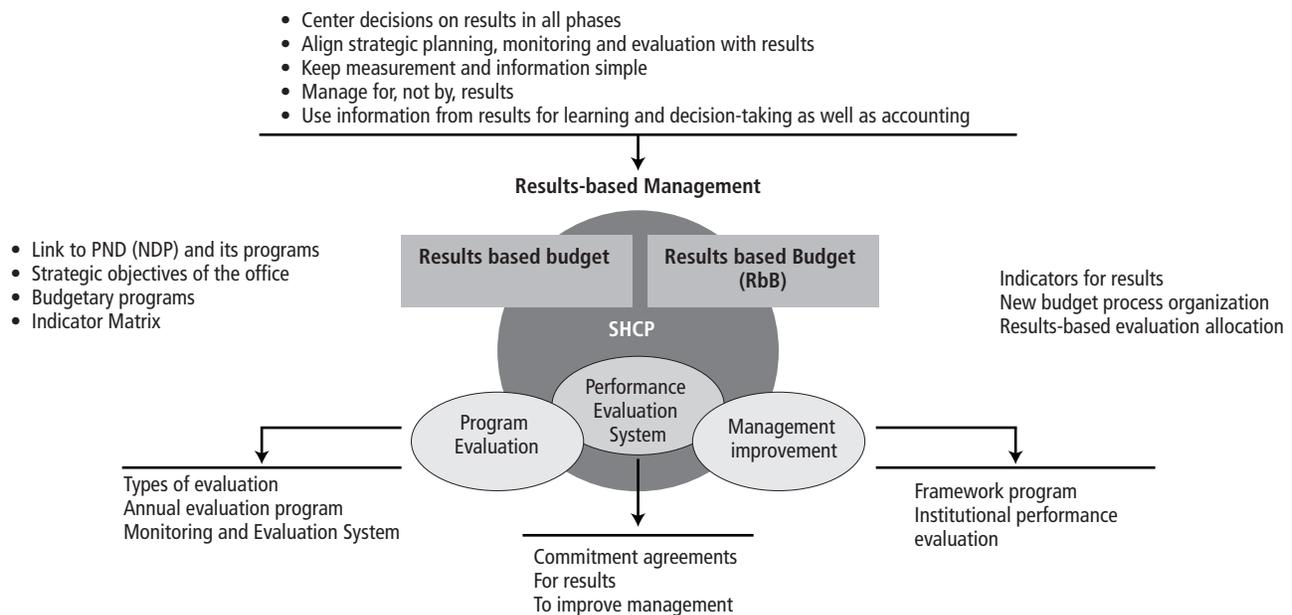
Performance evaluation system: results-based management strategy

With the leadership of Secretariat of Finance and Public Credit, it was determined that the programming-budgeting process would be carried out under the results-based management (RBM) strategy, which uses performance information to improve decision making, incorporates strategic planning tools, applies logical models, and monitors and evaluates results. The strategy focused on (1) concentrating decisions on results, and aligning strategic planning, programming, monitoring, and evaluating with results; (2) orienting information systems toward results; (3) focusing management on results; and (4) making use of information for learning, decision making, and accountability. The strategy was conceived as a strategy with the capacity to be used in all public entities.

At the same time, the performance evaluation system (PES) was conceptualized as “the set of methodological elements that allows for making an objective assessment of program performance, in line with principles focused on verifying the degree to which goals and objectives are being met, and based on strategic and management indicators that make it possible to determine the social impact of programs and projects.” The two main components are (1) the objective evaluation of federal programs, and (2) the ongoing evaluation of the functioning and results of public institutions. To strengthen this system, important legal and institutional measures were carried out, and the necessary support was provided to turn this system into a starting point for major administrative reform. Beginning in 2007, the implementation of this redefined system was promoted actively by the federal government, with support from the National Congress, and aimed at making the constitutional and legal reforms necessary for its implementation to be comprehensive. The participation of multilateral development banks, both the IRDB and the IDB, has been important. In June 2007, the Secretariat of Finance and Public Credit, representing the Mexican federal government, signed the Non-Reimbursable Technical Cooperation agreement (No. ATN/OC-10476-ME) with the IDB-

FIGURE 12:

Results-Based Management Strategy



Source: Author.

to carry out the implementation of a Results-Based Budget System (RBBS) in the federal public administration.

The RBM strategy to carry out the RBBS-PES was part of a much more comprehensive effort in the country, accompanied by a state vision that promoted efficient, effective, and accountable government. The Comprehensive Reform Initiative of the Public Treasury, presented by the federal executive branch to the National Congress on June 20, 2007, for its consideration, included measures for strengthening tax collection and greater coresponsibility through participation by state-level governments, as well as better and more transparent public spending. This initiative was approved on September 13 of that same year by the House of Representatives, and was given the name “Treasury Reform for the Most Disadvantaged.”

Behind this reform is the idea that taxpayers should have greater certainty that the taxes they pay are used well, and that they are applied to government programs and investment projects that offer tangible results to the population, as proposed in the 2007–12 National Development Plan (NDP), the 2008–12 National Program for Development Financing, and the Live Better strategy. The following are the most outstanding elements in the reform:

- To grant constitutional status to the obligation for the three levels of government (federal, state, and municipal) to ensure that the economic resources at their disposition are evaluated by technical entities created by the federal and state governments.
- To establish that federal economic resources used by states and municipalities will be subjected to the provisions in regulating laws, specifically the Federal Budget and Fiscal Responsibility Law, the Fiscal Coordination Law, and an assessment by technical entities. In other words, the use of resources will be subjected to stipulations for performance evaluations aimed at verifying the degree to which goals and objectives are being met, on the basis of strategic and management indicators, among other aspects.
- To establish the foundations for the uniform presentation of information on income and expenses, as well as assets, by the federal, state, municipal, and Mexico City governments. One of the main achievements is the General Law on Government Accounting of December 31, 2008. With the har-

monization of the government’s accounting system for the three levels of government, and earned value accounting records, there will be a solid source of information for strengthening budget allocation and accountability processes.

Results-based budgeting

Results-based budgeting (RBB) is a central element in the implementation of the PES. The process of developing the Federal Expenditure Budget (FEB) Project for fiscal year 2008 made adjustments to the programmatic structure. Two adjustments especially worth mentioning were the creation of the budgetary program to replace the institutional activity category, and the implementation of the objective-indicator-goal model. The RBB was defined as “the set of activities and tools that make it possible for the decisions involved in the budgetary process to systematically incorporate considerations regarding the obtained and expected results from the use of public resources, and that motivate government offices and entities to achieve such results, with the objective of improving the quality of public spending and accountability.”

Institutional framework

With the Federal Budget and Fiscal Responsibility Law, constitutional reforms of public spending and supervision, and the subsequent legal reforms in the area of the RBB-PES, an institutional framework was established to provide greater legal and economic certainty and transparency in the overall budgetary process. Simultaneously, a combined M&E system was established, with the necessary networking and coordination elements based on provisions currently in force, with the participation of the Secretariat of Finance and Public Credit, the Civil Service Secretariat, the National Council for the Evaluation of Social Development Policy (CONEVAL), and the House of Representatives, through the Federal Auditing Department.

At the same time, federal public administration offices and entities have the responsibility, through higher clerkships or the equivalent, to plan and evaluate budgetary programs and administrative units responsible for public policies and related federal programs. In accordance with the normative structure of the RBB-PES, regulators and those responsible for making expenditures collaborate. In practice, collaboration has been increasing as training programs

have advanced. In this collaboration between regulators and expenditure-makers, another factor has had a decisive impact—that is, the importance placed on performance evaluation in expenditure budget projects as an important element in budgetary decision making.

The PES's combined institutional design has important advantages. On the one hand, it makes it possible for the executive branch to exert decisive influence on programming by linking the NDP and its programs with the strategic objectives of government offices and entities, and with budgetary programs, and by adopting qualitative improvements in programming structures. It allows for considering results when making budgetary allocations; improving the management and quality of public spending; providing follow-up to budgetary management; improving policies, programs, and institutions; and strengthening the accountability process. In addition, it makes adequate external supervision possible, conducted by the Federal Superior Auditing Office, through the Special Performance Auditing Office, which, according to the Federal Budget and Fiscal Responsibility Law, covers the control and evaluation of federal public spending, among other tasks.

With the recent publication of a September 2009 decree that reforms, adds to, and repeals various provisions in the regulations for the Federal Budget and Fiscal Responsibility Law, PES coordination is enhanced through the creation of a permanent working group in evaluation and monitoring of budgetary performance, and the results of that performance. This group is composed of the Secretariat of Finance and Public Credit, the Civil Service Secretariat, the President's Office, and CONEVAL.

Implementation of the RBB-PES

Various actions in the implementation of the RBB-PES, carried out to date, make the following possible:

- Modifying the budgetary process to be able to identify the destination of public spending, and linking such spending with strategic objectives in national development planning. Budgetary programs were redefined to facilitate the measurement of results. The programmatic category of "Budgetary Program" was created to replace the previous "Priority Activity" category, which consisted of four generic modalities. The new programmatic category

makes it possible to identify the allocations and final destination of federal public spending, based on the classification of groups and modalities.

- Regulating the evaluation of federal programs. With the issuance of "General Guidelines for Evaluating Federal Programs of the Federal Public Administration," the Annual Evaluation Program (AEP) was created. Its objective is to refine performance evaluation as an instrument for gradual improvement in the quality of public spending.

This work made it necessary to establish the necessary types of evaluation as well as the development of a set of methodologies for performance evaluation that would facilitate the systematic, comparative evaluation of budgetary programs. It was also established that the evaluations of budgetary programs would be conducted by independent consultants. Also established was the requirement to make evaluation results public, making them available on the Web sites of the Secretariat of Finance and Public Credit, the Civil Service Secretariat, and CONEVAL. It was also necessary to

- Establish mechanisms for monitoring evaluations. Implemented in October 2008 were "Mechanisms for monitoring aspects susceptible to improvement, derived from external reports and evaluations of federal programs."
- Develop the necessary systems for monitoring the goals for indicators established in the FEB programs. The functionality of validation, review, definitive updating, and selection of budgetary program indicators was developed in the RBB-PES module included in the Secretariat of Finance and Public Credit's Web site.
- Train and educate public servants for participation in the RBB-PES implementation process, through long-distance training and regular courses on Logical Framework Methodology.

Matrix of Results-Based Indicators

Encouraging changes have been achieved through the policy of strengthening PES. For fiscal year 2008, 398 of the 1,183 budgetary programs had matrix of results-based indicators (MRIs), with a total of 3,290 indicators. The fed-

eral public administration offices and entities identified 265 indicators, corresponding to 217 budgetary programs, as most representative for measuring and assessing performance. For 2009, 512 of the 1,108 budgetary programs have MRIs, with a total of 4,634 indicators. Of these, 246 strategic indicators corresponding to 160 priority budgetary programs were selected. During the process of developing the FEB Project for the 2010 fiscal year, federal public administration offices and entities, in collaboration with the Secretariat of Finance and Public Credit, Civil Service Secretariat, and CONEVAL, reviewed and improved the MRIs, particularly those with 2009 selected indicators, to summarize and thereby achieve greater certainty in the information generated. MRIs were registered for 550 budgetary programs, with 4,422 indicators that will serve as the basis for the PES during the 2010 fiscal year.

Of these indicators, 294 were selected from 199 programs, considered to be the most representative of the federal public administration, both due to the amount and

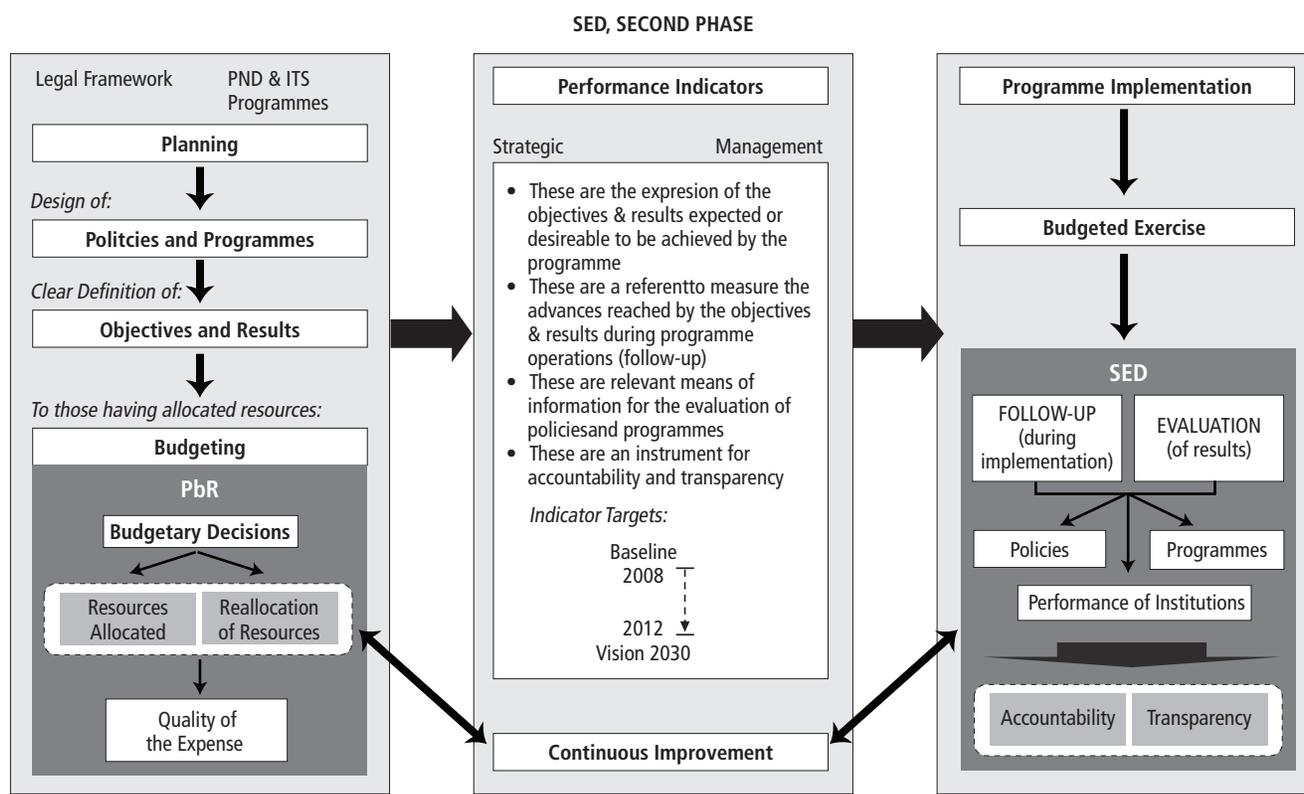
type of resources allocated, as well as the proposed results for the benefit of society. In addition, 294 strategic indicators that most adequately measure the objectives of these programs were identified and included in the 2010 FEB project. It is important to point out that the indicators selected correspond to the most relevant budgetary programs in federal public administration, to the degree that they directly or indirectly provide federal subsidies to target populations.

Annual Evaluation Program

The AEP brings together the evaluation proposals from the Finance and Public Credit and Civil Service Secretariats and CONEVAL, addressing the evaluation strategies used by each of them. In 2007, 2008, and 2009, a total of 126, 192, and 196 evaluation strategies have been used, respectively. While design evaluations dominated during 2007, specific performance evaluations began to take shape in 2008 and 2009, demonstrating

FIGURE 13:

Performance Evaluation System, Second Stage



Source: Author.

progress in meeting program goals and objectives on the basis of a summary of information contained in the PES and through an analysis of strategic and management indicators that made it possible to verify the results of the use of budgetary public resources. The AEP for 2010 has a strategy different from those used in previous years. The purpose of the 2010 AEP is to expand the scope of the evaluation of the federal public administration, by programming evaluations of areas not previously included. Examples are the Interior Secretariat (Area 4), the Foreign Relations Secretariat (Area 5), the Secretariat of Communications and Transportation (Area 9), and the Federal Attorney General's Office (Area 17). Also acquiring importance in the 2010 EAP is the strategic evaluation of social protection policies, microcredit policies, employment policies, and highway subsector policies.

Continuous improvement

In the context of the implementation of the RBB-PES, a series of improvements have been incorporated into the budgetary process. Together, these improvements make it possible to link planning, programming, budgeting, use, monitoring, evaluating, transparency, and accountability, exerting influence on the systematic feedback from programs and contributing to the analysis, discussion, and decision making regarding resources allocated to the programs. These improvements consist of the following:

- Planning: (1) alignment of budgetary program objectives with the objectives and indicators of the NDP and its programs; (2) Alignment of budgetary program objectives with the strategic objectives of the government office or entity.
- Programming: (1) defining the modality of budgetary programs; (2) integrating the budgetary program into the programmatic structure; (3) developing and updating the MRI for the budgetary program, including the creation and continuous improvement of strategic and management indicators.
- Budgeting: specific budgetary allocation for each budgetary program, bringing together the different expenditure-makers participating in each program.
- Use: linking the rules of operation for the budgetary program with the MRI.

- Monitoring: (1) monitoring budgetary programs, and their objectives, indicators and goals; (2) preparing reports on progress made in meeting the goals for budgetary program indicators.
- Evaluation: (1) systematic evaluations in accordance with the AEP, the results of which have improved the design of the budgetary program, and in some cases, its operation and impacts; (2) commitments for results and improvements in budgetary program management, based on aspects susceptible to improvement derived from evaluations.
- Transparency and accountability: public accounts that bring together information on the monitoring of goals for indicators in each budgetary program.
- Impact: for information on program performance, generated by M&E indicators, to have an impact on budgetary allocations, it is necessary to make ongoing efforts in consolidating the quality of the information incorporated into the PES, expanding the use of the MRI in analyzing program performance, and rooting results-based management in the daily activities of public servants, in the annual budget proposal, and in the program, and project review process.

Performance evaluation at the local level

At the local level, during fiscal year 2009, the initial performance evaluations observed significant advances in the implementation of strategic and management indicators. They were defined for General Area 33, Federal Contributions for States and Municipalities, allocated in the context of the existing schemes for intergovernmental coordination and with the participation of coordinating offices for each fund. A total of 29 indicators with varying frequencies were agreed on for the eight funds included in the area. Of these, 18 are strategic and 11 are management indicators. State and municipal governments report progress in the indicators through the information system designed by the Secretariat of Finance and Public Credit. This progress is reported to the National Congress through quarterly reports. In the near future, the MRIs and performance indicators for federal public resources transferred to local governments in General Area 23, Wage and Economic Provisions, will be defined for each of the funds focused on regional development. The strategy for implementing the PES at the local

level requires each state (1) to reform its constitution and local budgeting and accounting laws, (2) to make progress in implementing methodologies for results measuring using its own resources, (3) to strengthen the technical capacities and independence of the supervision entities of local congresses, (4) to evaluate the results from the use of federal public resources received on the basis of strategic and management indicators, and (5) to improve accountability.

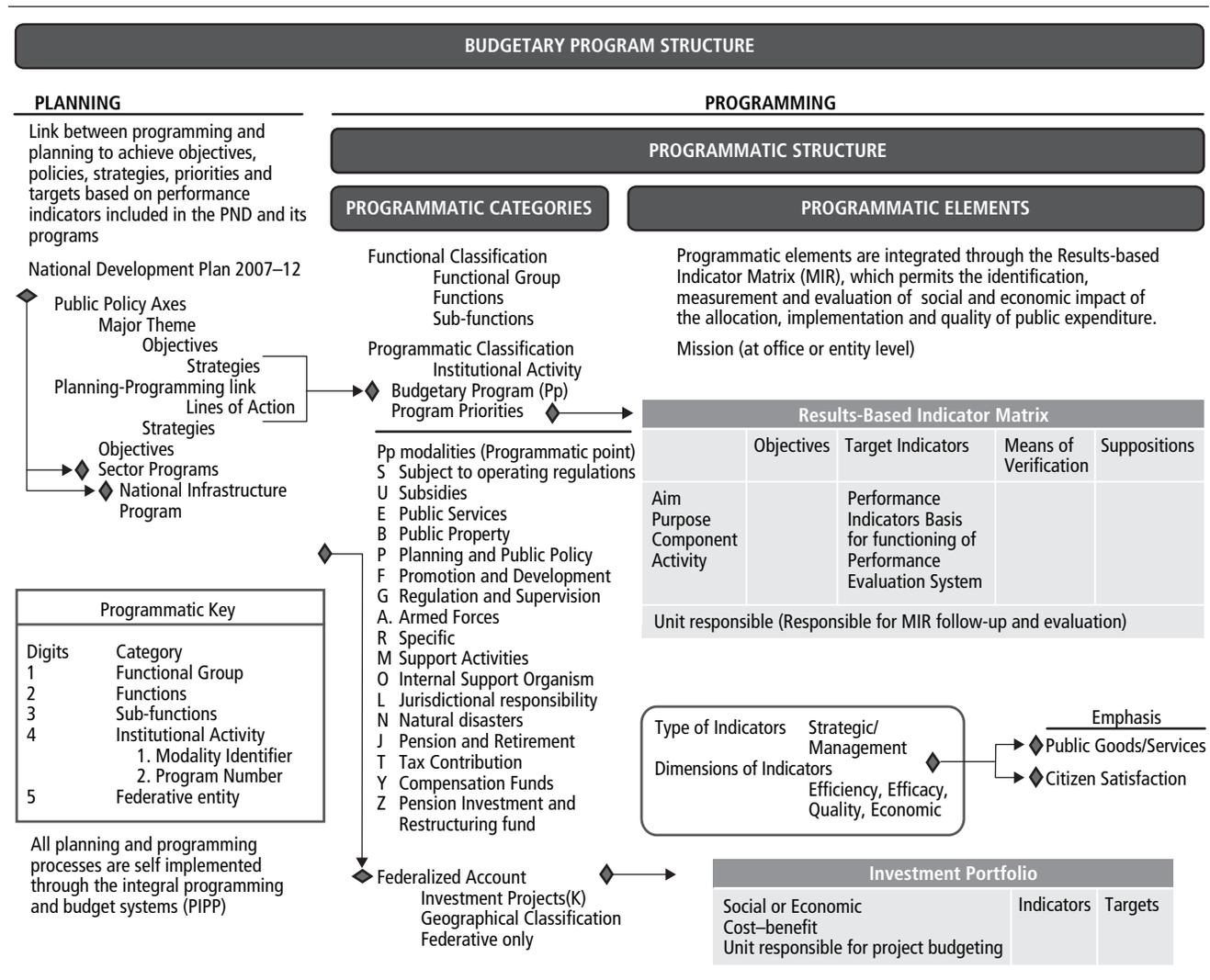
A critical assessment of performance evaluation in Mexico

To strengthen the PES, it is necessary to overcome a series of challenges, all of which merit examination. Without

question, the PES increasingly is considered to be a vehicle for change, as it exerts influence on parliamentary debates on the efficiency and effectiveness of public spending. Similarly, it has become important in relation to the welfare economy, the allocation of budgetary resources over time, and the very structure of economic policy, in terms of its capacity to implement economic policy to achieve the objectives proposed. It has become important as an element for evaluating public policies and institutions. Because the RBB-PES is relatively new, it is necessary to strengthen its institutional design, broaden its coverage in terms of programs and projects, and expand the context in which it is applied to include other branches and levels of government.

FIGURE 14:

Budget Programming Structure



Source: Author.

Various measures elevated the PES to the level of the 2010 FEB Project. This guaranteed the PES's sustainability. The government offices and entities responsible for programs must carry out a process of continuous improvement and must maintain the MRIs updated. To accomplish this, it is necessary to consider the advances and results obtained from monitoring whether program goals are met, whether evaluations are conducted in line with the AEP and whether the criteria defined by the Secretariat of Finance and Public Credit and CONEVAL are met. In addition, the MRIs should consider, in the programs in which it is necessary, the transversal focuses of equality in relation to gender, youth, disability, and ethnicity.

Improvements in the MRIs should align their objectives with the strategic objectives of government offices and entities, and with the strategic objectives of the NDP and the programs derived from that plan. The rules of operation should be consistent with the information in the MRIs, in response to their improvement and updating. Government offices and entities must develop a work plan for following up on the primary results from the evaluations conducted, and must include aspects that could be improved in the designing of public policies and the corresponding programs. Commitments will be formalized through specific instruments. The advances and results achieved through the PES will be reported and published in the terms of applicable provisions so that the information obtained from the monitoring of commitments on improvements, corresponding to the fiscal years 2007 to 2010, will be taken into account as part of a gradual, progressive process during 2010 and for the subsequent budgetary processes.

PESs are in a constant state of change. Challenges arise as their application is extended throughout public institutions, causing changes in the behaviors of institutional agents, and generating acceptance, indifference, or rejection. The main challenge lies in finding a way to motivate behavior favorable to the PES. A change in traditions toward conducting evaluations requires not only the training and dissemination that have been carried out on an ongoing basis, but also the generation of a scheme of incentives that rewards the institutions that operate programs with better results and allows them greater degrees of freedom.

To enhance the probability of success and permanence, it is vitally important for the RBB-PES to be accompanied by an adequate system of incentives and by the dissemina-

tion of results obtained with this new focus. To that end, it is vital to establish the necessary incentives to promote the fulfillment of strategic objectives in programs, policies, and institutions with the objective of innovating and modernizing how organizations work, and the budgeting and spending processes. Ultimately, the information obtained from monitoring the program indicator goals and the evaluation's results should be included in budgetary decisions and in program management.

The RBB-PES has achieved considerable success, particularly in modifying the normative framework and in generating and disseminating information on public management performance. Nevertheless, this success has been achieved due to the obligatory nature of the legal framework. Institutions and functionaries have not yet taken ownership of the entire system. It will be possible for the PES to make significant steps forward if a positive scheme of incentives is incorporated for encouraging the use of performance information in decision making, and if the rating of public servants can be incorporated into program performance. Institutional incentives also should be established, to allow public entities to benefit from the programs they carry out, through the savings obtained from applying efficiency and cost-effectiveness measures.

At the regional level, it is necessary to move forward with parallel actions that strengthen financial systems. Actions will align accounting systems in states and municipalities with those in the federal government. On December 31, 2010, at the latest, the federal government and states must have developed indicators for measuring physical-financial advances associated with federal resources and with different points in time during the accounting process. It is estimated that it will be possible to have implemented a uniform accounting system by 2011, and that by 2012 all records will be maintained under the new system.

The PES must be appropriate for its users. It is necessary to encourage the main actors in the budgetary process to use performance information, particularly what is generated in the framework of the RBB-PES, and to improve the clarity and transparency of the results presented to society. The challenge is to convince users of the positive aspects of using the system, and above all, in instilling the notion that a budgetary program is more valuable when the genetic code characterizing it is fully identified. Lastly, the RBB-PES is a fundamental part of the effort to build and

establish a budgetary tradition among the population, so they can be aware of the use, destination, and performance of public resources in a clear, simple, and direct manner, and so they can participate in the way that such resources are allocated.

Agustín Caso

*Deputy Director General, Performance Evaluation System,
Ministry of Finance and Public Credit, Mexico*

The Creation of an Evaluation System of Higher Middle Education in Mexico

Introduction

Information and knowledge are prerequisites for well-founded decision making, which in turn are necessary for continuous improvement or progress. In the public arena, information and analysis aid us in obtaining diagnoses that identify the main problems to be solved. Identification of the problem and its probable causes provides us with the necessary elements to design public policy programs or interventions. When time and the nature of the problem permit, implementation of the program (or intervention) should be preceded by a pilot program to discern its impact. Finally, once the public policy is implemented, it should be subjected to continuous monitoring and periodic measurement of the impact that the intervention had.

The three components necessary for the design and implementation of effective public policy (diagnosis, intervention design, and impact measurement) can be seen as part of an integrated evaluation program. Without one of these elements, the public policy decision runs the risk of not having the expected impact. Interventions without impact not only fail in their attempt to resolve the problem for which they were designed, but also waste limited public resources. Public policies, whose design and implementation are based on integrated evaluation, resolve a specific problem efficiently and reveal the effectiveness of the use of public resources, thereby increasing support for the intervention.

Today, Mexico does not have in place the institutionalized mechanisms or processes to evaluate public policies oriented toward the improvement of higher middle education (*Educación Media Superior*, EMS). Two public institutions evaluate public education; however, neither of these sets

forth as an objective the evaluation of higher middle-level education policies. The objective of the Planning and Evaluation of Educational Policy Unit (UPEPE) is “to develop valid and reliable evaluations for supplying pertinent and timely information to federal and state authorities.” However, the evaluations carried out by the UPEPE are focused on basic education (primary and secondary); thus, they do not evaluate education policies at the higher middle level. On the other hand, the National Institute of Evaluation of Education (INEE is in charge of evaluating the education system as a whole, without specifically evaluating HME public policies). As will be seen later, the UPEPE and the INEE are a basic part of the Higher Middle Education Evaluation System (HMEES). Additionally, the National Evaluation Council of Social Development Policy (CONEVAL) has as its objective to establish the rules for and coordinate the evaluation of social development policies and programs.

The creation of an HMEES (*Sistema de Evaluación de la Educación Media Superior*, SEEMS) comprises the first effort in this direction. SEEMS was created to comply with two general objectives:

- Improve the quality of EMS by means of decision making based on diagnosis, design, and measurement
- Inform society, through accountability, of the status maintained by the EMS and the manner in which EMS policies, programs, and interventions exert an impact on the quality of this education level

The creation of SEEMS arises within the framework of the Integrated Reform of Higher Middle Education (IRHME). IRHME sets forth an exhaustive review of the EMS system, including education content, the installment of a mechanism to assign and evaluate the education centers’ directors, and the design of an instrument to measure academic achievement, among others. Therefore, the entry of IRHME is a unique opportunity to carry out a complete diagnosis of EMS, to evaluate the impact of policies set in motion as part of the reform, and to launch pilot programs for designing and evaluating future interventions.

Evaluation within the IRHME framework

EMS in Mexico has been characterized by a large number of subsystems with different study plans, quality standards,

entrance and exit criteria, objectives, and goals. Consequently, there have been high dropout rates, poor results in cognitive achievement tests, and a weak link with the labor market.

The IRHME was put into action in 2007.²⁵ The objective of IRHME is to endow identity and order on a segment of the education system in Mexico that has been characterized by the absence of common policies in all subsystems and a lack of regulation in the supply. The manner by which IRHME seeks to achieve this ambitious objective is through the creation of a National Baccalaureate System (NBS), which defines IRHME entrance and exit criteria, standardized quality standards, and study plans that are common to all subsystems without sacrificing their diversity.

To pave the way for the creation of the NBS, federal, state, and other authorities defined the profile of the higher middle school graduate. With this in mind, they identified the minimal academic resources to guarantee that all NBS students, no matter the state, subsystem, or type of schooling, graduate with the necessary tools to face the challenges of the twenty-first century (the exit profile). The academic resources necessary include the education content, the education center's physical plant, the profile of the teachers,

and the transparency of the academic management of the center's director, among others. IRHME conceptual planning and practical implementation rests on four basic pillars (see Figure 15).

Education content

This pillar includes definition and construction of a competency-based common curricular framework. The common curricular framework allows for free movement of students among different subsystems without the need to begin studies from zero, as occurred until recently. Additionally, it includes the redefinition of a memorization to a competency-based teaching paradigm model.

Diverse supply options

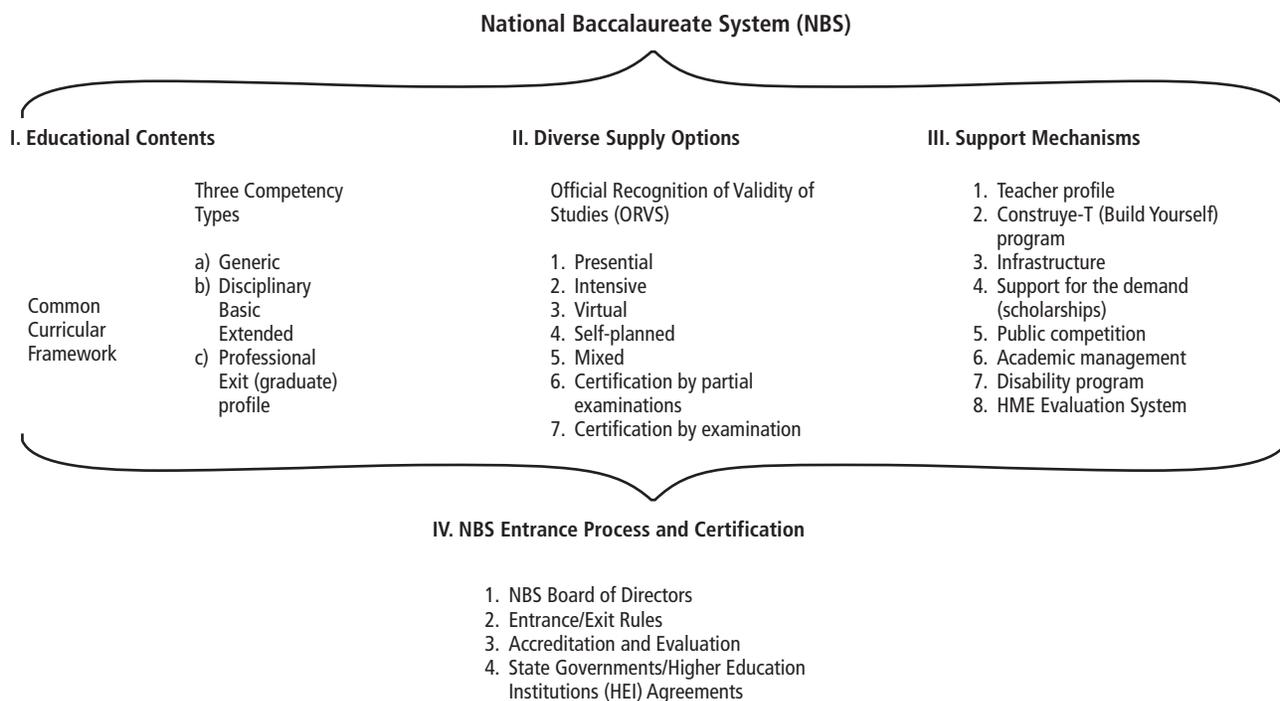
This pillar includes definition and recognition of the different supply options of EMS.

Operating mechanisms

This pillar includes instrumentation of programs and interventions for the construction of the NBS. These programs or interventions represent the operative part of IRHME,

FIGURE 15:

National Baccalaureate System



Source: Author.

that is, the pathway to achieve the NBS objectives. The programs or interventions included, among others, the following: the Teacher Training Program with the adequate profile, the Academic Infrastructure Program, the Director Training Program, the EMS Scholarship Program (support for demand), the Academic Management System, and the creation of the EMS Evaluation System.

NBS entrance and NBS certification

This pillar establishes clear rules of NBS entrance and certification, and national certification for all NBS graduates. The 2009–10 academic year marked the beginning of the NBS, following two years of work on the definition of the common curricular framework, NBS entrance and exit rules, and implementation of the necessary programs and interventions. In 2009, all centers demonstrating that they had participated in the programs and interventions required to guarantee the graduate profile could request their entrance into the NBS. As agreed to by all the education authorities, entrance into the NBS will be determined by independent evaluator organisms based on criteria pre-established by the board of directors.

The Higher Middle Education Evaluation System

SEEMS plays a fundamental role in IHRME in evaluating all policies and interventions carried out within the framework of reform (“operating mechanisms”). Additionally, the evaluation system designs and evaluates pilot programs for future interventions that contribute to IHRME objectives.

SEEMS includes four evaluation spheres according to their aggregation level. The four levels are (1) system and subsystems, (2) school or center, (3) teachers, and (4) students. Each of these areas of influence plays a basic role in the learning process; thus, it is necessary to possess a diagnosis and a referent for each. The diagnosis of each area together with the referent (or point of comparison) permits identification of the sphere in which EMS presents the greatest lag, that is, where we require redoubling of effort. For example, in the case of the NBS student area, the referent (the “should be”) is defined by the exit profile of the higher middle graduate. Having a comparison between diagnosis and referent for each ambit also permits us to evaluate the efficacy of the policies and interventions defined by IHRME and based on their capacity to focus resources on areas with the greatest lag.

Programs within IHRME

As previously mentioned, the operation of IHRME is accomplished through programs or interventions that seek to endow the centers belonging to the NBS with the academic resources necessary to guarantee the graduate’s exit profile. The programs or interventions can be divided into the following two broad headings: (1) programs that are in process, and (2) pilot programs.

Programs and interventions in process

- Infrastructure Programs: drive the growth of the educational supply in EMS; endow and strengthen the existing infrastructure.
- Teacher Training Program: develop the teaching competencies that are required for achieving the NBS exit (graduate) profile.
- EMS Scholarship Program: broaden education opportunities between students residing in a household that falls under the asset poverty line.
- Program of Attention to Persons with a Disability: ensure the equality of educational opportunities among persons with special needs.
- *Construye-T* (Build Yourself) Program: prevent risk behaviors among EMS youth and promote the construction of a life project.
- The *Impulsate con Inglés* (Drive Yourself with English) and Computer Program: promote the use of new technologies and the tools necessary in a globalized world.
- School Management System (SMS): foster transparency and accountability among EMS center directors.
- Public Competition for Assigning Center Directors: develop a merit-based employment-assignment system and thereby foster transparency in the use of resources.

Pilot-phase programs

- *Preparate* (Prepare Yourself): broadens EMS coverage by means of distance learning (online).
- Incentives-Alignment (IAL) Program to Strengthen Academic Achievement: identify a better incentives scheme to improve academic performance.
- Program on Perceptions and School Dropout: test the effectiveness of reducing school dropout

in an intervention whose sole component includes information on the benefits associated with schooling.

Advances in SEEMS consolidation

In 2009, the construction of SEEMS began with the participation of the National Institute of Evaluation of Education (NIEE), the Planning and Evaluation of Educational Policies Unit of the Secretariat of Public Education (*Secretaría de Educación Pública*, SEP), the National Center for the Evaluation of Higher Education (CENEVAL), CONEVAL, the Latin American Faculty of Social Sciences (FLACSO), and the Under Secretariat of Higher Middle Education of the SEP (*Subsecretaría de Educación Media Superior*, SEMS). SEEMS members designated a group of experts to coordinate the tasks of diagnosis and evaluation within each SEEMS area. In parallel, some of the programs or interventions are in the process of evaluation, in collaboration with university academics of international prestige, and even with multilateral organisms.

Evaluation of HME scholarship programs

Evaluation of the scholarship program is performed by Orazio Attanasio and Costas Meghir of the University College London, who work in evaluation issues that utilize an experimental design. The HME Scholarship Program (HMESP) consists of a monthly monetary transfer to all EMS students who comply with all of the requirements established by the program's rules of operation as follows:

- Mexican citizenship
- Enrolled in a participating public HME institution
- Under 21 years of age
- Live in a household that falls under the asset-based poverty line
- Request government aid by filling out an application form²⁶

The scholarship program's objectives are to increase absorption (the transition from secondary school to the baccalaureate level), reduce school dropout, and increase levels of academic achievement. In statistical terms, $D_{i,t}$ is defined in the following manner:

$$D_{i,t} = \begin{cases} 1 \Leftrightarrow \text{the individual "i" decides to leave the} \\ \text{EMs in time "t"} \\ 0 \Leftrightarrow \text{the individual "i" decides not to leave the} \\ \text{EMs in time "t"} \end{cases}$$

We can define the probability of dropout, $\Pr(D)$, as a linear function of a variable, TB_i , which indicates that individual i has an HME scholarship, and a series of controls \mathbf{X} . Therefore:

$$\Pr(D_i) = \alpha + \beta\mathbf{X} + \delta_1 TB_i + \varepsilon_i \quad (1)$$

Where β and \mathbf{X} are a vector and a parameter matrix, and independent variables, respectively; δ_1 is the parameter that measures the effect of the HME scholarship on the probability of dropout. That is, δ_1 is our parameter of interest. And last, ε_i is a randomized term with a mean of zero and a distribution that asymptotically approximates normal.

The validity of model (1) for obtaining the impact of the scholarships on dropout depends on the manner in which these have been assigned among treatment groups—those who received a scholarship—and the control group—those who did not receive a scholarship. If the treatment (EMS scholarships) was assigned randomly, differences between the two groups (treatment and control) can be attributed to the scholarship program; thus, we will know its impact on dropout, absorption, and academic achievement.

A questionnaire with all the necessary information to estimate the model (1) was designed to evaluate the program. In 2010, the National Public Health Institute (NPHI) survey will be applied to 3,500 households of youths belonging to the control group and the identical number of households in the treatment group. Final results will be available at the end of 2010.

Evaluation of the Construye-T (Build Yourself) program

The *Construye-T* program is made up of the following six special dimensions for the development, protection, and construction of life projects of youths in EMS:

- Knowing one's self: to develop the creative capacity, as well as the expression of one's inner world, to find answers when confronted by obstacles and

risks, to overcome fear in the face of new paths, and to recognize affections and emotions.

- School and family: to improve the relationships among male and female students and their families, and the teachers. This consists of two subdimensions: school environment, and family.
- Healthy life: to strengthen the resources and individual and social elements of protection for promoting health and obtaining a better quality of life. It embodies three subdimensions: consumption and health, consumption of addictive substances, and sexuality.
- Youth participation: to educate male and female students as citizens and to link their performance in school and in the community, for the benefit of the environment.
- Nonviolence: to offer elements for the development of a culture that promotes inclusion, solidarity, and cooperation for finding peaceful solutions to conflicts.
- Life project: to foster the development of competencies in male and female students for them to define and conduct their future, in such a way that they become the owners of their lives and they discover their values, capacities, potentialities, and the sense of their existence.

Evaluation of the *Construye-T* program is carried out by the United Nations Educational, Scientific and Cultural Organization (UNESCO) in collaboration with the International Institute of Education Planning in Buenos Aires. Variables of interest include terminal efficiency (the percentage of first-year EMS students who finish on schedule) and the risk factors to which youths pursuing EMS are exposed.

To carry out the evaluation, in 2007, the year in which the federal-subsystem centers program began operation, the SEMS and the NPHI implemented the first survey “Exclusion, Intolerance, and Violence in the Higher Middle Education.” At the end of 2009, a second survey will be conducted, with the objective of analyzing whether there progress has been made in federal centers compared with other subsystems. Evaluation equipment will analyze the difference in terminal efficacy and risk behaviors throughout time and between federal subsystems and others, that

is, the difference between the differences. If $ET_{f,t}$ and $ET_{nf,t}$ are defined as the terminal efficiency between federal and nonfederal centers in time t , respectively, then the impact of the program can be measured by means of the following equation:

$$\text{Impact} = [(ET_{f,t+1} - ET_{f,t}) - (ET_{nf,t+1} - ET_{nf,t})] \\ | (X_{f,t+1}, X_{f,t}, X_{nf,t+1}, X_{nf,t}) \quad (2)$$

where X is a vector of observable characteristics at the subsystem level. The same model can be estimated for the case of risk behaviors, with the difference that impact would be rendered at the student level, and not at the center level, as in the case of terminal efficiency.

The *Prepárate (Prepare Yourself)* pilot program

Online education in Mexico has been directly mainly toward the adult population and in many cases, it is already inserted in the employment area. The *Prepárate* pilot program attempts to determine whether providing EMS online to persons under 17 years of age can be an alternative to face-to-face education, that is, whether online education achieves at least the same learning, failures, and dropout levels. UNESCO is in charge of evaluation of this pilot program, whose four basic objectives are as follows:

- Given that distinct academic institutions with different programs of study participate in the program, attempts have been made to systematize these academic variants of program applications, while citing the common and different components and determining their importance. In this respect, the strengths and weaknesses of each variant will be identified, highlighting the lessons learned by the institutions involved and analyzing the program management design.
- To conduct a qualitative evaluation by means of workshops or focus groups in terms of student permanence and failure
- To carry out an evaluation on learning achievements through employment of the standardized Higher Middle National Evaluation of Academic Achievement in Academic Centers (ENLACE) test in which the face-to-face and online modalities of the institutions will be compared. To render these groups

comparable, the students selected in both groups will have similar socioeconomic characteristics.

- Finally, the development of abilities, practices, and study habits in relation to information technologies and communication will be evaluated. This evaluation will be performed through focus groups.

Incentives-alignment program

The objective of the IAL program is to identify the best incentives scheme (monetary) to improve student achievement in mathematics. This pilot program is the result of collaboration among the Inter-American Development Bank (IDB), the University of Pennsylvania,²⁷ and SHME. The IAL program operates in 88 federal technical-baccalaureate higher middle education centers, in which nearly 50,000 students form part of this academic achievement-conditioned transfer program. The 88 centers forming part of this program belong to the following four groups:

- T1: students receive monetary transfers based on their achievement in a standardized mathematics test.
- T2: teachers receive monetary transfers based on their students' achievement in a standardized mathematics test.
- T3: students, teachers, and directors receive monetary transfers in terms of their individual and group achievement in a standardized mathematics test (incentives-alignment).
- C: No monetary transfer is received.

The centers were assigned to each of four groups randomly, which guarantees that the differences in mathematical achievement between the control group and each of the treatment groups can be attributed to the incentives scheme. In addition, a pre- and post-course questionnaire is applied every year to control for changes in observables that could be treatment related (incentives).

The mathematics test with which the achievement of the students is evaluated was designed especially for this program by CENEVAL. Students, teachers, and directors receive the monetary transfer in a sole annual exhibition, in terms of achievement in the standardized test. The results of this important pilot program will be announced at the end of the 2010–11 school year.

Although the program seeks to increase student learning, it will analyze the impact on some education indicators of intermediate-results, such as the dropout rate, transition rates among school grades, terminal efficiency, university attendance rates, and cognitive achievement in mathematics.

Perceptions and school dropout program

From the viewpoint of economic rationality, the decision to drop out of school is determined by (1) the salary that the student can obtain in the labor market (without having finished baccalaureate studies), (2) the student's *perception* of the returns (or economic earnings) associated with education, and (3) the relative importance that the student assigns to present versus future remuneration. Thus, an increase in remuneration in workers without finishing baccalaureate studies will tend to increase school dropout in higher middle levels; likewise, an increase in the *perception* of higher middle students on the returns of education will tend to reduce dropout.

In the absence of restrictions in credit markets, that is, assuming that students have the opportunity to finance their studies, higher middle students will drop out when the labor-market salary is higher than the *perception* of the future income that they could acquire if they finished their baccalaureate studies. Therefore, part of the dropout problem could be associated with an erroneous perception, that is, with imperfect information. On the other hand, within a context of imperfect credit markets, school dropout can obey economic restrictions, which push the student into opting for present versus future consumption (for example, if students reside in households below the poverty line, their capacities and preferences for investing in their futures will be limited). This economic restriction can be eased by providing scholarships for poverty-level students.

The task of this project is to identify the role played by erroneous perception (or imperfect information) in the decision to drop out of school. We know that, in the case of poverty, economic restrictions can be more important than lack of information; thus, we differentiate between the impact exerted by erroneous perception on situations with and without budgetary restrictions (students with and without scholarships to study at the higher middle level). The methodology is divided into two stages:

- Learn the *perception* of students concerning the returns of education and the nonmonetary benefits associated with this, and comparing these perceptions with the returns observed.
- Carry out a later experiment, in which a group of students is provided with information on returns—monetary as well as nonmonetary—in education (measurement with the official information from the “National Survey of Occupation and Employment” in Mexico).

In the case of EMS, in which the dropout problem is concentrated on the first year of studies, the experiment will be carried out with first-entrance students. From among the universe of first-time students, two treatment groups and one control group will be created:

- TM1: receives information on the monetary and nonmonetary benefits related to EMS.
- TM2: receives information on the monetary and nonmonetary benefits related to EMS and a SHME scholarship.
- CM: does not receive information on the returns of education or on the SHME scholarship, but is given a survey identical to that administered to TM1 and TM2.

In statistical terms, D_{it} is defined as follows:

$$D_{it} = \begin{cases} 1 \Leftrightarrow \text{the individual “i” decides to leave the} \\ \text{EMs in time “t”} \\ 0 \Leftrightarrow \text{the individual “i” decides not to leave the} \\ \text{EMs in time “t”} \end{cases}$$

We can define the probability of dropping out, $\Pr(D)$, as a linear function of treatments TM1 and TM2, and a series of controls (\mathbf{X}). Therefore:

$$\Pr(D_i) = \alpha + \beta\mathbf{X} + \gamma_1\text{TM1}_i + \gamma_2\text{TM2}_i + \eta_i \quad (3)$$

where β and \mathbf{X} are vector and matrix of parameters and independent variables, respectively; γ_1 and γ_2 are parameters that measure the effect of information on the monetary and nonmonetary returns of education (TM1) and the effect of information in addition to a higher middle scholarship

(TM2) on the probability of dropping out or not entering higher education. That is, these are our parameters of interest. Last, η_i is a random term with a mean of zero and a distribution that asymptotically approximates normal.

Model (1) in combination with evaluation of higher middle scholarships will permit the evaluation of whether the presence of the two treatments exerts effects that are above those of the impacts that would be obtained in isolated fashion. In other words, the impact of dropping out on students who receive the scholarship and in addition receive information on the returns of education is greater than the sum of the impacts of the treatments:

$$H_0: \gamma_2 > \gamma_1 + \delta_3$$

where δ_1 is the isolated impact of the higher middle school scholarships on the probability of dropping out (see equation (1)).

The program will follow up students in treatment and control groups within one year of their receiving information to determine the impact that information on the returns of education had on school dropout. In addition, SHME will apply a standardized test to students who will participate in a pilot test to analyze whether information exerts an impact on academic achievement (possibly through the effort involved).

The challenges and the future agenda

During the initial stage of SEEMS, the challenge was that of integrating in a single conceptual document the ideas of a large and heterogeneous workgroup made up of diverse governmental and nongovernmental institutions. On the other hand, the benefit of having a working group of a number of academic and evaluation institutions is that the SEEMS agenda can complement other work plans under the heading of monitoring and evaluation. The remaining advantage of having a heterogeneous working group is that SEEMS can be nourished by the experience and information from other systems or institutions with greater experience. Additionally, possessing a workgroup composed of institutions within and outside of the federal government lends credibility to the results.

Another important challenge is that of finding an institutional mechanism by which SEEMS results (and those of other evaluation systems) are taken into consideration in

decision making. The fact that the under secretariat is the main promoter and a fundamental part of the SEEMS has increased the interaction between evaluation results and decision making. The true challenge is to create an institutional mechanism that guarantees that the evaluations and decision-making binomial is preserved in future administrations.

Perhaps we should begin creating suitable conditions to promote this link. At the present time, institutional design in the exercise of federal government expenditure in Mexico—and in many other parts of the world—leaves little space for rigorous evaluation-based decision making.

The work agenda that we have outlined is ambitious and transcends the present administration, and even the Integrated Higher Middle Education Reform. For the first time in the history of Mexico, we will have EMS with identity, clear, and common goals for all of the subsystems, and even its own system of evaluation in which cutting-edge interactions and decision making arise. We expect that the EMS results we will obtain during this administration will be the main promoter of the culture of evaluation in Mexico.

Rafael de Hoyos

Advisor Coordinator, Under-Secretariat of Higher Middle Education, Secretariat of Public Education (SEP), Mexico

Latin American Monitoring and Evaluation Systems and Their Governance

Introduction

Since their emergence in the early 1990s in Latin America, institutional arrangements adopted to install the functions of monitoring and evaluation (M&E) in public management results at the national level have been marked by profound transformations. Currently, there is a tendency toward convergence, which is likely associated with the increasing consolidation of a tradition of results-based management. However, there are some remnants from two types of different arrangements, which in their origins were linked to two different focuses with respect to the functions to be fulfilled by M&E.

Along these lines, it is clear that the countries that decided early on to institutionalize M&E in their public sectors did so by adopting one of two different paths. The first

path, the “budget model,” essentially seeks to use M&E to improve decisions for budgetary allocations and has therefore been promoted by treasury ministries. The systems originally implemented, in Chile and Mexico, for example, illustrate this model, which comes from the hegemonic current of thought on the modernization of public management, originating in the United States and Great Britain in the late 1970s. In the case of Mexico, in fact, a Program for the Modernization of Federal Public Administration (*Programa de Modernización de la Administración Pública*, PROMAP) was implemented in 1995 (PROMAP 1995–2000), and one of its main focuses was to initiate actions to evaluate its results and pave the way for linking these results to the resources allocated.

The second path, the “planning model,” in contrast, can be found in those countries that had decided to adopt M&E to improve political decisions at high levels of government and improve accountability. In this sense, it is no accident that the institutional designs guided by this type of motivation are focused on articulation with national development plans and are inspired by ministries of planning. Colombia’s National System for Performance Management and Evaluation (SINERGIA) and Costa Rica’s National Evaluation System (NES), in their origins, illustrate this type of model. Colombia, in fact, was the first country in Latin America to propose a design for a system to evaluate the management and results of public administration. It accomplished this through the Constitution of 1991, in which this task is assigned to the National Planning Department. In 1994 SINERGIA’s initial outline was disseminated, and between 1994 and 1999 it was validated by the World Bank and was promoted as a model to be followed. In Costa Rica, the National Development Plan (NDP) for 1994–98 promoted a new organizational tradition within the public sector, with an orientation toward results and an emphasis on making public functionaries responsible. It proposed the creation of a system for evaluating management as one of its fundamental pillars. The experiences of Great Britain and Colombia were used as the primary referents for NES’s conceptualization.

Brazil is the only country in which the original institutional design differed from the two models; it represented a hybrid model that would later tend to prevail. Its experience in results-based management began in 1996 within the framework of the 1996–99 Multiyear Plan although it

was defined by 42 strategic programs. In the Multiyear Plan for 2000–03, with the support of the World Bank and with Canada as a model, the following five elements were established as the focus for the public management model to be implemented:

- Organization by programs
- Strengthening of a strategic orientation
- Selectivity
- Management
- Evaluation

An important aspect of the Brazilian case was the integration, within the Multiyear Plan management methodology, of the teams from the areas of government planning and government budgeting. Thus, it appears to be possible to address one of the primary conflicts characterizing M&E institutional designs, although it was not totally resolved in this country. Current tendencies point to a search for solutions along these lines, at least partly explaining the reasons why emerging convergences appear to surpass the original models and bring greater governance to M&E systems. Some advances in this direction will be described and then pending challenges will be identified.

Some advances in governance

Toward shared leaderships

In countries in which M&E systems have been instituted under the jurisdiction of planning ministries, competition with treasury ministries has tended to limit the scope and effects of the systems. Evidence, however, suggests the possibility of building shared leaderships, although certainly not without some tensions.

In the case of Costa Rica, the trigger for possible coordination between the National Planning and Economic Policy Ministry (*Ministerio de Planificación Nacional y Política Económica*, MIDEPLAN) and the Treasury Ministry was the Law on Financial Administration and Public Budgets that went into effect in 2001. This law stipulated that both government entities, with prior consultation with the National Comptroller General, must issue on an annual basis the technical and methodological guidelines for formulating annual operational plans (AOPs) for all institutions. Since the law establishes that AOPs must be subject to the

priorities defined in NDPs, the Treasury Ministry and the National Comptroller General demand—as a requirement for the assessment and approval of the annual budget proposal—a resolution for linking the AOPs to the NDP issued by the MIDEPLAN.

Currently, some progress—at least expressed as a formal connection between the MIDEPLAN and the Treasury Ministry—is evident in the “Technical and Methodological Guidelines for Strategic Sector-based and Institutional Programming and Sector-based Monitoring and Evaluation,” developed jointly by the two entities in April 2009.

A key impetus for the construction of shared leaderships has been offered by the creation of joint entities, sometimes having important decision-making powers. In Costa Rica, an Inter-institutional Technical Commission has been in place since 2002. It functions as an arena for coordinating actions between these three entities. In this context, technical and methodological guidelines for developing AOPs are formulated and issued.

In Colombia, also in 2002, the Intersectoral Committee for Results-Based Management and Evaluation (ICRME) was created as an arena for technical coordination and a connection between evaluation and budgetary programming. It is composed of the Presidency, the National Planning Department, and the Treasury Ministry, in addition to those invited by each. The boost given to ICRME is evident in the most recent document CONPES 3515 (2008) published by the National Economic and Social Policy Council (*Consejo Nacional de Política Económica y Social*, CONPES), which strengthens and institutionalizes the functions and role played by this committee.

A concrete expression of the ties between the Treasury Ministry and the Planning Ministry could be observed in Colombia in 2007, when as part of the Framework for Medium-Term Spending process, 19 sector-based budget committees were created to establish the primary goals to be achieved in a medium-term budget context.

From isolated to more coordinated uses of information

In Colombia, coordination with other actors began to improve after the Inter-sectoral Commission on Information Policies and Management for Public Administration began to operate. The functions of this commission include eliminating duplicity in requests for information and resolving

situations in which two or more public administration entities have influence or control over the same information. Also having an impact was SINERGIA's reorientation toward M&E of exclusively macro-level results (sector-based policies and programs), as well as the creation of the programming and management by objectives and results system, which made it possible to simply and unify processes for obtaining information on management results from various public entities.

Transparency laws, which have been increasingly promoted in recent years, appear to be contributing to greater, more comprehensive use of information. In the case of Mexico, for example, the creation of the Federal Institute for Access to Public Information in the framework of the Law for Access to Public Information facilitates access to information on the performance of federal programs and federal public administration offices and entities.

Creating entities for negotiation

Relationships with sectors, particularly with the highest ministry offices, have constituted one of the most critical matters in a number of M&E systems, insofar as it sometimes has led to a weakening of these ministry offices, and consequently, of the political leadership over sectors. According to what was established at the Latin American Center for Development (*Centro Latinoamericano de Administración para el Desarrollo*, CLAD) Congress in 2008, the contrasts in the use of information on results between developed countries and Latin America constitute one of the aspects that limit results-based management in the region, particularly the institution of results-based budgets. Industrial countries tend to develop a "bottom-up" process based on the "responsibilization" of services, involving negotiations and agreements with ministries and allowing the use of results as a primary source of information. However, the process in Latin America tends to be "top-down" and based on mistrust, supervision, sanctions, and avoidance of responsibilities at the base level.

Concrete actions toward the creation of communication spaces that allow for negotiating M&E agendas as well as goals and indicators have begun to be carried out more vigorously in recent years. For example, in the case of Colombia, efforts began at the end of 2008 to develop Multiyear Evaluation Agendas, in which the idea is to avoid isolated evaluations subject only to the demands and interests of

each program. Along these lines, one of the most important advances in this component is that efforts were begun to define the programs to be evaluated on the basis of negotiations and coordination among the seven sectors of public administration considered to be of high priority, since most of the national budget was allocated to them. For each sector, the ICRME and the evaluation department (*Evaluación de Políticas Públicas*, DEPP) team define the programs they consider to be the most important for evaluating.

Brazil provides an example, in the strictest sense, of a real "system" in relation to the M&E of the Multiyear Plan. It is composed of a Monitoring and Evaluation Commission (MEC) as a joint interministerial entity, and Monitoring and Evaluation Units (MEUs) in each ministry. Beginning with the 2008–11 Multiyear Plan, the MEUs have been subordinated directly to the sector-based entity, signifying greater influence in the highest ministry offices. The MEC is the engine of the system, with representatives from the Civil Cabinet at the President's Office, the Treasury Ministry, the Environment Ministry, and the Ministry of Planning, Budgeting, and Management, where the Executive Secretariat is located.

From isolation to interconnection among M&E systems

One of the main problems reported by organizations is the repeated requests for the same information from different systems. In the end, they feel as if they are being invaded by requests. Their time is wasted and they are distracted from more important objectives. What is worse, the legitimacy of M&E is undermined. This situation is found particularly in countries in which various national systems have not been able to join together "at the base level."

Some institutions have attempted to resolve this type of problem. In Brazil, in addition to the Ministry of Planning, Budgeting, and Management that coordinates the M&E system in the Multiyear Plan, at least three other institutions systematically conduct M&E. The President's Office monitors a portfolio of goals associated with the Growth Acceleration Program (which is a set of high-priority initiatives that overlap to a significant extent with PAP programs, although are not limited to them); the Ministry of Social Development and the Fight against Poverty monitors and evaluates social policies and programs; and the National Accounts Tribunal carries out quick assessments of a

limited set of government programs, as part of its mandate for monitoring performance. What should be emphasized here is that the purposes of the different systems complement each other, and there are significant areas of overlap. However this does not signify unnecessary repetition or parallelism. In fact, both formal channels (particularly the MEC, given its interministerial composition) and informal channels of communication are broadly used, according to reports from sources. Furthermore, the monitoring of the Growth Acceleration Program is integrated at the operational level with the Multiyear Plan monitoring, since it uses data from its information system to monitor and evaluate its own initiatives. In addition, the evaluation of social policies and programs—since this involves external evaluations—provides inputs for the evaluation of related PAP programs, complementing and testing their self-evaluations.

In Mexico, M&E's close connection to public policies and programs is noteworthy. The National Council for Evaluation of Social Development Policy (CONEVAL), the Treasury Ministry, and, to a much lesser degree, the Civil Service Secretariat coordinate the Performance Evaluation System. One of the expressions of the joint efforts among these three entities is the 2007 establishment of guidelines for evaluating federal programs, initially only applicable to federal programs subject to rules of operation, yet later including other programs as well.

Greater focus on policies and accountability

In various M&E systems, there is a noteworthy change in focus toward the policies and strategies. This can be clearly observed in Colombia, Costa Rica, and Uruguay. In Colombia, the turning point for SINERGIA began in 2002, when the Uribe administration reoriented the system toward a focus on monitoring at the program and policies-sector levels. Also noteworthy is the emphasis currently placed on the function of accountability to society. In 2008, various efforts began to address the topic of a National Accountability Policy. This initiative was primarily focused on the system serving as a source of information for citizens, control entities and Congress, and not only as an information tool for decision making by the executive branch (CONPES 3294, 2004). Nevertheless, these areas have not yet been adequately developed, and it is important to continue to strengthen them.

New foundations for sustainability

Normative frameworks are not sufficient for sustaining M&E systems, but there is no doubt that they can contribute to this goal. Significant steps are being taken in the region along these lines. In Costa Rica, for example, the reforms in Article 11 of the National Constitution (2000) and the Law on Financial Administration and Public Budgets (2001) constitute a renewed platform for strengthening NES. Through these reforms, the evaluation of results and accountability is introduced as the guiding principle of public management. The new law seeks to coordinate the powers granted to the MIDEPLAN, the Treasury Ministry, and the National Comptroller General in the area of budgetary evaluation and institutional management evaluation.

In Colombia, through Article 132 of Law 1151 of 2007 (through which the NDP is regulated), a step forward was taken in recognizing the importance of evaluating the public sector. The Article, titled "Evaluation of Management and Results," indicates that

according to what is stipulated in Article 29 of Law 152 of 1994, in relation to the design and organization of systems for evaluating management and results, the Inter-sectoral Committee for Evaluating Management and Results is created as an entity for establishing priorities for program evaluation, approving methodologies and their scope and considering results that may contribute to improving processes for formulating policies. The National Planning Department will design mechanisms for efficient use of the corresponding information.²⁹

Challenges

Immediate challenges for the M&E systems

Some of the immediate challenges vary according to the system. For example, in the case of Costa Rica's NES, the importance of consolidating the self-evaluation process formally promoted since 2006 is recognized, specifically, in relation to the development and implementation of means of verification that offer transparency and legitimacy to the results reported. In Colombia, one of the challenges is to achieve the systems' institutionalization in the executive

branch and the other sectors of public administration. In Brazil, the challenge is to introduce multiple evaluation mechanisms, with participation by other actors, primarily users, for the purpose of diminishing the tendency toward self-complacency generated by self-evaluation.

Other challenges are shared, and basically involve expanding the positive tendencies that are beginning to appear, with the understanding that these are still unfinished tasks. The recurrent challenges involve achieving greater coordination between the processes of planning, budgeting, implementation, M&E, and specifically, the need to develop new mechanisms and strategies for coordination with treasury ministries to achieve a greater link between M&E information and budget decisions, as well as to ensure that the latter take political strategies more seriously into consideration. The fact that there are still cases of strict compartmentalization at the level of units of analysis certainly makes this task more difficult. For example, in Colombia, the unit of analysis used by the Budget Bureau of the Treasury Ministry and the one used by SINERGÍA often do not coincide. For some, the unit of analysis is a project, and for others it is a public program or policy, and this means that objectives, baselines, and indicators are difficult to share between them.

There is still a great deal of work to be done in terms of working in coordination with ministries and government entities, to ensure that they use the information generated by the systems. And there is a similar need in relation to improving the integration of different systems and their coordination with society. The need to build trust points to the need to replace general policies that return powers by selective mechanisms with those in which flexibility is part of what is recognized as good management. There is also a need to design mechanisms that permit the distribution of power among intersectoral entities.

Most of the challenges are not of a technical nature. They refer more to matters of systems governance, in which coordination, negotiation, and collaborative formulation of policies among diverse actors—and here it is important to identify them as political matters that affect the governance quality of M&E systems—become increasingly important. And of course, if we divert our attention to these, others of even greater magnitude come to the fore. One is clearly related to the political economy of M&E. Another concerns the compartmentalization and segmentation of

Latin American public administrations, and consequently, the type of institutions characterizing the region. These refer to the type of power and influence relations, and the values that have been consolidated through the various state reform processes.

Medium-term challenges: The new public arena and its values³⁰

And along these last lines, another set of challenges emerges. Without doubt, changes of uncommon magnitude have been generated in recent decades, and not only are we confronting a new universe of services that are public but in private hands, but also we are finding that the logic of the market has been established within the state.

In fact, most institutional incentives (new institutions) that have been promoted through state reform and public management processes are oriented toward reproducing the logic of the market within the public sector—and this is even acknowledged without hesitation by the promoters themselves. But these incentives go even further, seeking to change the state's role in the public sector, and specifically, to diminish state power in the sector. This is the real institutional change being prescribed.

One of the most noteworthy effects of the introduction of market mechanisms in the public sector is the change in its very nature. Recently, the Organisation for Economic Co-operation and Development (OECD) has even proposed a new concept, that of “public domain,” to define the public sector that has been newly configured, with the assumption that it no longer only includes entities belonging to or controlled by the government, but also includes services financed (directly or indirectly) by the government, yet provided by private organizations.

In the case of Latin America, the fragmented data available suggest a tendency toward growth particularly in the social services provided by commercial entities. Consequently, Latin America (with a few exceptions) has adopted the logic of the market for the provision of public services in more extreme terms than the countries in which this type of focus initially emerged. In this regard, in the United Kingdom, immediately after the period characterized by the deepest transformations in this direction, it was acknowledged that although the public sector might have lost its economic role, it had maintained a strong social role. It continued to finance and provide vital goods and servic-

es, including health, education, research and development, criminal justice, and social security.

Among the issues emerging since these new realities became apparent is the lack of data for identifying the costs, inputs, and, especially, the results of the new arrangements, which are needed to adopt duly established public policies in this regard. The premise, in all cases, is that the private sector, whether commercial or non-commercial, has comparative advantages compared with the public sector—although the indicators for regarding whether or not to support this statement are not specified. In fact each modality or instrument should be evaluated for efficiency and cost, quality and effectiveness, and coverage and fairness.

Another phenomenon that is largely a consequence of the application of a market focus in the public sector is the over assessment of measurements, with biases that border on formalism and with an emphasis that generates adverse effects on fairness. The assumption that what cannot be measured does not exist, and the tendency to manage “for” indicators instead of the opposite, constitutes some of its expressions. One effect of this is that the attention taken away from the *ends* becomes caught in the *means*. In fact, many measurements are currently being made, but in general we do not have access to real evaluations of changes in pertinent government capacities, or their impacts on society.

Nevertheless, what is most significant is the connection between incentives, private calculations of personal benefits, and measurements, which involve the logic of the market established in the state and its consequences for maximizing values. One of the main findings suggested by recent studies is that this connection can be translated into a lack of attention to fairness, as a maximizing value from the public sector. In addition, it has been found that goals can be converted into negative, strategy-structuring elements among actors, and can configure relations characterized by competition rather than collaboration.

All of this warns us of the type of tradition underlying the new institutional arrangements. It proves that it is not enough to focus on the positive result from recent transformations in the public sector, pointing to the establishment of a tradition of goals and indicators, knowing that this new tradition eventually may lead to negative effects on the values of the public arena and its governance.

M&E systems are not free from potentially negative aspects, in relation to the type of institutional arrangements and type of values established in the public arena, and in the end, in society. They do not exist in a vacuum, but rather are affected by both the features of the sociopolitical context in which they are being developed, and the patterns of reform in the public sector that frame their gestation and development.

A significant medium-term challenge is related to the type of state and public administrations under construction. The old pattern of the state’s configuration was inefficient, and promoted rentier societies and irresponsible bureaucracies. The new pattern, on the other hand, is leading to divided, individualist societies. Now that the state is once again recovering its importance, it is time to move away from both of these focuses. For this to happen, to begin to acknowledge the complex interdependencies between values, institutions, and behavior norms, as well as between the search for fair distribution and for production efficiency, requires more extensive research than proposed. Such research, which directly involves M&E systems, should begin by acknowledging that we currently are confronting a black box (public domain), in which a significant degree of the production of public goods and services takes place, and by admitting that no goals or indicators are totally free from outside interests, particularly if the underlying logic is one of the market, not of what is public. The governance of M&E systems is, without a doubt, an immense task that requires influencing patterns of state reform and public administration.

Nuria Cunill Grau

Special Advisor and Former Program Director, CLAD

Session 3 Notes

²¹ See the MIDEPLAN Web site, <http://www.mideplan.go.cr/content/view/87/438/> (accessed on October 1, 2009).

²² See the Deinsa Delphos Web site, <http://www.deinsa.com/delphos/index.htm> (accessed on October 2, 2009).

²³ See Baltimore City Web site, <http://www.baltimorecity.gov/Government/AgenciesDepartments/CitiStat.aspx>

²⁴ See White House Web site, www.whitehouse.gov/omb/blog/09/06/08/BuildingRigorousEvidencetoDrivePolicy

²⁵ All information on IHRME may be obtained at <http://www.reforma-iems.sems.gob.mx/>

²⁶ For information on ProbHME, see <http://www.sep.gob.mx/becasmediasuperior/index.html>

²⁷ The University of Pennsylvania research team is composed of Jere Berhman, Susan Parker (CIDE), Petra Todd, and Ken Wolpin.

²⁸ For further discussion of the M&E systems described in this paper, see the SIARE-CLAD Web site, <http://www.clad.org/siare>, particularly the section on the evaluation of results.

²⁹ Article 29 of Law 152 of 1994.

³⁰ See Cunill Grau (2009) for further development of the issues presented in this point.

Special Session: National System for Performance Management and Evaluation (Sinergia)

National Management and Results Evaluation System Achievements, Challenges, and Perspectives

During the Fifth Meeting of the Monitoring and Evaluation Network of Latin America and the Caribbean: Challenges and Opportunities for the Institutionalization of Regional Systems, a special session was dedicated to Colombia's National Management and Results System (NMRS) to commemorate the 15 years since its inception. During the session, the achievements, weaknesses, and challenges of the system were discussed and an analysis conducted of its advantages, roles, and aspects that have influenced its development.

The panel included Director of Public Policy Evaluation (PPE) of the National Planning Department (NPD) Diego Dorado; Presidential Advisor Claudia Jiménez Jaramillo; Budget Director of the Ministry of Finance and Public Credit Fernando Jiménez; Director of the National Association of National Financial Associations (ANIF) Sergio Clavijo; and Deputy Director of the NPD Juan Mauricio Ramírez, as chair.

This document includes the central points of these interventions, complementing their analyses and addresses with bibliographies by authors that have studied NMRS as a reference. The document is divided into four parts. It begins with an examination of the experience of NMRS during these 15 years of learning and then explores NMRS's links with other central government entities. It explores views on the system from outside government and ends with a description of the challenges and perspectives for NMRS for the next few years in a context of strategic realignment of the system.

Fifteen years of learning

The Colombian results-based model emerged with the Political Constitution of 1991, as part of an intense process of modernization involving a series of economic and political reforms. Regarding institutional issues, one of the changes stipulates that the state must have suitable monitoring and evaluation (M&E) tools and a system that combines the various institutional efforts of a reform-oriented government.

Three years after the new constitution, Law 152 of 1994 (Organic Law of the National Development Plan) established the guidelines for the follow-up of policies, stipulated that it must be based on the priorities and guidelines of the National Development Plan (NDP), and made the NDP responsible for the performance of these functions. These constitutional and legal mandates led to the emergence of Columbia's NMRS. This tool was designed to enhance the effectiveness of the formulation and implementation of policies, programs, and public projects to provide timely measurements of the management results of public administrators, improve public investment management, and provide appropriate information for assigning funds.

At the same time, the system would focus on follow-up and evaluation with particular emphasis on the achievement of results, as reflected in the 1994 National Economic and Social Policy Council (CONPES), document 2688/94, which stipulates the "introduction of an evaluation culture within organizations in order to boost public investment management."

Some participants suggested that the system can be divided into two generations of evolution: 1995–2001 and

2002–09. The so-called first generation (1995–2001) saw the development of the basic conceptual, legal, and instrumental framework of the evaluation function. Results-based management practices were incorporated into central administration and the continuous use of management indicators was introduced, with certain planning by goals and results tools in central government bodies. These plans included (1) the indicative plan (tools for following up the main NDP programs, designed for the strategic managerial sphere), (2) the action plan, (3) efficiency agreements (intended to improve coordination between the organizations responsible for the various levels of expenditure results, beginning with a clear distinction between their skills and the definition of what their contribution should be to emerging from a fiscal crisis), (4) management indicators, and (5) commitment and evaluation (special documents that included the decisions made on the sector goals between the ministers and the president on the basis of the priorities established in the NDP).

During this stage, the World Bank provided continuous support for NMRS, through the Modernization of Public Financial Administration (MPFA) project. This first generation ensured the methodological solidity and conceptual maturity of the system's M&E instruments. It permitted the creation of a space to encourage a results-based management culture and to learn more to ensure its proper implementation. It also provided a better diagnosis of the institutional context, which subsequently made it possible to design a more efficient follow-up system to meet the country's needs.

The second generation (2002–08) began with a complex scenario for NMRS, determined mainly by the following aspects:

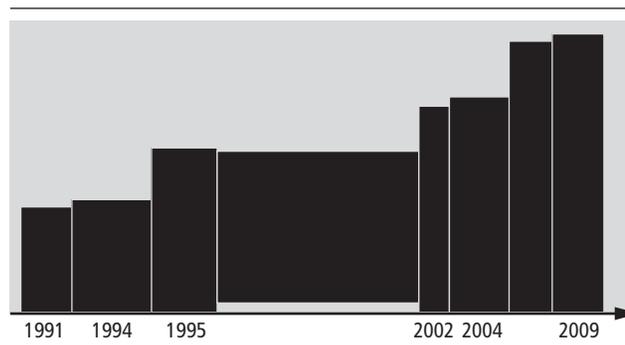
- The commitments and tools of the government bodies were detached from the budgetary assignments for each sector.
- The various links, tasks, and responsibilities of the system were not clearly defined.
- Data management was poor. Too many indicators (940 in 2002) focused on measuring management rather than results, and, in turn, discouraged the achievement of specific goals associated with this management. Their processing placed a burden on government bodies that was not reflected in productivity gains.

- There was no clear separation between the concepts of monitoring and evaluation.
- There was a lack of incentives to demand evaluation on the part of the various institutions and results were not disseminated, thereby undermining accountability.
- There was a lack of knowledge of the actual effects of investments.
- Knowledge for decision making was insufficient.

This scenario, coupled with Álvaro Uribe's assumption of the presidency, raised the need to restructure and strengthen NMRS's operating scheme. This point of inflection would enable the system to change its approach "from monitoring at the agency-sector level to the program and policy-sector level. On the basis of this reform, the system made a clear distinction between the monitoring and evaluation activities and stressed the results-based assignment of public expenditure" (Villareal 2007).

This new approach has been reflected in various important landmarks, such as (1) the implementation of the System of Government Programming and Goal Management (SGPGM) in 2002, (2) the Strategic Plan for 2006–10, and (3) the drafting of the 2004 National Council of Economic and Social Policy (NCESP) 3294 document, "Renovation of Public Administration: Results-based Management and the Reform of the National Evaluation System." The SGPGM permits real-time monitoring and provides citizens with access to the state of management as a function of the set goals and a base line, corresponding to the start of the administration. It identifies those responsible for carrying out every policy and program, while linking results and

FIGURE 16:
Milestones of Reform, 1991–2009



Source: Author.

TABLE 9:

Changes Introduced through NCESP 3294

	NCESP 2688, 1994	NCESP 3294, 2004
Objectives	Introduce a culture of evaluation within the departments to boost the managerial capacity of public administration.	Help improve the effectiveness and impact of policies, programs, and institutions; improve the efficiency and transparency of programming and the assignment of resources; and promote transparency in public administration, by actively incorporating civic control.
Components	The system consists of two modules. The first includes the mechanisms designed for the self-evaluation of departments regarding the management and results of the implementation of policies, programs, and projects. The second includes mechanisms designed for evaluation by the corresponding ministries and the NPD, of the projects, programs, and policies periodically selected for this purpose.	Monitoring or follow-up of results. Focused evaluations. Dissemination of results for accountability.

Source: Head Office of Public Policy Evaluation, NPD.

Note: NCESP = National Council of Economic and Social Policy; NPD = National Planning Department.

goals to the investment resources assigned in the General Budget of the Nation. The Strategic Plan promotes the institutionalization of NMRS within public administration. Table 9 summarizes the changes introduced through the NCESP 3294 document.

For its part, NCESP 3294 is the policy guideline that made it possible to unify concepts and methodologies. It is the road map that currently governs the system and defines the scope and limits of the development of its functions. In general terms, the modification of the evaluation and follow-up function in Colombia sought to achieve three objectives:

- Improve the effectiveness and impact of the policies, programs, and institutions
- Increase the efficiency and transparency of the programming and assignment of resources
- Encourage transparency in public management, by actively incorporating civic control

The reform divided the system into three complementary components or strategic lines, by establishing conceptual and instrumental differences between them (see Figure 17). It then explored each of these lines or components in greater depth.

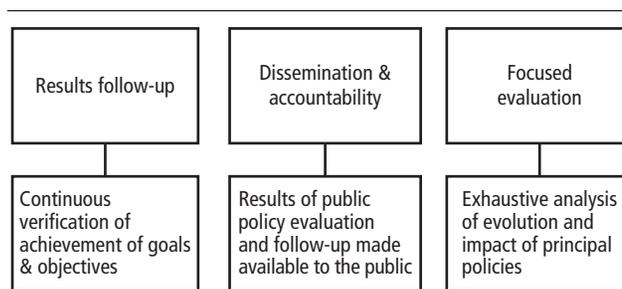
Follow-up of results

The goal is to verify continuously the achievement of the NDP's goals and objectives. This component establishes

the most significant commitments for each department at the start of each administration. The tool for the follow-up is the System for the Follow-Up of Government Goals (originally known as SFGG and now known as SISMEG). This system records the indicators for the 20 departments of the executive branch that are responsible for achieving the goals that will contribute to achieving the NDP goals. Within these departments, a "goal manager" is designated to assume responsibility to record progress toward the established goal.

Its presence online and its virtual consultation environment enable the system to obtain up-to-date information on physical and budgetary performance in real time from those directly in charge. This can reduce the transaction costs inherent in the process, and make management and follow-up more transparent. It also enables decision makers

FIGURE 17:

Results Road Map

Source: Head of Office of Public Policy Evaluation.

to obtain timely information without having to resort to the usual, traditionally established routes.

Through this system, citizens, Congress and entities responsible for control are able to exercise social control and undertake consultations related to the commitments made with those responsible for policies and programs and the progress achieved in their respective indicators. The excessive number of indicators that once made up the system has been reduced.

After the system records information on the indicators' progress, a monthly report is drawn up to be analyzed by the President's Ministry, which ranks achievement of the indicators by sector and organization, and gives details of poor or satisfactory levels of achievement and performance. Every three months, virtual publications are disseminated on issues linking goals and investment by specific sector. An annual report, known as the Results Sheet, is submitted to the president of congress on each completed contract. The state of progress of certain indicators is given in Figure 18. Likewise, the follow-up of results has achieved the following:

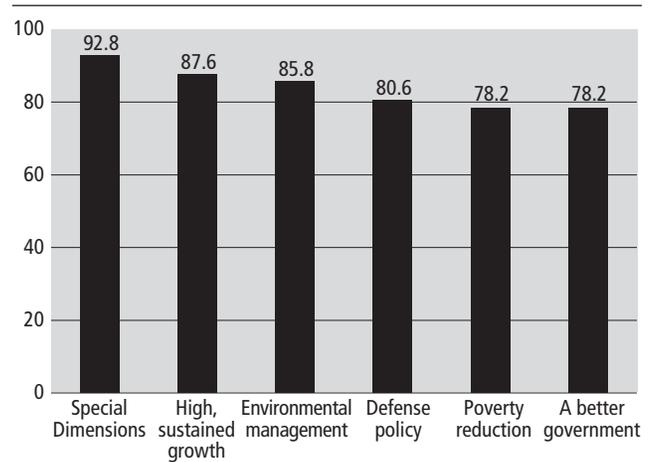
- Two development plans have effectively been monitored (2002–06 and 2006–10).
- Early warnings are issued on the plan's progress (short-term follow-up).
- 624 indicators are used.
- Monthly reports are available (average level of updating, at least 92 percent).
- Over the past year, 285 follow-up meetings have been held.

Focused evaluations

According to NCESP 3294, evaluation is defined as an “exhaustive evaluation of the causality between a state intervention and its effects in determining its relevance, efficiency, effectiveness, impact and sustainability.” In this respect, NMRS has gradually incorporated annual and four-yearly evaluation agendas defined by the Inter-Sector Committee of Evaluation and Results-Based Management (see Figure 19).³⁶ The scheme of prioritization of evaluations includes such criteria as (1) the amount of resources used by programs, (2) the characteristics of the target population, (3) the number of persons affected, the importance of the program for a specific sector, and (4) the innovative na-

FIGURE 18:

National Development Plan Goals



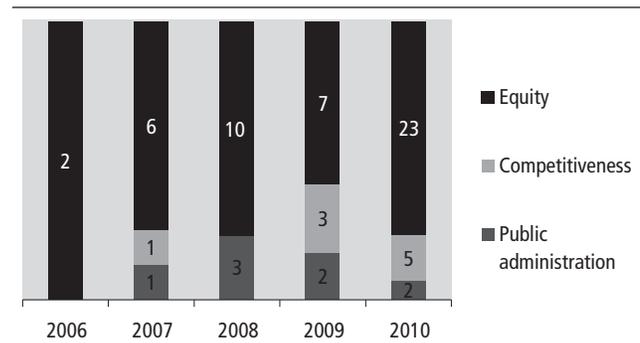
Source: Head Office of Public Policy Evaluation, NPD, cutoff date: October 31 2009.

ture of the activity and its potential to be replicated or reproduced.

All the evaluations carried out to date use contracting systems provided through a competitive bidding process, in which consultants or consulting firms outside government develop the required evaluation and the Public Policy Evaluation Office gives technical support and monitors the evaluation. In addition to carrying out robust impact evaluations using experimental and quasi-experimental models, NMRS has developed less costly evaluations depending on the program being analyzed. Thus, management, process, institutional, results-based, and executive evaluations complement the range of methodologies used by the system to evaluate programs.

FIGURE 19:

Annual Evaluations, 2006–10



Source: Head Office of Public Policy Evaluation, NPD.

To date, the NMRS has achieved considerable progress in evaluation issues. This has been borne out by the increase in the number of evaluations carried out per year (two evaluations on average in 2006, 12 evaluations in 2009, and 30 planned for 2010), the links established between the evaluation results and policy formulation and reformulation, the increase in investment in evaluations (since 2006, nearly US\$15 million have been invested), the diversification of the agenda, which now includes competitiveness and public administration issues, and the fact that evaluations are carried out by widely recognized consulting firms independent from government.

At present, as part of its process of strategic realignment, NMRS is attempting to incorporate a component called RADAR to compile and collect the public policy evaluations carried out in the various public organizations, territorial departments, research centers, and universities in the country. This initiative will make it possible to improve and optimize the use of evaluation studies, encourage coordination and complementarity in this sphere, use decision making to achieve more effective management, and produce information for the various social actors.

Dissemination for accountability

The main purpose of this component is to promote interaction between the citizens, the control entities, the executive branch, and the other branches of government. The component will implement the principle of practice in public management. In other words, it means making the results of the evaluation and follow-up of public policies available to the public.

As far as dissemination is concerned, the aim is to provide appropriate dissemination for NMRS products, by improving the production of documents in terms of quality and availability and ensuring that the information produced is timely, clear, and reliable for its users. Thus, information has been delivered on the evaluation results carried out on various national government programs and the follow-up of the NDP goals.

The various tools for dissemination include sending documents physically to organizations, academia, and civil society organizations, and by e-mail to more than 9,000 addressees. Likewise, forums and seminars are held, some on an international scale, to disseminate the work carried out by the Head Office of Public Policy Evaluation as the NMRS

Technical Secretariat and thereby create top-level technical and academic debates on the issue of evaluation, follow-up, and accountability regarding public management.

Moreover, the NPD Web site has incorporated an NMRS link that includes documents, products, and services as well as control tools and input for accountability. NMRS has a virtual documentary center of nearly 500 articles related to evaluation, follow-up, and accountability. This progress has been made available in (1) NCESP documents on the progress of public policies and evaluation reports, (2) evaluations in the "Public Policy Evaluation" series, (3) an annual management report by and for the president of congress, and (4) technical support to ensure national government's accountability to the community through televised Council of Ministers meetings. In short, the dissemination of results and the availability of information have enabled the government to direct management toward results. Likewise, thanks to the administration's technical support and credibility, during these 15 years of learning, NMRS has managed to build a strong M&E system in government.

This is particularly noteworthy because much of the information produced by the system is intended for those responsible for implementing the policies evaluated. In this respect, it is essential for governments, as policy implementers, to have the instruments for measuring management results and to participate in this process (Zaltsman 2006).

NMRS and intergovernmental coordination: A tool for guiding political and budgetary decisions

NMRS's links with other government bodies was crucial to the institutionalization of M&E throughout the government. This link is explicitly generated, insofar as the system produces quality technical information that the government uses as input in political decision making. At the intergovernmental level, the use of the information produced by the system operates within the context of the president's exercise of political control over the government cabinets, and the assignation of investment spending as an input for budget planning in the Ministry of Finance.

Political control

NMRS enables the executive and, more specifically, the president, to verify the achievement of NDP goals, monitor

the results of government management and those responsible, improve state planning on the basis of priorities, and generally make results visible to make management more transparent. Figure 20 illustrates this coordination.

For the President’s Office and the Council of Ministers, the information produced by NMRS is particularly important insofar as it is the primary source used for management controls. This use contrasts with the use it could be given by ministries in other scenarios to support decision making.

The case of “Families in Action” has shown that the information produced can be used to carry out structural or rapid changes in the implementation of the operation. The most obvious structural change occurred because of the evaluation results, whereby the number of beneficiaries tripled and virtually all the municipalities in the country were included. This last change involved the program’s implementation in cities and the introduction of the population displaced by violence as the new beneficiaries of the program.

Budgetary planning

The relationship with the Ministry of Finance and Public Credit has yet to be formally institutionalized, although

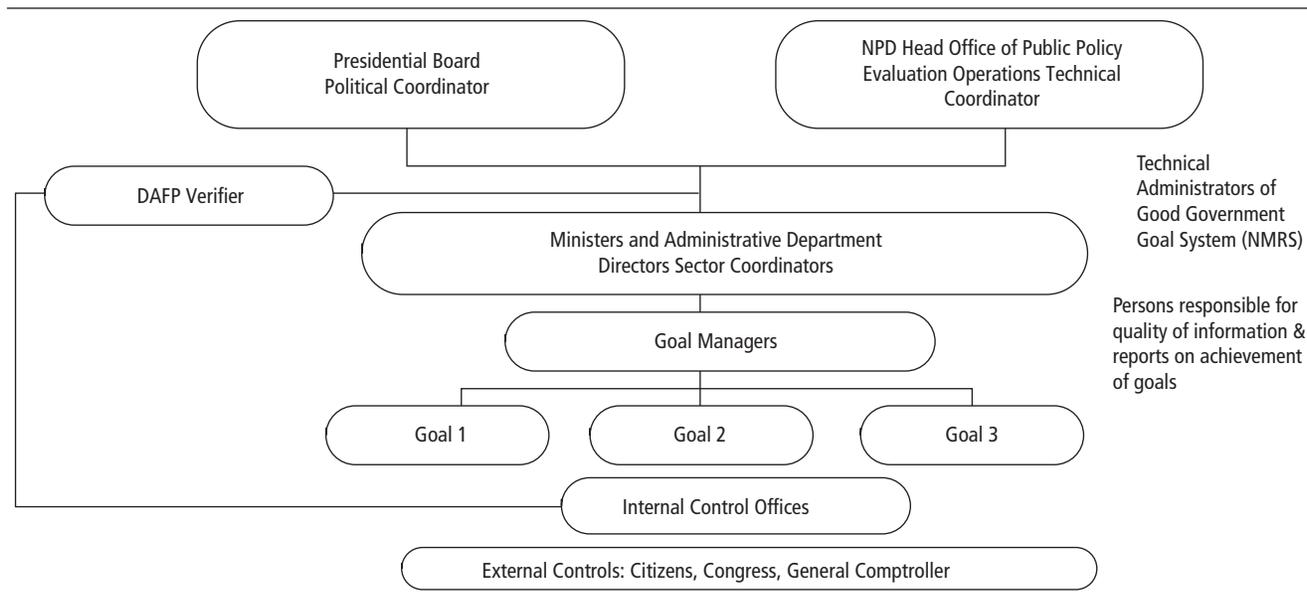
certain approaches have been established for follow-up. This relationship is established through coordination with the follow-up group in the System for the Follow-up of Government Goals. In addition to the ministry’s obligation as a government entity to provide sector indicators, it uses the information available in the system and links it to annual goals with the assigned investment budget, as a first step toward the programmatic classification of expenditure.

Another approach is in the organization of a sector expenditure committee. This committee undertakes a medium-term expenditure framework exercise, follows a strategic projection process, and prioritizes expenditures. This facilitates the achievement of fiscal goals that look beyond the annual horizon and permit better results within the time required to assign resources, and it ensures that the resources are used more efficiently. Lastly, where evaluations are concerned, the Ministry of Finance forms part of the Intersectoral Committee of Evaluation and Results-Based Management.

Links between the NMRS and the Finance Ministry have been weak because, among other things, the relationship has not been institutionalized. Although the information on follow-up and evaluation may exert a limited influence over the distribution of the national budget in the

FIGURE 20:

Goals Management System



Source: Head Office of Public Policy Evaluation, NPD.

Note: NMRS = National Management and Results System; NPD = National Planning Department.

short term, considerable potential could exist for the findings in the medium and long term, particularly with clear evidence of the government's performance in achieving presidential goals.

It is essential, then, to strengthen relations between the Ministry of Finance and NMRS, because the influence that M&E information can exert on decision making regarding the budget is still unclear. Although under the current scheme it is the NPD's responsibility to design the investment budget, it has absolutely no influence over the operating budget or current spending designed directly by the Finance Ministry. Likewise, coordination could be a prerequisite for achieving a budget and planning by results.

One of the challenges raised by the new stage of NMRS's strategic realignment is to work in conjunction with the Ministry of Finance. To achieve this end, NMRS and the Ministry of Finance proposed an initiative to establish a ranking of efficiency and effectiveness functioning in public organizations to increase transparency and the optimization of resources. Table 10 summarizes the channels of communication between NMRS, the Ministry of Finance, and the Presidency.

Outside perspectives: How credible is the information produced by NMRS?

An external diagnosis of the reliability and usefulness of the information produced by the system indicates that the outlook is not that gloomy. In fact, to say that it is dangerous to carry out self-evaluations, to the extent that the "critical" counterpart is absent, may be controversial. First, NMRS performs self-evaluations to the extent and in the way that results-based management requires. The government carries out an internal follow-up of the NDP and its goals ac-

ording to the constitutional mandate in Article 343 and Article 29 of the 1994 152 Law. Undertaking them is not necessarily "dangerous," however, because follow-up and evaluation are enabling it to measure the degree of progress of the commitments set at the beginning, seeing how its management goes, and trying to correct mistakes in time.

Likewise, obtaining this information from the executive branch may be more practical insofar as ministries, implementing organizations, and administrative departments are the owners of administrative information, whereas the production of primary data by an external party would be virtually unviable.

As for goal setting, the government's definition does not necessarily mean that they measure nonessential things. On the contrary, goal setting is the main instrument through which a democratically elected government achieves the mandate it has been given by its citizens. Goals will reflect the priorities, capacities, and resources made available by the government plan, as a condition of delegation to the executive branch. One should not forget that the process of constructing indicators and goals as an objective unit of measurement is agreed on at both the technical and the political level. This implies not only the consideration of methodological and technical aspects, but also introduces political validation by the top decision-making bodies that give the process political muscle, together with a high level of appropriation.

Likewise, NMRS has been interested in citizens' opinions and perceptions, as borne out by the fact that its virtual platform allows readers to make judgments about the government's performance and the progress of its commitments and the accessibility, quality, and usefulness of available information. On the other hand, in 2006, the PPE

TABLE 10:

Channels of Communication

	FOLLOW-UP	EVALUATION
PRESIDENT'S OFFICE	Management Controls, Council of Ministers and public accountability. Presentation of Report to Congress.	Inter-Sector Committee of Evaluation and Results-Based Management coordinates the definition of the evaluations agenda for every four-year period and the evaluations agenda for every fiscal period.
MINISTRY OF FINANCE	Functional Investment Committees to link goals and investment budgets. Ranking of effectiveness and efficiency of public entities.	Inter-Sector Committee of Evaluation and Results-Based Management technical coordination and liaison between evaluation and budgetary programming.

Source: Head Office of Public Policy Evaluation, NPD.

undertook a survey on the quality of the results information, asking more than 3,000 households in 13 capitals about transparency, degree of response, accountability, and availability of information, among other issues. One of the results that emerged suggests that nearly 47 percent of citizens regard public information as useful, which raises a challenge and establishes a commitment to guaranteeing the availability of more and better information.

Second, in evaluation issues, in recent years, NMRS has played the role of judge or third party in contracting evaluations (carried out by external consultants, as mentioned earlier). This is done to avoid the wrong sort of incentives, because the ministry or executing entity that pays the evaluator is being evaluated as well.

Thus, the selection of evaluating firms or the consultant is done through competitions, tenders, and public bidding sessions that use rigorous selection schemes endorsed by the World Bank. Applicants usually include think tanks, universities, prestigious research centers, and consultancy firms, with teams of well-known professionals. For academics who wish to protect their reputations and studies, being hired by a government does not exempt them from the scrutiny of academia. This fact should guarantee the academics' objectivity.

Third, the Head Office of Public Policy Evaluation is currently working on the development of an M&E network in Colombia. This development provides a master plan for the analysis of the achievements and difficulties found during the institutionalization of M&E systems. It is also a way to strengthen intergovernmental relations and provide civil society with information about public policy evaluation that is not necessarily carried out by the NPD or even the national government.

Challenges for NMRS over the next few years

Over the next few years and in the process of its strategic alignment, it is essential for NMRS to focus its efforts on consolidating the work it has carried out with the remaining entities in the public sector, particularly with the Ministry of Finance. Likewise, the main challenge is to produce relevant, quality information for citizens. The following are

some of the strategies NMRS has proposed for the next few years:

- Increase links between NMRS and other NPD areas
- Increase follow-up of the development plan through budgetary follow-up
- Link the design of follow-up indicators to the development plan with a long-term development view
- Diversify the evaluation agenda, including sectors that did not traditionally participate
- Identify the evaluations that other government bodies have submitted to guarantee international standards for evaluations (RADAR)
- Link the evaluations to broader analysis dimensions than the program that has been evaluated
- Promote policy evaluations in regional and local governments
- Define an inclusive policy for internal, external, and social accountability
- Create spheres of exchange with the territorial level to continue expanding the culture of results-based management (RAM).
- Promote training in follow-up and evaluation at both the national and territorial level.

Diego Dorado

*Director, Evaluation of Public Policies (DEPP),
DNP, Colombia*

Special Session Notes

³¹ See the NPD's Web site, <http://www.dnp.gov.co/PortalWeb/Default.aspx>.

³² On the NCESP, see <http://www.dnp.gov.co/PortalWeb/tabid/55/Default.aspx>.

³³ 624 indicators are used. Cutoff date October 31 2009.

³⁴ This committee is composed of the presidency, the NPD, the Ministry of Finance, and the respective ministries.

³⁵ See the Web site, www.dnp.gov.co/sinergia.

³⁶ Although NMRS did not take part in this exercise, sector committees do constitute a coordination and negotiation link between the Ministries of Finance and Planning in the effort to consolidate a results-based budget.

Session 4: Monitoring and Evaluation Systems: Capacity Building for Monitoring and Evaluation and Academia-Government-Civil Society Partnerships

Conceptual Framework for Developing Evaluation Capacities Building on Good Practice in Evaluation and Capacity Development

Introduction

This presentation contributes to a better understanding of the complexities of evaluation capacity development and suggests ways of approaching them, within an institutional context or at national level. It aims to make use of “good practice” in capacity development to ensure attaining “good practice” in evaluation. It builds on established “good practice” in these two professions and on the experience of evaluating capacity development.

Both professions—evaluation and capacity development—are multifaceted and evolving, which means this chapter captures only a segment of the whole at a certain point in time. Capacity development practitioners continue to develop and refine concepts and practices. The Organisation for Economic Co-operation and Development’s Development Assistance Committee (OECD/DAC 2006) publication on capacity development recognizes the ongoing challenges. In a similar way, the evaluation profession continues to grow.

The demand for evaluation is high, but increasingly critical of evaluation practices that are not demonstrating independence, credibility, and utility. In addition, following the Paris Declaration and the Accra Agenda for Action, demand is increasing for evaluation capacities in partner countries, and the evaluations of the “Delivering as One” initiative of the United Nations will be country-led. Therefore, developing evaluation capacities is important for a

range of stakeholders in the development and humanitarian communities. The chapter aims to provide these stakeholders with a conceptual framework to support their efforts. It is divided into three main sections: one to discuss evaluation capacities at three levels, which overlays the evaluation principles and capacities at three levels, one to discuss the capabilities and processes to instill them, and finally one to discuss examples and draw out suggestions for stakeholders working on evaluation capacity development.

Evaluation capacities at three levels

Evaluation principles

The evaluation principles date back to a meeting of evaluation professionals in 1991 and probably to earlier meetings. They have become integral to the evaluation profession at large. For instance, independence has taken central stage in the multilateral development banks when the evaluation function at the World Bank started reporting to the executive board rather than the president; a change that was replicated in regional development banks and that stimulated discussions in, for instance, the United Nations. In 2008, the framework for peer reviews raised three principles to be central to evaluation: independence, credibility, and utility (DAC/UNEG 2007).

Independence forms the bedrock of good evaluation practice (Chelimskey 2008). It is fundamental to attaining credibility and utility of evaluation. It should lead to impartiality reflected in, for instance, the choice of subjects of evaluation or the selected evaluation method. It is achieved through means at structural, institutional, and individual levels.

Credibility of evaluation is enhanced with greater independence but needs to be accompanied by competence of the evaluators, transparency of the evaluation process, and impartiality of evaluators and process.

Utility of evaluation is not guaranteed with independence and credibility, but requires that commissioners and evaluators undertake the evaluation with the intention to use its result, and undertake it at a time when the results can meaningfully inform decision-making processes and that evaluations are accessible. This principle exists to ensure evaluations are undertaken to influence change.

Together these principles are markers of quality evaluation and ensure “good practice” in evaluation. Independence, credibility, and utility are three equal sides of the triangle at the center of which the quality of evaluation rests. Their interrelationship, including inherent complementarities and tensions, is illustrated in Figure 21. Each of the central evaluation principles has other, supporting ones.

Complementarities exist among these principles: an evaluation that is biased toward the perspective of one stakeholder (meaning not independent or impartial) tends not to be credible with others. If an evaluation lacks credibility it is unlikely that stakeholders will pay attention to it, let alone use it. Ignoring the stakeholders, including their interest or timelines, will reduce the utility of an evaluation but also the credibility of its process. By maintaining a good balance among independence, credibility and utility, these three principles will reinforce each other and enhance the quality of the evaluation.

But the evaluation principles also have tensions. For instance, independence can lead to isolation of the evaluation

function, which reduces its usefulness. Or, some stakeholders might perceive evaluations to be useful solely when they report only successes (and omit information about failures), or when it serves fund-raising or publicity purposes. Or, policy makers may want to pursue a certain course of action that they do not like to see derailed by evidence that would advise otherwise. Such evaluations have little credibility with other, well-informed stakeholders and are of little, if any, use for resolving problems, as they fail to discuss problems. In other examples, stakeholders believe that credibility can derive exclusively from close familiarity with the subject under evaluation, but they ignore the possibility that such a close link eliminates independence and reduces the chance that the evaluation will produce an impartial and credible evaluation that provides new insights.

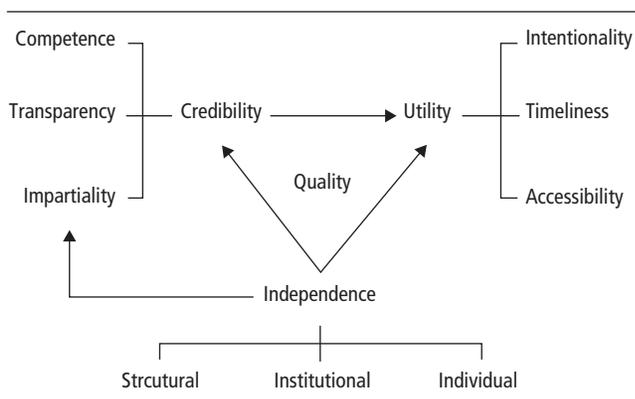
Each evaluator must act in line with these evaluation principles. Measures are needed to institute the independence, credibility, and utility of evaluation to ensure that adherence to these principles is less dependent on individuals. Evaluators may be subjective. They may lack an understanding that independence is granted to ensure impartiality and not for other reasons, and they are at risk to be put under pressure from stakeholders who want to influence evaluation findings. Therefore, it is important to have an institutional framework that holds evaluators accountable and protects them from undue influence, and an enabling environment that supports the notion of evaluation, learning, and accountability.

Capacity development: working at three levels

Capacity development is recognized by many as central to development and progress. The Accra Agenda for Action is the latest such commitment to strengthening capacities in support of ownership of development processes. For many years, investments into developing capacities have been manifold and have been supported by many different actors, many of whom struggle with similar challenges. The concept of capacity development evolved from a focus on training of individuals, through institutional development, when it was recognized that individuals worked within the context of their organizations and therefore more than training was needed to be successful (UNIDO 1990). Capacity development evolved to acknowledge that organizations do not work in isolation but rather require an enabling environment that consists of, among other things,

FIGURE 21:

Evaluation Principles



Source: Author.

policies, networks, and an attitude of engagement. Capacity therefore goes beyond an individual or an organization, but includes all three levels, as illustrated in Figure 22.

Evaluation capacities at three levels

Capacity development practitioners highlight the importance of working at three levels, which—when applied to the world of evaluation—illustrate the need to institutionalize the evaluation principles with measures that go beyond the individual, and span the institutional framework and the enabling environment for evaluation.

Enabling environment

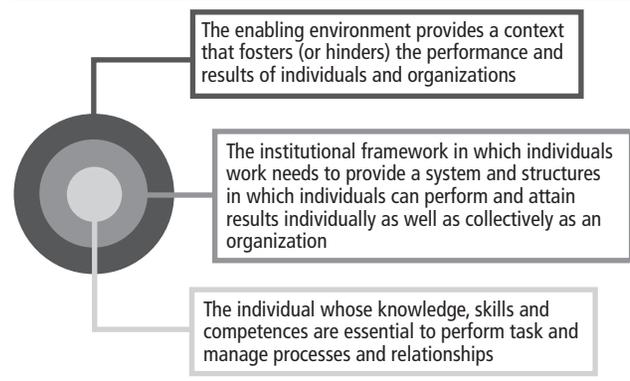
The enabling environment for evaluation is determined by a culture of learning and accountability, meaning the degree to which information is sought about past performance, the extent to which there is a drive to continuously improve, and the need to be responsible or accountable for actions taken, resources spent, and results achieved. Such culture is embedded in tacit norms of behavior, the understanding of what can and should—or should not—be done, in many cases behaviors being role-modeled by leaders. They are, or should be, codified in government legislation or an evaluation policy that expresses the leadership's or an organization's commitment to learning, accountability, and the evaluation principles. An enabling environment is supported or created through governance structures that demand independent evaluation, whether through parliaments or governing bodies, and is further enhanced through professional associations and networks that set standards and strive toward greater professionalism in evaluation. Finally, the structural independence of an evaluation function is important to setting an enabling environment. Placing the evaluation function in such a way that it does not report to the person or function responsible for the policies, strategies, or operations that are evaluated creates an environment that enables greater independence, credibility, and utility. Ideally, the enabling environment is such that decision makers proactively demand impartial evaluations to inform their debates and choices, which increases the usefulness of evaluations.

Institutional framework

The institutional framework for evaluation ensures that a system exists to implement and safeguard the independence, credibility, and utility of evaluation. It reduces two

FIGURE 22:

Levels of Capacity Development



Source: Author.

risks: (1) that declared commitments to independence are revoked through personal actions—systems can be observed and, once established, more difficult to reverse without agreement of all stakeholders—and (2) that inherent in the dependence on the individuals, who need a system of checks and balances to ensure accountability and protection. Such an institutional framework

- Includes a system of peer review or assurance that the evaluation function is set up to safeguard and implement the principles of independence, credibility, and utility.
- Establishes safeguards to protect individual evaluators—evaluators, evaluation managers, and heads of evaluation functions—when exercising their independence, including transparent and credible processes for the selection, appointment, renewal (if applicable), and termination of contract of the head of evaluation and assurances that evaluation staff do not suffer in promotion exercises.
- Ensures that a multidisciplinary team exists in the evaluation function or on an evaluation team that ensures credibility of evaluation by understanding multiple dimensions of evaluation subjects and combining the necessary technical competence.
- Secures the independence of funding of evaluations, at an adequate level, to ensure that necessary evaluations are carried out and that budget holders do not exercise influence or control over what is evaluated and how. Therefore, funding should be

BOX 2:

Structural independence

Structural independence requires that an evaluation function or office does not report to the same individual who is responsible for the subject of evaluation. Otherwise, the evaluation function might be under political or organizational pressure that interferes with the independent and impartial planning, conduct, and reporting of evaluation findings. For example,

- National systems evaluators report directly to parliament to ensure the highest degree of independence and ensure this decision body benefits from impartial evaluation insights;
- Many of the development banks, starting with the World Bank, this structural independence translated into evaluation functions reporting directly to the executive boards of these banks;
- The UN system for which the Evaluation Norms suggest that a reporting line to the governing body and/or the head of the organization were adequate to ensure structural independence; and
- The practices of the Food and Agriculture Organization or the United Nations Development Programme, which introduced dual reporting lines: one functional reporting line to the governing body and an administrative line to the head of the organization.

Structural independence does not guarantee impartiality. It requires the personal or behavioral independence of evaluators, and institutional measures that help safeguard independence and ensure it leads to impartiality.

under the direct control of the head of the evaluation function, and should be adequate for a reasonable work program. The adequacy of funding for evaluation can be determined through a ratio of evaluation resources (human and financial) over the total program size (some organizations aim to allocate 1 to 0.8 percent of their total budget for evaluation) or the coverage rate of evaluations (for instance, evaluating 25 percent of the overall program, or to evaluate a minimum of 30 projects per year to have a valid sample size).

- Combines measures for impartial or purposive selection of evaluation subjects to ensure impartially on the one hand and increased utility on the other by making deliberate choices linked to decision-making processes.
 - To avoid biases in the selection of subjects for evaluation, objective selection criteria are used to ensure the evaluation sample is representative of the whole. This principle would apply to the sample selection in a project evaluation

when choices are made for site visits, the selection of projects or programs for inclusion in an evaluation work plan of an organization or of initiatives of a national program. As a result, the findings of the evaluation(s) will be as representative as possible, and therefore indicate strengths and weaknesses that are likely to apply to the project, program or portfolio as a whole. There is no bias to select only the good performers or only the problem cases.

- To ensure utility, for instance, of strategic or thematic evaluations, these may be chosen purposively (rather than to establish a representative sample) to link the conduct of the evaluation to the information needs and decision-making processes of those using evaluation findings. In these cases, a consultation process to determine what are the most important and/or strategic topics for evaluation is important to ensure various stakeholder needs are considered.

- Sets out a system to plan, undertake, and report evaluation findings in an independent, credible, and useful way. To increase objectivity in the planning and conduct of evaluation, systems are needed that increase the rigor, transparency, and predictability of evaluation processes and products. Such systems can include more or less detailed process descriptions or guidelines for the design of evaluations, for preparatory work that needs to be conducted and reported on, and for reporting findings. The processes should have built-in steps for communication, consultation and quality assurance, and be communicated to stakeholders to enhance transparency and to secure their willingness to share information.
- Institute measures that increase the usefulness of evaluations, including the sharing of findings and lessons that can be applied to other subjects. Evaluations should be undertaken with the intention—of stakeholders and evaluators—to use their results. For projects or programs, this intention can derive from plans to use evaluation findings in the design in a future operation. Collectively the findings of these evaluations can be used to understand the performance of an organization with the intention to replicate well-performing aspects and rectify systemic problems. The timeliness of planning and conducting evaluations and presenting their findings is equally important to ensure utility. Finally, accessibility means that evaluations are available to the public, can be retrieved (for instance, through a user-friendly Web site), are written in ways that are understandable (clear language and limited jargon), and are distributed to a wide group of relevant stakeholders.

Individuals

Debates about professional standards and accreditation of evaluators and evaluation managers have been ongoing within professional forums for years. Even with structures and systems in place, the independence and impartiality of evaluation depends on the integrity and professionalism of individuals as evaluators, evaluation managers, and the head of evaluation.

The profession and its ethics require limiting personal biases to the extent possible. In some circumstances, exter-

nal evaluators are believed to exercise greater independence than those who work in an organization, because they are less exposed to institutional or peer pressures or have not absorbed the corporate culture in an unquestioning way. However, individual or intellectual independence *is* dependent on the individual who demonstrates in his or her behavior the adherence to and practice of the evaluation principles: avoiding conflict of interest, acting with integrity and with an independence of mind, engaging in evaluations for which they are competent, acting impartially, and undertaking an evaluation with a clear understanding of the clients or decision-making process and how these need to be informed. Many evaluation associations and evaluation functions of national and international organizations adopted codes of conduct for evaluators, and the United Nations Evaluation Group developed additional ethical guidelines (UNEG 2007). Table 11 provides an overview of the points discussed above and illustrates the intersections among the three levels of capacities and the three evaluation principles.

From capacities to capabilities:

The process matters

The above discussion centers on individuals and entities, such as governing bodies, evaluation units, or evaluators and how they collectively form an evaluation capacity of mutually reinforcing ingredients. To be effective, these evaluation capacities need to be capable of delivering evaluation services—producing independent, credible, and useful evaluations—and much more.

Five capabilities

The five-year project on capacity development (ECDPM 2008) concluded that capacities, at whichever level they exist, needed to be able to achieve the following:

- Commit and engage: volition, empowerment, motivation, attitude, confidence
- Carry out technical, service delivery, and logistical tasks: core functions directed at the implementation of mandated goals
- Relate and attract resources and support: managed relationships, resource mobilization, networking, legitimacy building, protecting space
- Adapt and self-renew: learning, strategizing, adaptation, repositioning, managing change

TABLE 11:

Evaluation Principles and Levels of Capacity

	Independence	Credibility	Utility
Enabling Environment	<ul style="list-style-type: none"> • Culture of Accountability and Learning • Government Legislation or Evaluation Policy • Community of Practice, Networks, Associations for Evaluation • Governing or oversight body that seeks independent credible advice • Structural independence 	<ul style="list-style-type: none"> • An attitude of providing access to information that allows credible evaluation 	<ul style="list-style-type: none"> • An attitude that entails the intention to use evaluation findings and recommendations • Willingness to change
Institutional Framework	<ul style="list-style-type: none"> • Institutionalized process of "peer review" to assess the independence, credibility, and utility of the evaluation function • Independent budget • Impartial selection of evaluation subjects • Independent planning and conduct • Noninterference in reporting of findings • Measures to protect evaluators from repercussions • Ethics Guidelines • Code of Conduct 	<ul style="list-style-type: none"> • Evaluation quality assurance system to ensure credibility, transparency, and impartiality • Multidisciplinary evaluation team that works well together 	<ul style="list-style-type: none"> • Understanding of the use of evaluation • Consultation processes • Timeliness • Accessibility of evaluations • Active sharing of lessons from evaluation • Utility demonstrated by the actual implementation of evaluation recommendations
Individual	<ul style="list-style-type: none"> • Avoidance of conflict of interest • Behavioral independence • Integrity 	<ul style="list-style-type: none"> • Competence (technical knowledge and evaluation skills) • Impartiality 	<ul style="list-style-type: none"> • Client orientation • Communication • Coordination with different stakeholders and with ongoing processes

Source: Author.

- Balance coherence and diversity: encourage innovation and stability, control fragmentation, manage complexity, balance capability mix

Evaluation capabilities

This recognition is important for moving from having, for instance, a written evaluation policy that means little, however, because it is not backed by capabilities to commit and engage. An evaluation unit may exist and carry out evaluations, but may lack the capability to relate to stakeholders and attract resources, or to adapt and self-renew by learning to provide new evaluation types in response to changing needs.

In addition, these capabilities need to be interpreted differently depending on the stakeholders concerned. For instance, the commitment and engagement of stakeholders in the enabling environment will set the culture for learning and accountability, whereas the capability to commit and engage for an evaluation unit entails setting and abiding

by evaluation principles, and for the evaluator entails an attitude of behavioral independence, among other things.

The application of these capabilities to the three levels at which evaluation capacities need to be established is summarized in Table 12.

The enabling environment for evaluation needs to be committed to a culture of learning and accountability. It should adopt an evaluation policy in line with the evaluation principles and use evaluation findings and insights in its policy making, performance improvements, and organizational renewal. It accepts that the independence of evaluation needs to be safeguarded, including its funding. In a national context, this could entail embedding evaluation into legislation and government policy. In an organization, the corporate culture, evaluation policy, and resources for evaluation form the context in which an evaluation function operates.

The evaluation function should be committed to the evaluation principles, protect these and the evaluators from

TABLE 12:

Evaluation Capabilities

Capabilities	Enabling Environment for Evaluation	Institutional Framework for Evaluation	Evaluators, Evaluation Managers, Heads of Evaluations Units
To commit and engage	A culture of learning and accountability that empowers individuals and organizations to reflect on their practice, take stock of what works and what does not, and take necessary action.	The commitment to evaluation principles, and the motivation and attitude to follow them through in daily practice.	Commitment to evaluation principles, ethical guidelines, and code of conduct. Behavioral attitude of independence and impartiality. Motivation to conduct evaluations in a credible and useful way.
To carry out technical, service delivery, and logistical tasks	An evaluation policy that codifies evaluation principles and good practice.	Ability to conduct I/C/U evaluations to support organizational adaptation and renewal.	Competence to conduct evaluations, both from an evaluation and a technical point of view.
To relate and attract resources and support	To legitimize evaluation through the evaluation policy and actions, to recognize and ensure the need for adequate and independent funding.	The ability to demonstrate legitimacy based on impartiality and credibility, to protect evaluation principles and standards and the evaluators, and to generate adequate funding.	Capability to conduct evaluations transparently and credibly, and to communicate and relate with stakeholders.
To adapt and self-renew	Recognition of evaluation playing a role in adaptation and self-renewal.	The ability to make strategic choices about evaluation's engagement in organizational renew and to relate evaluation findings to stakeholders to ensure learning. Frequently adapt and update evaluation methods and approaches.	Learn new evaluation skills.
To balance coherence and diversity	Manage change, using evaluation evidence and recommendations when available. Avoid fragmentation of systems for learning and accountability.	The ability to ensure evaluation processes and approaches are systematic but flexible, balance capabilities on the team, and employ a mix of approaches to develop an evaluation culture.	Find the right balance between systematically applying evaluation guidelines and seeking opportunities for innovation.

Source: Author.

pressures, and have the motivation to implement them in their daily practice. It should have a system that safeguards and institutionalizes independence, credibility, and utility.

In its relationships with others, it needs to demonstrate legitimacy based on impartiality and credibility (including technical competence) and it needs to contribute to the adaptation and renewal of the context within which it works—an organization or a wider, say, public sector reform beyond just one organization. It has to have the capability to work systematically (to ensure transparency) but flexibly (to ensure credibility and utility), use a mix of tools to foster an evaluation culture, and keep its methods and approaches updated. The evaluation unit needs to have a mix of capabilities on its teams and has to have the capability to determine and negotiate adequate funding.

The individual evaluators—as staff of an evaluation unit or as consultants contracted to conduct a particular evalua-

tion, as evaluators, evaluation managers of heads of evaluation units—have to be committed to the evaluation principles, the ethical guidelines, and the code of conduct. He or she has to demonstrate independence and impartiality, and conduct evaluations in a credible and useful way. Evaluators are expected to be technically competent as evaluators or competent in the field that is being evaluated, or both, and keep updated in his or her knowledge of evaluation methods and techniques. He or she should have the ability to follow guidelines, but seek possibilities for innovation.

The process matters

The process of capacity development is, in itself, an important part of the outcome, because capacities and capabilities are developed through a process of interaction and dialogue, mixed with specific initiatives to set up systems, install hardware, and train people. The process combines parts that are

well planned, based on participatory diagnostics of capacities and their weaknesses, and others that are flexible and opportunistic. Just the number of stakeholders and the range of capacities and capabilities discussed above illustrate the complexities of capacity development and flag the need for planning and implementation tools that are able to address complexity, something that blueprints tend not to be good at. Instead, space needs to be created for analyzing capacity gaps in a participatory way, agree on common goals and translate them into a joint strategy, while at the same time maintaining the flexibility and capability to take up opportunities as they arise, learn from experience, and change tactics and work on various levels of capacity at the same time.

Capacity development practitioners arrived at the conclusion that 10 process matters are important for greater success in capacity development. In this chapter, these process issues have been regrouped around (1) the drive from within, (2) development and agreement of clear expectations, and (3) “ordered chaos” or combining high degrees of flexibility with systematic approaches.

The drive from within has to do with the realization that capacity development needs to come from within or have strong internal champions, whether it is when developing a national system or a function in an organization. If an idea is imposed from the outside, chances are that capacity development is not likely to succeed. The drive from within relates to the capability “to commit and engage” and includes the following:

- Ownership, which manifest itself in having local or internal champions, resource allocations, and engagement. It is, however, difficult to measure, not homogenous across internal stakeholders, and not constant over time
- Leadership, which is important, especially for setting the culture of learning and accountability, but also must fit with the context and its culture
- Collective action, motivation. and commitment, meaning that capacity development will not occur if it is one person writing documents, but requires engaging various stakeholders whose capacities and capabilities will be developed

The process of developing and agreeing on clear expectations in capacity development is in itself part of the capacity

development process. Developing a common understanding of capacity weaknesses through a participatory, structured diagnostic—using tools like the conceptual framework in Figure 22 or the five capabilities in Table 12 can serve as an analytical framework to move attention from resource gaps to recognize broader issues—and a shared vision of the needed capacities are important steps in channeling resources toward clear objectives. Very often the many actors in capacity development have tacit understandings of what capacity is and how its development should happen. Few of them have explicit strategies, which makes it more difficult to find a common ground and work in the same direction. A diagnosis of existing capacities is necessary to find entry points to start capacity development. This diagnosis benefits when it is developed in a participatory way that develops ownership at the same time.

The “ordered chaos” or combining high degrees of flexibility with systematic approaches entails a recognition that capacity development, while systematic, does not (always) follow an ordered process. Instead it includes planned as well as incremental and emergent approaches. For instance, when developing specific skills, a well-structured process would be appropriate. On the other hand, when trying to influence culture to become more supportive of learning and accountability, opportunities might arise at short notice, in various settings, and without prior notice, and include an informal conversation with a key stakeholder. In these contexts, capacity development initiatives would be incremental and emergent in response to opportunities as they arise. Capacity development also requires recognizing the less tangible aspects of capacities and combining small and large initiatives, depending on context and opportunities. It requires finding the right balance between an operating space that allows capacities to evolve and accountability for capacity development results; a balance that is hard to attain. And, finally, capacity development takes time, so it is necessary to stay the course (even through adverse times), but build *quick wins* into the process (to keep up motivation and build on success) and allow time to reflect and evaluate whether progress is being made.

Where to start? Illustrative examples: What does it mean for stakeholders?

There is no right answer to the question “Where to start?” Depending on the situation, a diagnostic could be the best

entry point to understand where the strengths, weaknesses, and gaps are. Sometimes, however, it might be necessary to first have a champion within who provides ownership and leadership of the process and generates interest in developing evaluation capacities. In other situations, it might be clear that the quality and credibility of evaluation has to be improved before anything else can happen. Or, in other cases, the incentives for accountability and learning need to be improved before anything else can happen. And, in many cases, it requires a mix of all of these things to develop capacities, but these choices are situation specific.

An entry point for developing institutional capacities is a diagnostic, which can be done using the conceptual framework suggested here, and through peer reviews in the bilateral system, the UNEG and Evaluation Cooperation Group of the multilateral development banks. This expert review can support a diagnostic process and lend it further credibility, but the evaluation unit concerned needs to be involved in and own the diagnostic process. The diagnosis could lead to a revision of policy, as has happened in the case of the United Nations Development Programme, or to the formulation of a strategy to develop evaluation capacities.

Few organizations have articulated an evaluation strategy, even less so one that explains how evaluation capacities will be developed. Yet, good practice in capacity development argues for articulating such strategy to have a common vision and bring stakeholders together to work toward a common goal. Such strategy would employ linear planning approaches for some activities, for instance when implementing a training incremental approach, for instance, when gradually expanding the evaluation requirements in a national system as discussed in the Sri Lanka case study to emergent approaches, when a number of flexible, informal initiatives are needed, for instance, to respond to opportunities when possible and necessary.

The following paragraphs highlight examples to draw out suggestions for stakeholders and how they might work toward developing evaluation capacities. It is structured around the three levels—the enabling environment, the institutional framework, and the individual—as these three levels associate well with identifiable stakeholders. At any of these levels, it is advisable to understand the existing capacities, the capabilities, and reluctances to adopting cer-

tain standards to develop a corresponding and relevant capacity development strategy.

Placing evaluation capacities in the context of good governance

Governance and evaluation are interdependent. Good governance creates an enabling environment for evaluation, while evaluation reinforces good governance. Sri Lanka provides a good example of how the commitment from the highest political level—the prime minister—created an enabling environment for evaluation. By demanding reports on results, stakeholders in ministries were enabled, even required to record and critically reflect on performance. This strong leadership was combined with a strong champion to translate political commitment into practical action.

Ideally, governments or chief executives of organizations seek independent, credible, and useful advice through an evaluation function. Demand for evaluation creates a strong enabling environment for evaluation. Decision-makers demand feedback on the use of resources and results achieved with them and set, through their demand for such information, an environment that enables evaluative thinking and practice. However, political and institutional power structures can limit the environment in which evaluations are conducted, especially if vested interests resist evidence that might demonstrate the weaknesses of political choices. In these cases, it is important to counterbalance political power structures with other stakeholders and to establish legislation or an evaluation policy that legitimizes and commits the country or organization to the evaluation principles. An increasing number of United Nations organizations are adopting evaluation policies. In the case of the World Food Programme, the evaluation policy establishes the commitment to the evaluation principles and explains measures to safeguard them.

An environment that is unresponsive to or even fearful of evaluation might originate in a lack of understanding of the value of evaluation to make more informed decisions. In these cases, it is for the evaluation function to deliver credible, high-quality evaluations to demonstrate the value added of evaluation. In addition, it will be necessary to explain the role and usefulness of evaluation in the context of governance and performance of the organization as a whole. Table 13 provides tips for stakeholders.

TABLE 13:

Tips to Reinforce an Enabling Environment for Evaluation

Governing Bodies (Parliaments, Executive Boards, and so on)

- Be aware of the threats to independence, credibility and utility of evaluation and demand for measures to safeguard these principles
- Adopt and oversee the implementation of legislation or policies that institutionalize the independence, credibility, and utility of evaluation
- Exercise oversight over the quality of evaluation
- Demand for evaluation and demonstrated use of evaluation findings and recommendations

Head of State, Minister, Policy Makers, Chief Executives

- Demonstrate leadership in setting a culture of learning and accountability
- Seek and use evidence from evaluations to validate the attainment of goals and objectives and to improve performance whenever possible
- Understand evaluation as part of good governance that aims to use public resources effectively and efficiently

Evaluation Function

- Demonstrate value added of independence, credible, and useful evaluations
- Raise awareness of stakeholders about the role and importance of evaluation and of the evaluation principles
- Contribute to evaluative thinking through awareness building, dialogue, and training

Professional Evaluation Networks, Associations

- Set standards as benchmarks that can be used to convince other stakeholders about the importance of the evaluation principles and measures to safeguard them

Source: Author.

Delivering evaluation services independently, credibly, and usefully

The importance of developing an institutional framework for evaluation, rather than using ad hoc arrangements, lies in the safeguarding of the evaluation principles, the provision of a framework that protects and holds evaluators accountable, and the predictability and transparency of this framework for all stakeholders. Such a system needs to be built in ways that it has the capability to adapt and self-renew, one of the five capabilities, to respond to new challenges with, for instance, a focus on new issues or by introducing new types of evaluations that address information gaps.

Generally, the evaluation functions need to lead efforts to develop the institutional framework for evaluation, setting up systems, processes, and guidelines that are in line with the evaluation principles and fit with the work of their organization or the national context.

These systems may include human resource management measures that ensure the selection of an independent and qualified head of the office of evaluation. The best systems will not—and should not—eclipse the human factor, but this makes it important to select and appoint heads of evaluation functions that are capable of safeguarding the evaluation principles and managing the evaluation function so that it delivers independent, credible, and use-

ful evaluations. The process should assure governing or oversight bodies and chief executives alike that an impartial candidate has been chosen, and the selected candidate should have contractual arrangements that protect him or her from repercussions, should critical evaluations be unwelcome. In addition, it is important that a system is set up by which evaluators are held accountable for their actions through a code of conduct or ethics guidelines, such as those established by the UNEG.

To set up a planning system that is representative for evaluations of operations, the Office of Evaluation at the United Nations World Food Programme (2008b) analyzed the portfolio to identify factors that should be used to ensure a selected sample of operations is representative of the total. This resulted in a system that combines the size and number of operations to reflect the considerable variation in these factors: a relatively small number of operations represent a large proportion of the financial value of the portfolio, while the rest is spread over a large number of small operations. Treating each operation, regardless of size, equally would have overemphasized the smaller operations in the sample; selecting the sample by size would have underrepresented the smaller operations in the overall sample. The weighted numbers are applied to determine how many operations in each geographic region should be selected for

BOX 3:

Selecting the head of evaluation

Selecting the Head of Evaluation

- The process should be agreed upon between key stakeholders.
- The process should be documented and publicized for reasons of transparency.
- The composition of the panel needs to ensure credibility in the eyes of key stakeholders, requiring the representation of stakeholder groups and the profession.
- The panel should have clear terms of reference that specify whether the panel “chooses” or “suggests”, and if it is the latter who has the final decision-making power.
- The panel should be involved in the entire process from agreeing on the vacancy announcement, the selection criteria, agreeing on interview questions and criteria for assessing answers, interviews, and so on.
- Appointment should be made for a fixed term period, including a contract for the same duration.
- Contract renewal should follow a process that involves the panel to ensure the renewal is based on performance and not dependent on producing biased evaluations.
- The termination process, while following standard human resource procedures, should again involve the panel to ensure termination of contract is based on performance issues and not a repercussion for critical evaluations.

evaluation. By contrast, a different system was needed for the strategic evaluations to ensure that they address issues that are important to key decision makers in the governing body and management, and revolves around a consultation process with these stakeholders. The utility of all evaluations is enhanced by linking them to decision-making processes.

An example of a system that ensures complete independence of finance is that of the International Fund for Agriculture Development (IFAD). Its evaluation policy determines that the budget is prepared independently of that of the fund's secretariat. The process is managed by evaluation office and involves discussions and approval of the

TABLE 14:

Tips to Develop an Institutional Framework for Evaluation

Governing Bodies (Parliaments, Executive Boards, and so on)

- Get briefed about the evaluation system to understand whether the institutional framework includes adequate checks and balances and to become a discerning reader of evaluation reports.

Head of State, Minister, Policy Makers, Chief Executives

- Get briefed about the evaluation system to understand whether the institutional framework includes adequate checks and balances and to become a discerning reader of evaluation reports.
- Introduce processes for the selection of the head of evaluation that ensures his/her independence.
- Provide secure, separate, and adequate funding for evaluation.

Evaluation Function

- Develop and document systems for the selection, design, conduct, and reporting on evaluations.
- Provide briefings on these standards to increase transparency and confidence in the process and products, which enhances credibility.
- Develop mechanisms to ensure lessons from evaluation are systematically shared and integrated into debates and decision-making processes.

Professional Evaluation Networks, Associations

- Set professional standards and good practice standards.
- Develop and implement professionalization, accreditation, and credentialing systems.

evaluation committee. The work program and budget of the evaluation office appears as a separate, clearly identifiable section in the fund's overall work program and budget.³⁹

A further system that is important to develop is one that governs evaluations processes and sets standards for evaluation products. They often take the form of guidelines or handbooks. Placing information about the process into the public domain increases transparency of evaluation processes. The production of guidance materials often needs to go hand in hand with briefings: between the evaluation manager and the evaluation team, and between the evaluators and the managers of operations to clarify expectations in evaluation quality and process. The process needs to include formal feedback processes and quality standards and assurance mechanisms. If evaluations are conducted in a decentralized way—that is, not managed by the central evaluation office—training will be needed to ensure that standards are understood and implemented.

Evaluation functions need to have capacities to incorporate findings and recommendations into decision-making processes as well as to share lessons from evaluations. These systems can include search engines on Web sites of repositories of lessons from evaluations, incorporating evaluation lessons into guidance materials that program managers use, or proactive efforts to share evaluation lessons throughout the evaluation process.

Impartiality, knowledge, and skills of individuals

The best enabling environment and greatest institutional arrangements will not replace the individuals involved in evaluation. Their integrity to plan, undertake, and report evaluation findings in an impartial way is essential for the

profession. They determine whether an evaluation is credible and useful. Their professional competences—as evaluators and in the technical area that they evaluate—and their ability to listen, understand, and weigh the different views of stakeholders are central to the quality of evaluation.

The importance of training as a tool for one part of the capacity development process is well understood. The chapter illustrates the importance of a structured training program that is based on specific training needs assessments and training goals. In addition, it is important to coach and mentor evaluation managers and evaluators, especially when they are working for the first time in this field.

Conclusion

This chapter demonstrates the synergies that exist between the evaluation profession and that of capacity development practitioners. A combination of good practice from both sides promises results in greater effectiveness to develop evaluation capacities that will safeguard and use the evaluation principles of independence, credibility, and utility and also promises higher returns on investments in evaluation of capacity development. It comes at an important time when evaluation needs to professionalize, and when demand for evaluation and for developing such capacities is high.

The consideration of three levels of capacity (the enabling environment, the institutional framework, and the individual) are important to institutionalize evaluation capacities in ways that it supports and protects individual evaluators, evaluation managers, and heads of evaluation functions, while holding them accountable. These three levels, together with the capabilities that are needed, provide a conceptual framework that moves the debate from training individuals for improved skills or calls for resources to a more sophisticated set of ingredients that promise to be mutually reinforcing capacities. Applying the conceptual framework to specific examples provides a number of pointers for stakeholders who want to work on evaluation capacity development.

Caroline Heider

*Director, Evaluation Office, World Food Program,
United Nations*

TABLE 15:

Tips for Stakeholders to Develop Skills and Knowledge of Evaluators

Evaluation Function

- Invest in training
- Build teams that are multi-disciplinary
- Encourage team work and cross-fertilization

Providers of Evaluation Training

- Continue to develop and offer training courses
- Increasingly move towards university courses to build a professional tradition

Professional Evaluation Networks, Associations, etc.

- Provide opportunities for professional exchanges and networking

BOX 4:

Definitions**Definitions**

Accessibility means that stakeholders have unrestricted access to evaluations that can be retrieved easily from their storage place. It also requires that these evaluations are written in ways that are easily understandable and that findings are shared with stakeholders and interested parties.

Accountability is the obligation to account for (and report on) work carried out and results achieved.

Competence of evaluators means that they have and can demonstrate a good understanding of the subject under evaluation and of the evaluation principles, rigorous data collection and analysis, including ethical principles for evaluation.

Credibility is the extent to which evaluation findings and conclusions are believable and trustworthy. Credibility is determined by objective factors, such as the transparency of the evaluation process and accuracy of the evaluation report, and subjective factors such as the perceived or demonstrated impartiality and competence of the evaluators.

Learning means that lessons are drawn from experience and, accepted and internalized in practices, thereby building on success and avoiding past mistakes.

Impartiality is the absence of bias; it entails due process, methodological rigor, and the consideration and presentation of achievements and challenges, success and failures.

Independence means that the evaluation is free from influences—political or organizational pressures or personal preferences—that would bias its conduct, findings, conclusions or recommendations. It implies that evaluations are typically carried out or managed by entities and persons free of the control of those responsible for the design and implementation of the subject of evaluation.

Intentionality is a clear intent to use evaluation findings. In the context of limited resources, the planning and selection of evaluation work has to be carefully done.

Transparency means that consultation with the major stakeholders is an essential feature in all stages of the evaluation process. This improves the credibility and quality of the evaluation. It can facilitate consensus building and ownership of the findings, conclusions and recommendations

Timeliness means that evaluations are chosen and undertaken in a manner that their findings and recommendations are available in time to inform decision-making processes.

Utility of evaluations means that they aim to and do affect decision making. Therefore, evaluations must be perceived as relevant and useful and be presented in a clear and concise way. Evaluations are valuable to the extent to which they serve the information and decision-making needs of intended users, including answering the questions posed of the evaluation by its commissioners.

Sources: EC Evaluation Glossary (online), OECD/DAC 2002, DAC/UNEG 2007, UNEG 2005, UNEG 2007, United Nations World Food Programme 2008a.

The Experience of the Latin American and Caribbean Institute for Economic and Social Planning in Evaluation Training

The Latin American and Caribbean Institute for Economic and Social Planning (*Instituto Latinoamericano y del Ca-*

rite de Planificación Económica y Social, ILPES) is a permanent body with its own identity, forming part of the Economic Commission for Latin America and the Caribbean (ECLAC). Created in 1962, its mission is to support governments in the region in the area of governance, through

research, consultancy, and, in particular, training. The institute has worked since late 2002 as a central training body of ECLAC and, in that role, it helps to disseminate the work of the ECLAC system, particularly that of an institutional nature.

To perform its role as the central organ of ECLAC training, ILPES has at the United Nations headquarters in Santiago, Chile, three fully equipped classrooms and 10 workshop rooms with 22 computers and six printers available to students. In addition, the auditorium, with a seating capacity of 55, is equipped with facilities for simultaneous interpretation. ILPES has all the necessary audiovisual equipment, such as video projectors and facilities, to conduct video conferences. ECLAC has three large-capacity conference rooms that can be used particularly for the opening and closing sessions of a course. To perform its job, ILPES is structured around two thematic areas and a training unit. The thematic areas are Budgetary Policies and Governance and Local and Regional Management. The main topics of each of these areas are as follows:

Budgetary Policies and Governance

- Fiscal and budgetary policy
- Socioeconomic evaluation and project management
- Evaluation of public programs and management

Local and Regional Management

- Regional development and territorial competitiveness
- Decentralization
- Local and regional management
- Planning and prospects.

The Training Unit provides all the logistical and administrative support necessary to conduct the various courses organized by the institute at its Santiago headquarters and in cities across the continent, either in partnership with other agencies or national, regional, or local governments.

To illustrate the efforts made in training by the institute in its 45 years of existence, it has taught various types of courses that have involved approximately 24,000 participants mainly from the public sector in all the countries in the region. It also has had the benefit of significant participation by teachers from universities and academic centers

in Latin America and the Caribbean including people from private sector institutions and civil society.

The training courses offered by ILPES/ECLAC revolve around specialized topics not usually addressed by academic programs in Latin America and the Caribbean. They combine theory with practice and systematize the experiences of countries in the region. Moreover, the international nature of many of the courses and teamwork offer participants an interesting exchange of views and problems connected with Latin American economies.

Each year several courses are open to international participants, as well as national and subnational courses. The language of instruction is Spanish and the courses normally last between one and five weeks.

The teaching staff consists mainly of ILPES experts and technicians from various substantive divisions of ECLAC. The ILPES takes pains, in turn, to recruit senior staff from governments, public agencies, universities, academia, and the private sector to teach or give lectures. In 2009, ILPES organized and conducted international courses on the following:

- Tax estimates
- Evaluation of government management and programs
- Evaluation of impact of programs to overcome poverty
- Management of project implementation
- Strategic management of local and regional development
- Strategic planning and construction of performance indicators
- Macroeconomic policies and public finance
- Budgetary policies and public management by results
- Preparation and evaluation of public investment projects

As for national courses within the framework of agreements with national or subnational governments of member countries, ILPES has carried out such activities in almost all countries of the region. In some cases, it was a single course or a few courses, but in others, it has developed extensive training programs for public sector officials, sometimes with more than 100 courses. This is the case of training programs held in Chile (1996–2000), Colom-

bia (1989–1990), and Mexico (ongoing since 2007). This year, courses have been conducted within the framework of agreements with national governments in Argentina, Brazil, Costa Rica, Mexico, and Peru.

For example, in Mexico's case, the latest agreement is intended to strengthen the technical capabilities of those responsible both at the federal and state levels for the design, implementation, monitoring, and evaluation of the matrices for performance indicators to ensure that these matrices are consistent, that they do not show any incongruence in their logic, and that their indicators are constructed with the highest possible technical rigor and supported by quality information. The products of this agreement are as follows:

- Workshop courses for improvement in construction of matrices for indicators of federal programs. Designed for staff of federal government programs and agencies, and at officials of state programs.
- Courses on awareness of the results-indicator matrix for senior officials of federal programs and federal entities.
- Technical assistance for improvement of indicator matrices and their indicators. Aimed at federal and state programs.
- Workshop courses for improving the matrix indicators of federal programs. Designed for federal government officials responsible for the design, monitoring, and evaluation (or related areas) of performance indicators of budgetary programs.

The ILPES long-distance education program, provides online training to students who cannot go to training venues in person. Since 2005, more than 1,500 people have taken the various types of courses offered, including participants from Latin America, India, South Africa, Bosnia, and Haiti, among others.

This approach uses the Internet, and the course design is based on the “learning by doing” approach, which considers the practical application of content delivered. Currently, the two courses are “Basic elements of local and regional development” and “Logic for the formulation of development projects.” Preparations are under way for other courses that refer to specific ILPES-related topics. The institute's extensive experience in training courses and programs allows us to give some recommendations for a training pro-

gram in evaluation to have the desired impact. These are as follows:

- Training should be theoretical and practical.
- There should be courses at different levels of specialization.
- Training should be inserted in a process.
- Training should include managers.
- Training should be ongoing.

Theoretical-practical training

Students quickly forget what they hear, but learn from what they do. All the training activities undertaken by the ILPES therefore combine theory and practice. Generally, mornings are devoted to presentation of the theoretical aspects of the issue being addressed in the course, either by ILPES or ECLAC experts, public sector professionals, or consultants. In the afternoon, students work in groups to apply to real cases what they have been taught in the morning.

In the formation of these working groups, the greatest possible diversity of professions, diversity of nationalities in international courses, and gender balance should be sought. This greatly enriches the learning process for the important exchange of personal experiences that occurs. It helps to achieve better results in workshop activities because participants with different educations provide different perspectives on the problem to be solved or issues to evaluate. In general, working groups should have no more than five or six members, because when a group is larger, some members probably will not keep pace with the rest (social pressure decreases when responsibility in a more numerous group is diluted).

Finally, practical experience has taught us that it is better for the learning process to work with real projects or programs than with cases prefabricated for the course. Prefabricated cases facilitate the workshop, but they cannot prepare students for the reality to be faced on returning to their jobs and having to evaluate a project. Instead, working with real projects or programs teaches students how to solve the difficulties they will face, such as incomplete or incorrect information.

Courses with different levels of expertise

Countries that have been most successful in the creation and operation of national public investment systems have

developed large-scale training programs in preparation and evaluation of projects, procedures, and the system's use of information about projects. This normally will include a series of courses at basic, intermediate, and advanced levels, requiring a large number of basic courses, fewer intermediate courses, and a few advanced courses.

The goal of basic-level courses is to provide trainees with a culture of project evaluation. This culture aims to establish a common language, to know the objectives of project evaluation and their value, and to create the skills to implement simple methods of formulating and evaluating projects.

The intermediate-level courses continue to train trainers in the development and evaluation of projects, as well as technical experts in project evaluation. Their role is to provide basic-level training and technical advice within the public sector, thus contributing to the sustainability of the SNIP and quality of the results obtained.

Finally, advanced training, often delivered through partnerships with universities, aims to form the top layer of system professionals, who will take on the responsibility of developing the investment system and fulfill duties as trainers or experts in evaluation.

Given the high staff turnover that often occurs in the public sector, these training programs should be permanent. To this end, the entity responsible for organizing the investment system should be concerned with organizing courses at the different levels, either directly or through agreements with national and international entities. Undoubtedly, it will require a large number of basic courses, some intermediate, and a few advanced level courses.

Training inserted in a process

To cause a significant and lasting impact, training in evaluation should be inserted in a process of reform and modernization of the state. If the student returns to the job and is not faced with the need to apply the acquired knowledge, he or she is likely to forget the material. In other words, having an evaluation training program that is not accompanied by reforms in resource allocation procedures to make evaluations compulsory makes no sense. In addition, to ensure that courses taught in the training program meet the requirements stipulated, it is essential to have an adequate methodological backup. This must be found in manuals that serve as teaching material in courses and as guides for practical evaluation work.

In short, an evaluation-training program is justified only if inserted into a process of modernization of public administration in terms of investment management. It should include methodological developments and procedural and even legal and institutional reforms.

Training authorities and managers

Having spent many hours carrying out a socioeconomic evaluation of a project or obtaining the results of a program, only to find that whoever runs the institution does not use the evaluation results, or worse, demands a different result, can be extremely frustrating for any public official. So if you want to create a culture of evaluation in the public sector, it is essential to consider factoring into the design of the training program lectures for senior directors and authorities who will receive the evaluation results.

For example, when Mexico began to implement results-based budgeting and the performance evaluation system and decided to use the logical framework methodology to evaluate the results of federal programs, talks were scheduled in the conference room of the Ministry of Finance for the directors of programs and government ministries in charge. More than 800 directors attended these two-hour lectures. This allowed them to have a basic knowledge of the tool to use, thereby facilitating dialogue with the technicians who would later prepare the matrices for the programs.

Ongoing training

In most countries in the region, there is a high turnover of public officials, due to their seeking better wages in the private sector, movements between government ministries, or changes that occur with a new government administration. This results in a significant drain of personnel trained in evaluating projects and programs. To compensate for this drain, training programs should be permanent. In fact, the most successful experiences in relation to public investment systems have been accompanied by major programs and training that has been sustained for decades. Such is the case, for example, of the National Investment System of Chile. The Ministry of Planning, responsible for its operation has been concerned about training technicians with various levels of expertise in evaluation. So it has held courses on its own account, had an agreement with

ILPES for several years under which more than 20 courses were held annually, and has signed agreements with various universities to train public officials at intermediate or advanced levels.

These training programs, which usually begin under agreements with international agencies that finance the creation of the system or support its operation, cannot continue to rely on external support always, either financially or academically. It therefore is essential to involve universities in the process, so that they can offer courses on evaluation, which are conducted in partnership with the government or included in the curricula of the courses they teach.

Universities are a key source of knowledge and experts in the field of evaluation. Whenever ILPES has organized large-scale training programs in different countries, university professors have been part of teams of consultants who have been allowed to develop them. Universities also are a source for creating studies and knowledge on evaluation of projects and programs.

The evaluation courses that universities offer traditionally have been limited to diplomas or master's degrees lasting one or more years. Conversely, a massive training program aimed at public officials requires the enactment of numerous short courses (one to four weeks). This requires the public sector to engage with universities, signing deals to offer the courses that government needs. Examples of this type of agreement can be found in Chile and Colombia.

An advantage of universities in the organization of courses by international bodies is that they can award diplomas that are valid for the purposes of government career development. The courses improve résumés and enable attendees to apply for other positions in the public sector. International bodies such as ILPES cannot provide such diplomas, but have the advantage of providing a broader view of the subject and the experience gained in other countries. Therefore a significant synergy exists in partnerships between international agencies and academia in conducting training programs in evaluation.

Eduardo Aldunate

*Expert, Budgeting Policies and
Public Management Area CEPAL, Chile*

Strengthening Monitoring and Evaluation Capacities through the Design and Implementation of National Results-Based Monitoring and Evaluation Systems: Lessons Learned in Africa, Asia, and Latin America

Abstract

Strengthening monitoring and evaluation (M&E) capacities is more effective and efficient if a midterm training plan is designed to meet internal monitoring needs for piloting public programs and the objective evaluation of results and performance of the public sector. This plan should include various components such as more decentralized, effective, and diverse training; complementary technical assistance; support for information systems; sensitization discussions; policy discussion and exchange platforms between professionals; institutional support for national, sustainable M&E units; development of external evaluation capacity; and a culture of accountability. It is a road map defined in a coordinated fashion at the national, sectoral, and local level to progress in a timely, modular fashion in various M&E systems and to contribute to the professionalization of both M&E. In a mature national M&E system with information-rich and learning organizations, the borders between M&E persist but in a complementary fashion with overlapping areas.

Context

M&E is one of the pillars in the implementation of Management for Results (MR), also known as Management for Development Results (MfDR), in developing countries' public sectors. The 2005 Paris Declaration on the effectiveness of aid highlighted the importance of improving the M&E of interventions for development (High Level Forum 2005). The Accra Action Agenda of 2008 reinforced the commitment of developing countries and donors to demonstrating results through increased accountability and transparency (Third High Level Forum 2008). Developing countries pledged to improve "the quality of policy design, its implementation and evaluation by improving information systems." Developing countries and donors have agreed to work to develop cost-effective instruments for MR to evaluate the impact of development policies and adjust them if necessary. The MfDR team has assembled an MfDR Capacity Scan tool kit to help assess

progress in national MfDR capacity, including M&E capacity (OECD/DAC 2009). Strengthening evaluation capacity building (ECB) and improving the capacity for internal monitoring of public programs has become a priority on the development agenda. In a context of limited public resources exacerbated by the financial crisis, the importance of improving national M&E capacities with suitable, inexpensive, sustainable approaches cannot be overemphasized.

We should recall two basic considerations. First, the domain of M&E covers a variety of subareas. Second, ECB covers far more than just training. ECB entails reinforcing or constructing M&E systems, particularly national systems, so that M&E is regularly conducted and used by the countries and organizations (OECD 2004).

Several ECB initiatives have been implemented as a component of development projects or by specialized institutions (World Bank Institute and Institute for Development in Economics and Administration, IDEA) and specific programs (International Programme for Development Evaluation Training, IPDET, Paris 21). Unfortunately, few ECB initiatives have been carefully evaluated. The few significant evaluations carried out, such as the self-evaluation of ECB, conducted by the World Bank Operations Evaluation Department (OED) in 2004 (World Bank 2004) and the evaluation of training in World Bank Projects and the World Bank Institute by the World Bank Independent Evaluation Group (IEG) in 2007 have seriously questioned the effectiveness and impact of existing ECB initiatives, particularly those based exclusively on training.

This document is based on the experience of several experts in the IDEA network in supporting institutions in developing countries, programs, and projects in the implementation of results-based management (RBM) in Africa, Asia, and Latin America. It outlines lessons learned and identifies some of the best practices in capacity building in M&E.

Developing ECB strategies based on lessons learned and good practices

The following ECB strategies should be regarded as complementary basic components. They pursue four basic objectives to obtain better results.

- Highlight the importance of M&E for decision makers to encourage their demand for M&E products

- Improve quantity and quality in the supply of M&E products
- Ensure the cost-efficiency relationship of M&E products
- Promote the sustainability of M&E systems and institutional mechanisms

Establishing ECB as part of the implementation of the M&E System

M&E training must not be given simply for the sake of training but rather should be considered a catalyst for implementing M&E systems for results. Incorporating ECB into the current implementation and improvement of the M&E system for results gives meaning to ECB's final purpose, which is contributing to better public policies and programs. Conversely, making progress in the implementation of the M&E system for results provides an opportunity to learn by doing, thereby strengthening and building capacities within a given organization.

The design and implementation of an M&E system should be progressive as part of various M&E components that are gradually implemented. There are no "off-the-shelf" solutions and the approach should be adapted to the needs and particular situation of the existing M&E. The now-standard reference in this field by Kusek and Rist (2004) is showing the way.

Tailor-made by taking into account the institutional and political context, needs analysis, existing M&E systems, and available resources, this plan has been used in various assignments conducted by experts from the IDEA Institute and tested to provide a relevant, effective framework for ECB. A good example of this is the design and implementation of the Results-Based Monitoring System (RBMS) at the Secretariat of Social Development (SEDESOL). The good practices drawn from this experience included the following:

- Close involvement by the organization's personnel in the implementation of the M&E system that boosts know-how within the organization and ensures the sustainability of the system. Staff involvement in the process of implementing RBMS at SEDESOL not only served as a group of resources for other programs in the secretariat but also for developing the National M&E System, currently be-

ing designed and implemented by the Civil Service Secretariat with the support of the National Evaluation Council for Social Development Policy and the Secretariat of Finance and Public Credit.

- Participation by all the actors at all stages of the design and implementation process of the M&E system to guarantee their support and ensure the success of the project, thereby achieving greater sustainability.
- Designing and implementing the M&E system through a pilot phase in 4 of the 27 programs to learn from the experience of boosting the organization's capacity to expand the system, which would therefore result in less dependence on external resources.
- An ECB approach based on MR principles, including focusing on results and a process that involves evaluating needs, formulating strategies to boost M&E capacities, designing and implementing action plans to boost capacities in M&E, and M&E of the implementation of strategies and plans of action to boost M&E capacities and enhance training effectiveness.

The lessons learned include the following:

- It is important to clearly understand the needs of all those involved in terms of data, information systems, reports, and the like. Often the clients themselves do not know exactly what their specific needs are. They regard information as a free asset and have infinite needs. Or else they have a narrow perspective based on specific short-term needs linked to their role in the organization, with little concern for the results. It is the advisor's responsibility to help clients understand their specific needs and, where necessary, suggest different options as a starting point.
- Acquiring a good knowledge of the existing M&E system can be a challenge, particularly in middle-income countries, where the ministries have developed a variety of M&E mechanisms and information systems over the years. These systems are spread over various levels of the organization, without much coordination, harmonization, or even dis-

semination of the information. For example, the diagnosis conducted by SEDESOL in 2004 identified 33 information systems in the organization. These systems did not always provide relevant information for strategic decision making, or they were difficult to interpret because they did not always meet the needs of the users of this information.

- A diagnosis should try to evaluate the organization's M&E capacity and build an M&E training plan as a central component of the project for implementing the M&E system to ensure the personnel's participation in the entire process, maximizing the number of "champions," and limiting the number of potential "opponents."
- Program personnel involved in the design and implementation of an M&E system sometimes may see the new system as a source of additional responsibilities and work that may cause frustration or indifference. They therefore should offer incentives to guarantee the full participation of those involved and of the personnel throughout the process.
- Boosting M&E capacity requires bearing in mind the capacity of the other pillars of MR⁴¹. Problems found at the M&E level often result from a limited design and implementation in other pillars of MR. At SEDESOL, the logic of the logical frameworks was reviewed; performance indicators were selected on the basis Clear, Relevant, Economic, Adequate, and Monitorable (CREAM+) criteria; and goals were based on Specific, Measurable, Achievable, Relevant, and Time-bound (SMART) criteria. In many countries, M&E is not seriously considered by many actors until the results have consequences for future budget assignments, which means linking the M&E process to the budget preparation process.
- M&E goes beyond technical skills. To serve as agents of change, M&E officers should possess leadership and communication skills to convince their colleagues in technical departments as well as cabinet ministers of their own interest in implementing RBM and insist on carrying out the reforms.

Clearly identify the expected results of ECB

The expected results of ECB should be clearly identified and a consensus should be reached between interested parties.

Performance indicators obviously have to go beyond results to include effects and, whenever they can be measured, impacts. What matters in the end is that the M&E reports and control boards are accessible to strategic (cabinet) and operational (program and project) decision makers and used for policy design, program management, and determining future budget assignments. Decision makers often are confused about results. In Senegal, several training workshops clarified what they meant by results at various levels. For example, the director of primary and secondary education realized that the classrooms built and the teachers trained were the result of building and teacher training projects, respectively, but that they also were input for the most strategic level. But, above all, they should be interested in graduation, dropout rates, and double costs, which were effect indicators, as well as in the success of their former students at the subsequent education level, graduate placement rates, and the development of self-employment on the labor market.

Key results must be accessible to other interested parties for the sake of accountability. In Senegal, a great deal of M&E information was available, but with little coordination and validation from the Head Office of Statistics. This led to frustration among government employees and civil society, which had limited access to these results and the reports. A new National Agency for Statistics and Demography was set up, with greater autonomy and power, and with an information system on national statistics in a Web environment. Using the results chain as an analytical framework helped identify four levels of training evaluation with the corresponding means of verification.

Key factors for effective training

Since training is an important component of an ECB strategy, attention should be focused on key factors for effective

training. Based on the experience of the IDEA Institute, we agree with IEG that

One of the most powerful determinants of training success is the organizational context in which training is carried out. For training to be successful, participants must have the resources and incentives to implement the skills and knowledge they have acquired. When these resources and incentives are not present before training, training must be accompanied by suitably designed interventions in order to take organizational and institutional limitations into account. (IEG 2008)

Experience has shown that demand-driven capacity-building activities are usually more sustainable and effective. They tend to respond to more specific felt needs, and the participants in these training activities tend to be more motivated to learn and then to apply what they have learned. A demand-oriented approach obviously entails designing a training program on the basis of the training needs identified in a participatory fashion during the diagnosis and preparation evaluation. Table 16 suggests a structured approach for measuring performance gaps and evaluating training needs that has been successfully used by the IDEA institute on several occasions.

This also involves the careful selection of participants for each type of training activity. An ECB plan should consider various types of training activities designed to match various categories of participants, objectives, and limitations. Strategic-level participants want to focus on the overview and have little time. M&E professionals want a deeper understanding of concepts and a better command of the instruments and techniques and have more time. The IDEA

TABLE 16:

Training Evaluation Levels

Level	Measures	Common means of verification
Level 1	Participants' satisfaction	End of course questionnaires for participants.
Level 2	Learning products	Post tests, compared with pretests.
Level 3	Performance of change of effects	Multiple, including observation, and interviews with participants, colleagues, and supervisors.
Level 4	Organizational impact/results	Multiple, including comparisons with organizational baseline, performance measure, surveys, and interviews with key informants.

Source: IEG 2007.

Institute has therefore organized one- or two-day sensitization and high-level discussion seminars with a small number of cabinet members, delegates, and high-ranking donors. For professionals with sufficient time and money who require official recognition of their skills in the form of an U.S. university diploma, Laval University has organized a Diploma Course in Management for Results, which includes an M&E certificate. To make them more accessible, these courses are offered in the three languages (English, French, and Spanish) and in various locations (Dakar, Dar es Salaam, and Panama City). For professionals with limited time and money, the institute has organized several tailor-made technical workshops at its regional centers or in the country interested in specific M&E issues based on expressed demand. What is astonishing is that most of the participants in these workshops secure their own financing and demand is growing, which proves that a market-oriented approach can work.

Training a critical mass of government officials in M&E has several advantages for a government. It offsets staff rotation and burnout by maintaining a minimum of ECB in M&E units. This creates a common approach among participants who share the same vision, language, and tools and who therefore can communicate better with each other in the community of practices. In several countries, the government asked each ministry to assign part of the budget to building capacities in MR, particularly in M&E and to send their executives on the same training course to create this critical mass and to create momentum in terms of mobilization toward reforms for MR.

A key factor in effective training is choosing a teaching strategy that has been adapted to the context, the participants, and the expected objectives. The institute's experience suggests that participants appreciate several teaching approaches, a minimum of readings, and a maximum of participatory methods. Instructors, therefore, not only require M&E skills and field experience, but also teaching skills. The institute, with Laval University, uses a skills approach whereby a group of skills are identified for each target training group, together with the current and target level of command for each skill. A training plan subsequently is designed to establish links between the training modules and their contribution to improving skills.

In terms of training methods, executive training requires distancing oneself from theoretical training and fo-

cus on practical training. This does not mean, however, that training is transformed into a collection of anecdotes. There must be a solid analytical framework, like the 10-step guide. Participants must learn how to use evaluation instruments professionally to establish their relevance and credibility. However, practical training methods use many examples that are presented and discussed through debates with professionals with experience. Success stories are discussed in working sessions, which helps participants develop their own M&E systems and tools during training. Some groups take study trips. For example, the Vietnam delegation traveled to Canada to learn more about measuring municipal performance; and the Benin delegation visited Panama to learn more about large project M&E. Complementary training strategies include participation in practice communities and meetings with international associations such as the Global IDEAS Conference.

Selecting instructors is another key factor in effective training. Any training process is based on the instructor's ability to elicit the participants' interest and to encourage them to keep listening and adopt the right mental structure to accept new concepts, tools, and approaches, which will effectively transfer their know-how to learners. This requires a rare species of instructor who combines a high level of education with an analytical framework and related concepts and has a good grasp of evaluation instruments. This person has extensive field experience in a variety of institutional and cultural contexts so that he or she understands the content, has the relevant teaching skills, and, lastly, develops the human skills required to be able to handle high-level participants with a variety of characteristics. Academics tend to be, well, academics. Researchers tend to focus on advanced technical subtleties of their own interest and the few peers they do have are too advanced to be relevant to most M&E professionals working in the context of a developing country. Long-term consultants and practitioners tend to talk excitedly about their own valuable experiences but may be out of date as regards methodologies and information and communication technologies (ICT). Civil servants and national and regional advisors (1) may have problems distancing themselves from the specificities of the context and the tools they know, (2) are not necessarily experts in a variety of modern M&E instruments and approaches, or (3) do not necessarily have the teaching skills required of a professional instructor. So, the second option

is sometimes to use a team of complementary instructors and ensure that they work together so that training is not just an unsystematic collection of uncoordinated, individual contributions.

A key challenge to the M&E community is how to increase access to quality M&E training. For example, in Canada, the International Programme for Development Education Training (IPDET) annually trains approximately 275 participants in English. This institute trains M&E professionals who will act as champions of change when they return to their jobs. Most M&E practitioners are unable to receive IPDET training in Canada. An initial solution would be to reproduce IPDET in various localities and languages. Two similar initiatives have been carried out: SHIPDET in Shanghai, offered twice a year, and Czechdet in Prague, offered on an occasional basis, both in English. IDEAS and the IDEAS Institute are discussing the possibility of organizing a DIPDET in Dakar in French for Francophone Africa and a PIPDET in Panama in Spanish for Latin America.

Another solution would be to design a Trainer Training Program. The institute recently used this formula in the Democratic Republic of Congo. During the first two-week phase, IDEA experts trained 50 participants. The best were selected for a second phase of Trainer Training, during which they expanded their knowledge and developed their teaching skills. The recently trained instructors then gave the following training courses with the technical support and advice of the IDEA experts. The result was that in six weeks, 100 participants and five instructors were trained, who are now qualified to repeat the training at both a central and a decentralized level.

A third solution is to build alliances with regional and national training institutions to develop executive training programs leading to certification. The IDEA institute has developed this kind of alliance with the eight countries of the West African Economic and Monetary Union (Senegal, Guinea Bissau, Cote d'Ivoire, Mali, Burkina Faso, Togo, Benin, and Niger) to train public finance officers at the *École Nationale des Régies Financières*, in Ougadougou. It also provides training in project management at the Tanzania Public Service College in Dar es Salaam.

Another factor in effective training is follow-up. In IDEA's experience, the most effective, sustainable strategy is to design and implement a medium-term (two to three

years) ECB in which the institute uses a combination of training, technical assistance, and support to help a national team improve data quality and information systems. Activities are organized according to a flexible road map that includes the production of intermediate products required by the national team, such as progress in policy matrix objectives to be prepared for the World Bank on a six-month basis; review missions, national development plans, or enterprise research planning evaluation; impact evaluation studies of major programs; annual performance reviews aligned with the next budget preparation; and processes to update the midterm fiscal framework. The program usually begins with training, followed by technical support missions and subsequently in-service training programs.

Improving M&E information and knowledge management systems

Information systems can be an excellent entry point for ECB. They focus on data-gathering, processing and analysis and the dissemination and use of information as well as data quality. All these elements have enormous implications to permit full evaluations and their credibility.

A good practice is a series of new applications that have been developed by various organizations and firms such as automated control panels. The IDEA Institute has developed control panels at the sectoral level (such as the Secretariat of Social Development in Mexico) and at the program level (such as the National Land Administration Program in Panama). It has developed an attractive computer solution called e@satisfaction that provides a reliable means of measuring, processing, and analyzing the quality of the services provided on the basis of online surveys and of obtaining rapid feedback on the Web in a user-friendly way to enable decision makers to conduct their programs or projects better.

Another good practice is to establish knowledge management and learning units in the organizations. Many national and international organizations have realized that experiences, lessons learned, and good practices were not shared among members of the same organization and were lost once these members retired, were transferred, or left the organization. These units play a crucial role in the preservation and dissemination of institutional memory by gathering available information on M&E, correctly documenting metadata, storing data safely, and providing selective access

to authorized users. They can make it easier to share experiences among internal and external experts, participate in impact evaluation studies on large programs to learn about failures and success, orient future policies and programs, and train the organization's personnel in M&E methods and their results. In their new mandate, these units can take advantage of advances in ICT and electronic documentation management systems (EDMS). The IDEA Institute has helped implement these systems in various institutions to protect an enormous amount of administrative and technical documentation with user-friendly retrieval methods. The following are lessons drawn from various experiences at the IDEA Institute:

- In several large public bureaucracies, information is scattered, meaning that it is not always easy to know where to find the information, who has it, or how it is used. It is important to identify all information sources, whether they include sophisticated information systems, Excel spreadsheets, or handwritten reports. This information can be identified through interviews but also may appear as summaries in the design and implementation process of the M&E system.
- The point should not only be to try to determine where the information is found but also how this information was gathered or calculated. This is extremely important in ensuring the quality and integrity of the information to be incorporated into the system. Otherwise, the information might be outdated or, worse, incorrect. The Millennium Challenge Account (a U.S. initiative administered by the Millennium Challenge Corporation) sponsored data quality evaluations in various countries. This initiative has been welcome, as have the efforts to revitalize and improve national statistical systems within the framework of the 21 initiatives of the Paris declaration (the Paris 21).
- An information system for monitoring and evaluation—managing for results (M&E-MR) should be designed by computer experts under the supervision of M&E specialists. M&E specialists should listen to computer specialists who are aware of the latest technological developments particularly as regards database management terms and business in-

telligence tools to design a user-friendly M&E information system. Nevertheless, M&E specialists must focus on the objective, namely, that the system should work, and be robust, profitable, and regularly provide a minimum amount of the information requested by decision makers. There are too many horror stories and the world is full of ineffective, costly white elephants and database cemeteries. It is useless to have a complex information system with the latest technology if it is not used for the program's everyday activities or in decision making. Computerizing a poor information system only achieves faster GIGO (Garbage In, Garbage Out). It is more useful to begin on a small scale but slowly and surely update and improve the M&E information system.

Sensitization and dialogue forums on policies and professional associations

The IDEA Institute has been associated with the establishment, participation, and strengthening of various types of platforms and mechanisms for sharing experience that contribute to ECB. At the country level, the use of the Joint Progress Review (JPR) has increased between the government and donors, sectoral thematic groups, and forums to create a common strategic vision. In Vietnam, the JPR mechanism implemented for the discussion of the 135 Program, a large-scale program to alleviate poverty, has contributed to a better policy dialogue between government, making the latter more effective and efficient. In Cambodia, the Poverty Forum proposed by the institute helped share a great deal of scattered information on poverty and contributed to a policy dialogue between public organizations and a vibrant NGO community. Sensitization initiatives such as brief seminars or cabinet retreats may increase acceptance at a high level and create a momentum for the implementation of M&E-MR under national leadership. The organization of top-level conferences at both the national and regional level, involving a combination of ministers, legislators, high-level technical personnel, and consultants may help to compare national experiences, carry out benchmarking, and orient participants toward a common understanding and strategy. Several communities of M&E practices have arisen at the national, regional, and international level, with varying degrees of success

and sustainability, which, if properly handled, can certainly contribute, in a profitable way, to sharing lessons learned and best practices. Last but not least, professional evaluation associations have a significant role to play in ECB, and IDEAS (as the only international association for development evaluation) has a unique niche and a contribution to make to sharing knowledge through geographic regions and linguistic and cultural zones.

Building sustainable M&E units

Beyond more effective training, improvements to the information system, sensitization, policy dialogue forums, and professional associations, ECB seeks to build sustainable M&E units. The term “unit” is used in a generic sense and may cover a variety of international systems according to the country and organization involved.

To improve the performance and sustainability of M&E units, the IDEA Institute has encouraged management approaches, all of which are designed to improve institutional performance and results. The National Institute of Standards and Technology (NIST) Balridge National Quality Program (National Institute of Standards and Technology 2009; Balridge National Quality Program 2009), the equivalent in Quebec, called the *Qualimetre (Mouvement Québécois de la Qualité 2009)*, and the Balanced Scorecard, initially developed for the private sector, were all adapted to the public sector (Kaplan 1999).

When applied to a public organization as an M&E unit, institutional support focuses on the mission and target groups of the M&E unit and its balanced consideration of human resources, financial resources, and processes of the unit. Its links have proved to be extremely useful if the aim is to develop the institutional sustainability.

In fragile states, an important consideration for institutional sustainability involves whether the M&E unit is inside or outside the public sector. If incentives and morale are low, for example, implementing the unit within ministries will be counterproductive. Implementing M&E units within external project implementation units (PIUs) is not a sustainable approach and could seriously undermine sustainable national institutions by attracting the few competent human resources available in the public sector through higher salaries and better working conditions. One approach used by the Institute in Cambodia has been to support a poverty monitoring technical unit in the Planning

Ministry, established with approximately 20 local young professionals, each acting as a focal point with a sectoral ministry. This arrangement achieved a reasonable balance between the need to produce M&E products in the short term and the desire to increase capacity.

External evaluation and accountability

In addition to the key role of M&E units within the sectoral ministries, the institutional device should include other public institutions as well as others from the private sector and civil society. In addition to organizations for ex ante and ex post internal control, an external public evaluation organization should have the legal independence and the legal power to obtain needed information as well as the human and financial resources to produce quality work. The experience of several industrial countries such as Canada has proved its significant contribution to the efficiency and effectiveness of public spending. Behind the organization of powerful auditing, the culture of results develops on both the demand and the supply side through the implementation of a management-by-results system in the public sector with performance contracts, and through the following: (1) the public accountability of the results of each actor in comparison with his or her goals and commitments, (2) the dissemination through the media and members of civil society to the population of the follow-up provided by justice when necessary, and (3) the development of evaluation capacity by the private sector and universities, and the like. The demand from both the private sector and civil society for accountability from the public sector and the demonstration of the public value created goes hand in hand with improving the quantity and quality of accountability.

In a mature M&E system, the limits of the two pillars evolve. This means, in other words, internal monitoring for piloting programs and external evaluation for accountability. The incorporation of a results culture into public organizations leads to the development of internal evaluation perceived as a tool for improvement and one that will develop an information-rich and learning organization. In addition to the classic monitoring reports on physical and financial performance and base, intermediate, and final evaluations, other tools have been developed such as results monitoring, control panels and rapid warning systems, annual performance reports, performance auditorships, qualitative and quantitative or combined evaluations (Q2 approach), and

the like. The natural framework for progressively including all these components, modules, and tools is the national M&E system that is constructed as a result of a coordinated, midterm M&E plan.

Conclusion

Building M&E capacities involves several challenges. However, experience based on lessons learned and good practices shows that it is possible to achieve significant progress by using various mechanisms and tools in a complementary fashion. Training should be decentralized and made more effective and diversified. Investing in complementary technical assistance, support for information systems, and certain sensitization and policy discussion forums and professional exchange platforms as part of a strategy and a road map defined at the national, sectoral, and local level can help the M&E system to make a great deal of progress toward effectiveness and sustainability and contribute to professionalizing the evaluator's work. This process of continuous improvement must be perceived as a key element in the overall reform of the public sector toward MfDR. This requires strong leadership and greater commitment from all the actors in developing countries and donors to achieve a culture of results.

Gilles Clotteau
Marie-Helene Boily
Sana Darboe
Frederic Martin⁴²

Session 4 Notes

- ³⁷ The meeting was sponsored by the evaluation network of the OECD-DAC. The evaluation principles were included in *DAC Principles for Effective Aid* (published in 1992). Many of the principles discussed then also had been discussed at and were later integrated into international and regional evaluation associations, such as the American and African Evaluation Associations (AEA and AFREA, respectively), the Evaluation Cooperation Group of the multilateral development banks, or norms of the United Nations Evaluation Group. The underlying principles are the same, although the terms used may not be identical.
- ³⁸ The UNIDO report defined “institutional capability” to include services to be performed, demand for these services, capabilities (skills in necessary quantity and quality) to meet demand, hardware (including facilities, premises, and instruments), methodology, legislation, and management and coordination functions. These features, considered necessary for an institution to function, were integrated into project design training and the annual evaluation reports of UNIDO at the time.
- ³⁹ The work program and budget can be found on <http://www.ifad.org/evaluation/whatwedo/wp/index.htm>.
- ⁴⁰ The authors thank their colleagues Juan Abreu and Sylvain Lariviere for their suggestions and comments on this document.
- ⁴¹ IDB/PRODEV considers the following five pillars of MR/MfDR: Strategic planning, Budgeting for Results, Public Finance Management, Program and Project Management, and M&E (Garcia López 2008). The OECD/DAC/MfDR considers the following five pillars of MfDR: Leadership; Evaluation and monitoring; Accountability and Association; Planning and Budgeting and Statistics (MfDR 2009)
- ⁴² For more information, see IDEA's Web site, www.idea-international.org.

Session 5: Institutional Arrangements and Related Policies to Ensure the Quality, Access, and Use of Performance Information

Challenges in Monitoring and Evaluation: An Opportunity to Consolidate the M&E Systems

The litmus test of the success of a monitoring and evaluation (M&E) system lies in the quality of indicators that are used to capture a dimension or an attribute to show the results in the form of an assessment of the performance in a particular aspect of governance or public service delivery. The classical view of the quality of indicators in M&E systems is to ensure that they are objective and adequate to reliably measure the impact of an input or intervention. The indicators are supposed to measure and reflect change over a time and hence conform to the criteria represented by the acronym CREAM, which refers to the following:

- Clear: Precise and unambiguous
- Relevant: Appropriate to the set goal
- Economic: Available or computable with reasonable cost
- Adequate: Provides sufficient basis to assess performance
- Measurable: Quantifiable

The well-known SMART test uses a quick reference to determine the quality of indicators by ascertaining whether they are

- Specific
- Measurable
- Attainable
- Realistic
- Time-bound

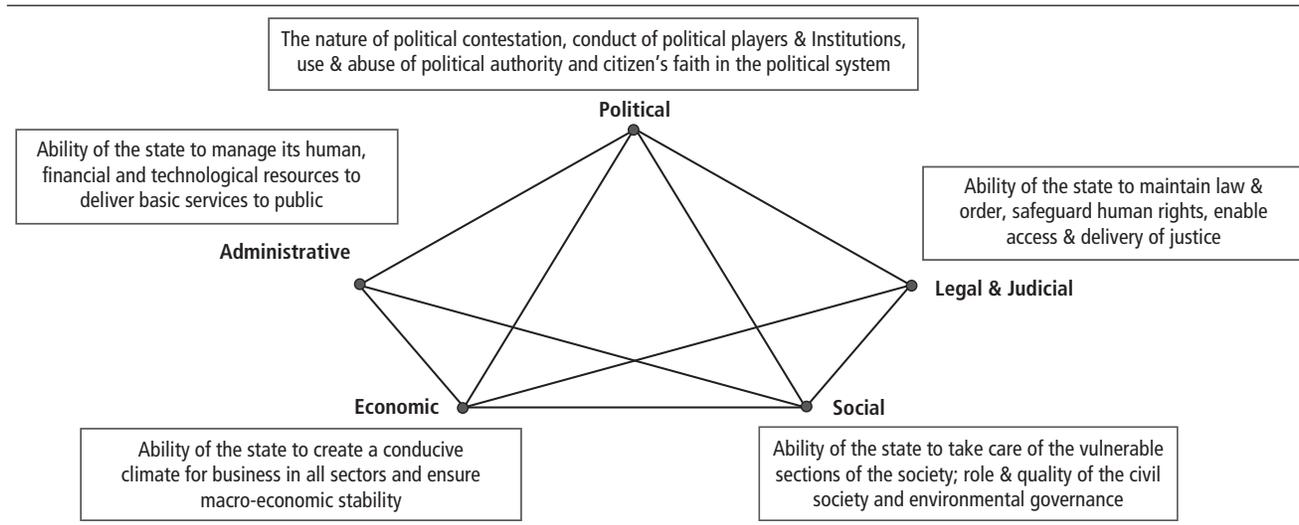
Essentially, these concepts provide a theoretical framework within which the quality of an indicator can be verified against the given criteria. But the mechanism of how to do it remains a challenge. In the context of public administration, the main actors associated with this mechanism, very broadly, are the state, market, and civil society. The state provides a political, legal, and economic framework, whereas the market creates opportunities for people, and the civil society mobilizes people's participation. Therefore, these three groups of stakeholders need to be brought into the process of feedback and consultation on the quality of indicators. This arises out of the need to move away from the M&E of mere processes, inputs, and outputs to an emphasis on measuring results, outcomes, and impact and that is the goal that should determine the quality of indicators.

Case study

When a recent study was undertaken to prepare a tool to monitor the state of governance in India,⁴³ the Centre for Good Governance initially took 271 indicators to shape up and describe five dimensions of governance (see Figure 23). These indicators, then, were put to a rigorous quality test based on the SMART criteria. The study conducted exhaustive stakeholder consultations at a number of levels. In the end, the number of indicators was reduced to 164, and were modified substantially, based on information from the stakeholders on availability and reliability of the data, frequency and dates of collection, the agency's background, and the like. The stakeholder consultations included politicians, bureaucrats, experts in the field, and ordinary citizens.

FIGURE 23:

Five Dimensions of Governance



Source: Author.

The mechanism adopted was to validate the robustness of indicators on a laid-out criteria by obtaining the opinions of the concerned stakeholders, including the judgment of experts on whether or not a given indicator would measure up to the criteria of quality. This qualitative analysis may not completely eliminate the bias in such assessment, but it certainly improves the validation process due to stakeholder involvement.

The methodology to identify the indicators included a review of available literature, different approaches adopted by different organizations to define the basic issues, and then a broad list of indicators that were put through a process of scrutiny and validation. Table 17 gives an idea of the

number of indicators that were considered to arrive at the composite score on the five dimensions mentioned above.

The approach adopted to validate the quality of indicators was based on the findings from field-testing and inputs from various stakeholders, including common people, experts, and professionals. Field-testing gave a lot of insight into the quality of indicators as quite a few things could be placed in perspective only after this “ground-truthing.” Some illustrations of the reasons for modifying or deleting an indicator are given in Table 18.

The quality, reliability, access, and use of information produced and used by M&E systems can be optimized by introducing a Citizen-Centric System of Social Account-

TABLE 17:

Governance Indicator

Governance dimension	Number of Indicators					In the final framework
	In the Initial Indicator Framework	Retained after testing	Deleted after testing	Modified/ merged after testing	Newly added after testing	
Political	46	18	19	9	—	24
Legal/Judicial	53	23	22	8	4	33
Administrative	88	29	35	24	11	52
Economic	39	19	19	1	5	25
Social	45	28	14	3	0	30
Total	271	117	109	45	20	164

Source: Author.

TABLE 18:

Modification of Indicator

Dimension of Governance	Indicator under validation	Source of data	Action taken	Reason
Political	Proportion of bogus voters in the electoral list	Secondary	Deleted	Found difficult to ascertain.
	Percent of women candidates fielded in election	Secondary	Deleted	Outcome—the proportion of women elected is not captured by this.
Administrative	Proportion of women in government service in senior cadres	Secondary	Modified	Modified as “Aggregate proportion of women in government service.”
	Percentage of the government budget earmarked for women	Secondary	Deleted	Difficult to collect the information as gender budgeting is not done.
Social	Percent of habitations not having primary school within one kilometer	Secondary	Modified	Percent of parents who say that the primary school is not easily accessible.

Source: Author.

ability in which, both in collection of information and its dissemination, the citizen is intensely involved. Some distinct tools have proven reliability to improve the effectiveness of M&E systems by actively involving the people. Some of these tools are as follows:

- Community scorecards
- Citizen report cards
- Social audits
- Participatory expenditure tracking, planning, and budgeting

Community scorecards

The community scorecard is a tool to monitor and evaluate the services provided by the state or public service providers in a joint and participatory manner. This tool is powerful enough to influence the quality, efficiency, and effectiveness of these services, particularly, at the local level. Community scorecards recently were developed by the Centre for Good Governance in health service provision and school education.⁴⁴ These are used by the community as a basic unit of analysis. They focus on monitoring public service delivery at the grass-root level and generate relevant information through focus group interaction in a participatory manner that provides feedback and facilitates a joint decision on the quality of service between the recipients and the service providers. This technique usually brings about a clear empowerment of community and promotes good governance. Such scorecards have been tried successfully in Ghana, Malawi, and Zambia. Various features in the construct of community scorecards are as follows:

- Preparatory groundwork
- Input tracking, sharing information, verifying record, comparing supply-side data, and finalizing the indicators
- Performance scorecard by the community—the scoring is done by community in the given format
- Self-evaluation by service provider—this involves giving score
- Interface meeting and action planning—feedback is given on evaluation of the scorecards
- Institutionalization—following up a repeating scorecard with benchmark data
- Consensus—finally, the community action plan involves both indicators given by the service providers as well as community to arrive at a consensus with mutual agreement

Citizen report cards

Whereas the community scorecards involve the community as a unit, citizen scorecards are administered through a survey. They can be used for local, provincial, or national assessment of services particularly in services like health, water supply, and utility. Citizen surveys help in benchmarking and comparing the level of services or activities for the government or unique service provider at a rating scale and help engage citizens and empower the monitoring process.

Social audits

This is a unique mechanism of creating social accounts involving the community and to audit a particular parameter

or activity. The use of social audits in the M&E mechanism can be effective in program implementation. This practice is being encouraged in important pro-poor programs, such as the National Rural Employment Guarantee Scheme and National Rural Health Mission. The essential steps in social audit involve training of social auditors, preferably recruited from the community, creating an environment of acceptability by the establishment for social auditors to operate, and a system of fixing responsibility and award of penalty or punishment for any irregularity. The key to successful implementation lies in bringing quality to the tools of measurement and design of indicators, so that only relevant information is captured.

Case study

In India, the social audit is ingrained in the design of the National Rural Employment Guarantee Scheme (NREGS) by a statutory provision. Therefore, all state governments are required to facilitate this while implementing the flagship program to provide guaranteed wage employment to the rural poor. Various models can conduct social audits. The state of Andhra Pradesh adopted a new approach and trained a large number of young volunteers from the wage earners' families and empowered them to conduct social audit. The government provided an enabling environment to facilitate the following processes.

- Both proactive and on-demand disclosure of information to the wage seekers under the Right to Information Act to give them access to muster rolls, pay orders, measurement sheets of civil works, engineering cost estimates, and sanction orders.
- Creation of requisite infrastructure for a two-day training for at least five volunteers constituting a social audit team in each village by providing venue and space for the training, training material, food, transport, and other logistics.
- Facilitating the verification on ground and comparison with official records to conduct social audit and summarize the findings over a period of four days.
- Hosting the public hearings in the presence of all officials and stakeholders to discuss the findings of social audit and to take immediate remedial action in case of irregularities.

The entire exercise has taken the shape of a massive social movement and, at the same time, all the components of the M&E system are standardized with well-designed indicators and a replicable process. The system has a Web-enabled support through a well-managed portal that has brought transparency and improved accountability. This has brought to light a good number of cases of irregularities in funds flow, and corrective action was taken. Another important aspect is the empowerment of the poor and enabling their participation and ownership in the developmental process.

Participatory expenditure tracking, participatory planning, and budgeting

Participatory expenditure tracking, participatory planning, and budgeting involve creating people-oriented accounts of public expenditure in schemes and programs and using participatory techniques in budget preparation. This is another example of involving people to bring useful information as input in decision making. This can serve as a monitoring tool, which is the basic function of a budget. The pioneering examples of Porto Alegre in participatory budgeting and Uganda in expenditure tracking are the role models in this.

What succeeds?

Intense civic engagement is the key to success in M&E as it improves the accountability system. It is important to create an overall atmosphere of mutual trust in government to citizen dialogue. An environment of social accountability and specific voice mechanisms for people to air their views needs to be incorporated.

Different networking arrangements, systems of standardization, and quality assurance in public service delivery such as the instrument of citizen's charter, simplification and demystification of processes and procedures, and unfettered access to information are necessary ingredients of success.

Social accountability is a nonelectoral system that still provides a straight control to bring sufficient accountability to the actions of government. Such mechanism can be supported by the state, or the citizen, or both and usually are bottom-up in nature. The use of a robust e-governance system and grievance redress mechanism to improve the interface between the citizen and the govern-

ment and to bring them close to each other is necessary in this context.

Poor quality of information for decision making in public policy results in negative fallouts. These fallouts include poor targeting of expenditures in various schemes and plans, corruption and leakages in the system, appropriation of the subsidies and government assistance by the undeserving, wasteful expenditure, and monitoring unimportant items at the cost of important ones. The core of decision making gets obscured if adequate information is not available. In such cases, the decisions are taken on the information brought by a close group of people who may be from a political or bureaucratic or business interest group.

Many countries have implemented policies and laws on “Access to Information.” The difficulties and challenges they each face differ from country to country. The challenge is to draw a line between individual privacy and freedom of information. A brief mention of the challenges faced by India includes a lack of adequate and effective record management systems, resistance of bureaucracy to part with the information, ignorance of people of the procedures to access information, and lack of training to make use of this law, including government officials and civil society. The implementation aspect of the Right to Information Act has provided a strong system of eliciting compliance in India by the government officials and public authorities, because hiding information attracts severe fines and reprimand.

Similar laws exist in a large number of countries now. Australia has an Ombudsman; Japan has a system of appeal to an “Information disclosure Board”; the United State has a system of appeals in case of denial of information to the concerned agencies, with a provision of judicial review; Mexico has a sound information regime; and Nepal has the Whistle Blower Protection provision.

In India, the regulatory framework for access to information generated by M&E systems is good enough in the sense that a citizen can ask for any information. But its effectiveness lies in the voluntary and proactive disclosure of information by the public authorities. The government should help citizens learn where the information resides and what he or she should request.

The voluntary and proactive mechanisms of disclosure have not fully matured in the government departments. The data in respect of such evaluation are not brought in-

to the open unless such evaluation is done by a third-party public agency like the comptroller and auditor general, Planning Commission, or watchdog bodies of legislature or executive.

The national system of M&E in India is multilayered, as indicated in Figure 24, and the information to civil society is given in the form of periodic and annual reports of various committees and ministries, national sample surveys, and the like.

The spread of information in the public domain is increasing, with easy access due to use of information technology and Web sites. A lot still needs to be done to create a baseline and present analysis over a period of time rather than giving barebones data. The Planning Commission of India and the Ministry of Statistics and Programme Implementation collect regular data on the government’s various programs and schemes. Most of these systems are at the macro level and there is a glaring gap between what is available at the national level and what the people at community level possibly would need. But this dissemination has to go far and beyond the federal and state levels up to the last villages and municipalities. This is possible only with intense engagement of community through social accountability mechanisms at the local level. The key to success, therefore, lies in a bottom-up approach in M&E, rather than a top-heavy macro management system. It may be said in the Indian context that the direction is positive and the atmosphere of transparency with citizen involvement is gradually finding a base at the grass-roots level.

Rajiv Sharma

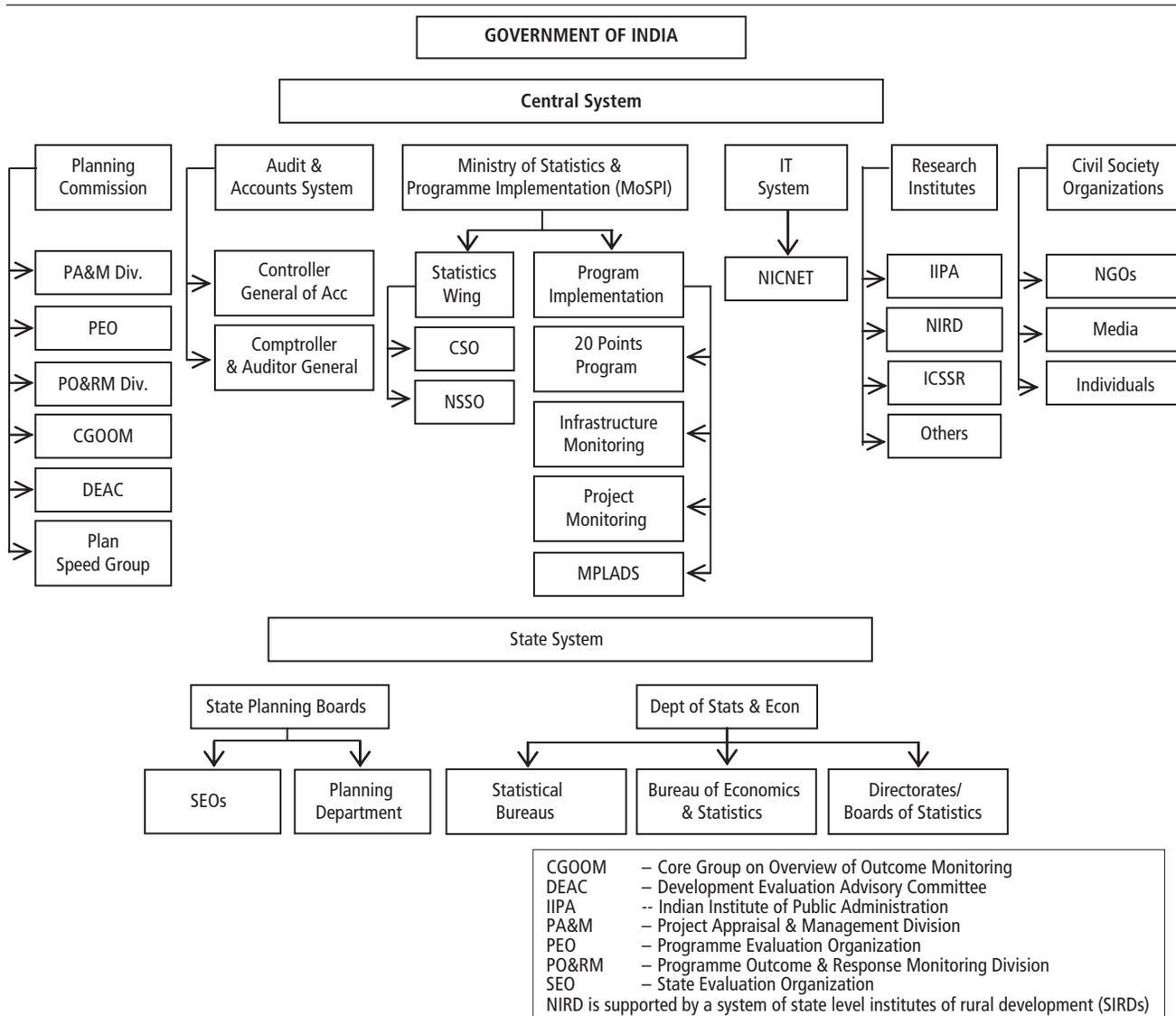
*Director, General and Project Adviser,
National Implementing Agency,
Centre for Good Governance (CGG), India*

Performance Information: How Web 2.0 Can Help

Measurement and evaluation (M&E) is challenged by a paradox: the ability to measure the effectiveness of a government program is negatively correlated with the ease of measurement. Inputs (what government consumes) and outputs (what government produces) are relatively straightforward to measure, but they do not give an indication of the effectiveness of the intervention. Measurement of impacts and outcomes is much more challenging—the tech-

FIGURE 24:

Structure of the Government M&E System in India



Source: Author.

niques to measure are more complex and take longer, the costs of measurement are higher and the challenges of attribution of outcome to intervention are well understood. As a consequence, evaluation of a specific intervention can rely on indicators that are readily measureable, although they may not reflect the outcome, or the underlying policy intent of the intervention.

Recent developments in information communication technology (ICT), in particular the phenomenon of Web 2.0, may provide a response to this paradox and improve the quality of M&E. This paper describes a framework for

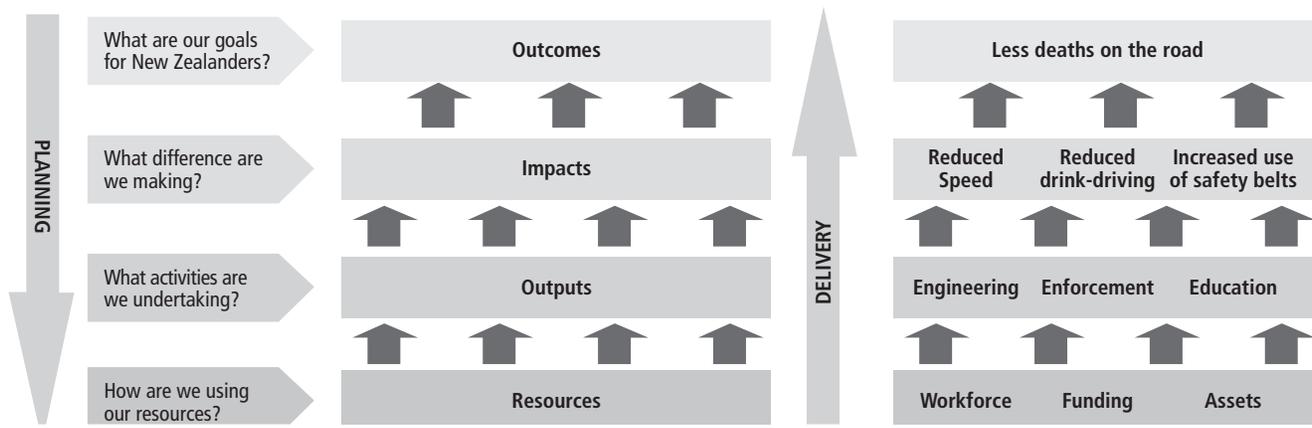
integrating a top-down centrally driven approach to performance management with bottom-up community-based reporting of citizen experience.

What is performance information and how can it best be measured?⁴⁵

This section describes the approach that New Zealand has taken to measure government performance. Performance in the public sector can be assessed at four levels: the consumption of resources by a government agency, the outputs (products) produced by the agency, the impact of those

FIGURE 25:

Linking Outputs, Impacts, and Outcomes



Source: Author.

outputs, and the resulting outcomes. The distinction between these levels and examples at each level are shown in Figure 25.

Measurement at each level requires different techniques, both in gathering the data that can be used to assess performance, and in the attribution of results to specific actions, or programs. The characteristics of measurement at each level are shown in Table 19.

In general, the lower levels are easier to measure, and the higher levels are more meaningful for evaluation purposes. The accountability regime for delivering outputs involves a range of legislative and assessment levers.⁴⁶ For lower level measurement, the New Zealand public management system has adopted a devolved approach, holding agency heads completely accountable for delivery of outputs.

A number of countries use a scorecard or dashboard approach to assess the performance of an agency. Three examples are as follows:⁴⁷

- New Zealand: Performance Improvement Framework
- Canada: Management Accountability Framework
- United States: The e-government scorecard from the previous administration has been superseded by the Performance Assessment Rating Tool

A similar framework, Government at a Glance, was adopted by the Organisation for Economic Co-operation and

Development (OECD) in their ambitious measurement and evaluation program. The initial report, published in October 2009, did not attempt to measure the higher levels of outputs and outcomes.⁴⁸ While this proved successful in focusing attention on improving individual agency performance, the bigger challenge—evaluation of the impacts and outcomes—has proved more elusive.

Trends in ICT

Web 2.0

The widespread use of the Internet has resulted in the development of software that enables cooperative working, and creation of new tools and approaches that enable groups to manage and share information. These tools have been given a generic title of “Web 2.0” or “social media”; examples include the following:

- Collectively authored information sources, wikis such as Wikipedia
- Social network sites such, as Facebook and LinkedIn
- User generated content sites such as Blogs, YouTube, and Flickr

The magnitude of these changes and the effect on cooperative working is set out in *The Wealth of Networks*,⁴⁹ which describes how social production is reshaping markets, while at the same time offering new opportunities for innovation and creativity. Clay Shirky (2008) describes how technology

TABLE 19:

Measurement Techniques

Type of Result	Level	Focus On	Examples of Common Measures
END or FINAL	Impact or Outcome	Aggregate improvement More equity Cost effectiveness Unintended effects	Greater health/wealth/happiness Less difference across deciles/areas Fewer deaths/accidents/kids in care Cost per unit of improvement in outcome Increased welfare dependency, risk, and so on
INTERMEDIATE	Impact or Outcome	Cognitive change Behavior change Risk reduction Lifestyle change Survival Unintended effects	Percent aware of risks/able to use new idea Percent investing/saving/quitting/ working Fewer drunken drivers/"bad" incidents Percent in jobs/new career/crime free Percent alive after 30 days/time event-free Graduates migrating or excessive uptake
NEAR TERM	Impact or Outcome	Completion rate Knowledge retained Reduction in queue Receipt of benefits Incentives changed Unintended effects	Percent finishing/getting qualified/in service Percent core messages remembered Average wait time/number in queue Percent impoverished with more money Percent believing regulatory change matters Higher incident or reduced survival rates
COVERAGE (or Reach)	Output	Coverage Targeting efficiency Access	Percent population in need receiving output Percent in "treated" group who met entry criteria Percent targets who did not access/use service Transit time (or other "big" barrier to use)
QUALITY	Output	Quality of delivery Timeliness Acceptability	Percent output fully meeting specification Percent ministerial/passports/etc. on time Percent who would use again/recommend use
QUANTITY	Output	Volume produced	People receiving training/rehabilitation Cases/complaints processed
EFFICIENCY OF PROCESS	Input Utilization	Efficiency Economy	Real output price trend (inflation adjusted) Price per unit, versus benchmarks Percent prison beds full/maximum capacity used Real input cost trend (for example, per policeman or nurse)

Source: Author.

enables group collaboration in a way that was never previously possible. Another example of Web 2.0 is the widespread use of aggregated rating as a method of assessment, combing the individual views of any product or service consumed by a wide range of people—news stories, books, buyers and sellers on eBay, hotels, music, and even American Idol contestants—to create a collective “wisdom of the crowds.” At its heart, the change is a shift from centralized production and control of information products to a distributed cocreation resulting in a richer and higher quality result.

Government 2.0

The term Web 2.0 was initially popularized by Tim O’Reilly in 2004. Tara Hunt took the O’Reilly tenets for Web 2.0 and applied them to government, and came up with the nine concepts in Box 5.⁵⁰

The last two years have seen the meteoric rise of social media on a global basis, and a significant effect within the political domain, most notably in the election of President Obama on the back of an Internet-savvy campaign, and a raft of open government initiatives from the administration in Washington, D.C.

Technology-based consultation is starting to be adopted more widely in the policy creation process, using social media tools to enable the collective development of policy options for consideration by lawmakers. One of the earliest examples of this was the use of a wiki to enable community input to a new Police Act in New Zealand in 2007.⁵¹

In 2009, the use of these tools increased significantly, for example, the development of the open government policy in the United States.⁵² A similar approach was taken in

BOX 5:

Comparison of Web 2.0 and Government 2.0**Web 2.0**

The Web as a platform
 The long tail
 Data: the next Intel inside
 Users add value
 Network effects by default
 Some rights reserved
 The perpetual beta
 Cooperate, do not control
 Software above a single device

Government 2.0

The government is my springboard
 The long diverse tail of citizens
 Data are tricky
 Going to the edges for feedback
 Citizen community is relationships
 Some rights reserved
 Evolution is an ongoing process
 Trust builds empowerment
 Government on the go

Source: Presentation by Tara Hunt, May 2007.

Australia to create Public Sphere—Government 2.0 Policy and Practice.⁵³

Beth Noveck, director of the White House Open Government Initiative, has described the three values of open government—transparency, participation, and collaboration. Each federal government agency must have a road map for open government, which not only covers technology, but also includes people and operations; agencies are required to develop a schedule for publishing data online in raw structured machine-readable formats.⁵⁴

This approach of sharing data and engaging in open collaborative dialogue enables governments to work with the citizens and the community to tap into wisdom and expertise outside government, strengthen trust in government, and reduce costs.

Open government, open data

One significant and relevant trend is the opening up of government data for use by commercial and nonprofit organizations. Most jurisdictions have laws that provide for public access to information, with exceptions to be subject to specific justification; although this legal position is clear, operational practice is often far from compliant with the law.

Government information is public data. While the government collects and stores a lot of information, it rarely has the time or resources to fully use the data. This means

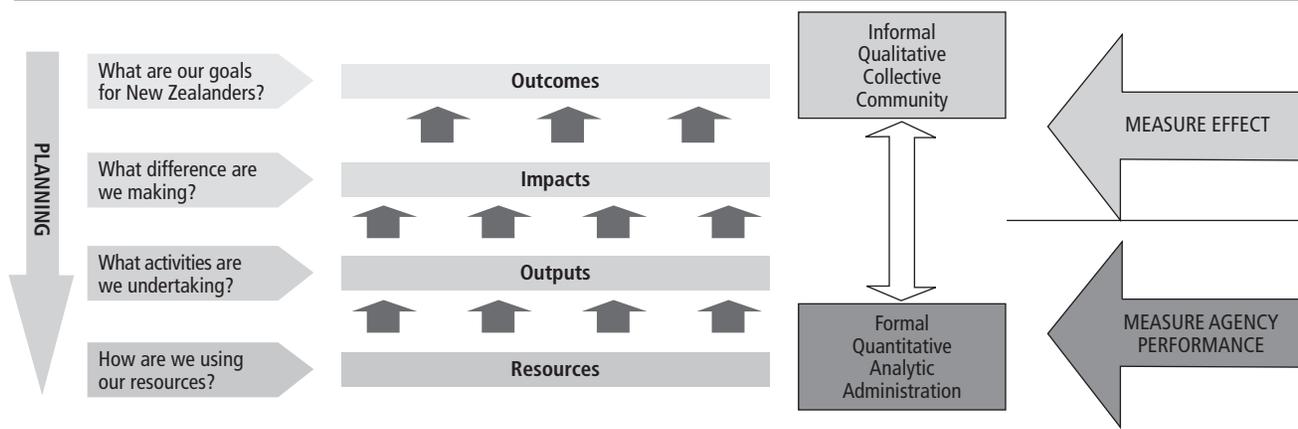
that the priority for government should be to produce machine-readable data (wholesale). Governments have a secondary role in interpretation (retail), which many others also can do—any interpretation of data is only one view. Government will need to shift from being a retailer of data to providing wholesale data—focusing on building an information infrastructure with a focus on the quality of data. This becomes increasingly important in the measurement of the impacts and outcomes of programs—government can provide the platform for civil society to operate as watchdog and innovator.

Open government data have been defined as data that are reviewable against eight principles: complete, primary, timely, accessible, machine readable, nondiscriminatory, non-proprietary, and with an open re-use license. Open data can deliver better citizen engagement, and strengthen trust in government. The Gapminder tool illustrates how official data can be used by the community in a way that never could be achieved within government.⁵⁵ Other publicly available free tools include ManyEyes, Swivel, and DataMasher. The supply of government data is increasingly being cataloged either by governments themselves or by civil society.⁵⁶

What are the implications of Web 2.0 for M&E?

As illustrated in Figure 26, measuring the effect of government interventions is likely to be more informal, qualitative, collective, and community based, than measuring

FIGURE 26:

Different Measurement Techniques

Source: Author.

inputs and outputs of individual government agencies. The characteristics of measurement of the effect of government programs are well aligned with the Web 2.0 collective assessment tools deployed in the commercial world. It is clear that government does not have the resources to effectively measure the breadth of impact of programs, and tapping into the wisdom of the crowds is a potentially powerful tool to achieve better measurability.

Examples from the New Zealand education sector

This section contains three specific examples from the New Zealand education system, ranging from formal evaluation based on legislation to crowd-sourced assessment of teacher performance.

Formal assessment of schools

The Education Review Office (ERO) was established in 1989 to evaluate and report publicly on the education and care of students in schools and early childhood services. ERO investigates and reports to boards of trustees, managers of early childhood education services, and the government on the quality of education provided for children and students in individual centers and schools.

Schools and early childhood services are reviewed on average once every three years. Reviews are undertaken more frequently in cases in which the performance of a school or center is poor and there are risks to the education and safety of the students. Reports on individual schools and early childhood services are freely available to the public.

ERO's findings inform decisions and choices made by parents, teachers, managers, trustees, and others, at the individual school and early childhood level and at the national level by government policy makers.

Based on the findings from individual school reviews, ERO also prepares national reports that evaluate different aspects of the education outcomes sought by government. Recent reports include Progress in Pacific Student Achievement, Teaching of Health and Physical Education, and curriculum materials to support teaching and learning of Te Reo Maori. This approach can be described as systemwide performance evaluation by professionals giving broad performance information on education outcomes.⁵⁹

Publication of data on student and school performance

The use of the data on individual student performance is currently an issue of significant debate in New Zealand. In 2010, the government will introduce literacy and numeracy testing of individual students in year one through eight against national standards, a move that is opposed by some professionals within the education sector.

One area of concern is publication of such data, when combined with emerging ICT tools, means that the national standards can be used to compare the performance of schools through "league tables." The subject is controversial in other countries such as Australia, Japan, and the United Kingdom.⁶⁰ This approach can be described as ag-

gregating individual professional assessments of achievement, to provide comparative performance information.

Community-based assessment of teachers

Rate My Teacher is a Web service that provides students with an opportunity to rate individual teachers. While there are drawbacks to such a service—harsh and unfair comments often can emerge from anonymous assessments—it provides a new and unique insight into the effectiveness of teachers from the perspective of the “customer” that previously was not available. This approach can be described as collective and subjective performance information of individuals and institutions.

Conclusion and implications for M&E

Evaluation is not an exact science; the challenges of measurement and attribution mean that effective measurement always will be a composite of a number of indicators. In 2005, at the first of these meetings, the challenges of M&E were laid out.

Technical developments in social media, initially applied in the commercial and personal arena, provide an exciting opportunity to add an additional lens to our evaluation tool kit—one that enrolls stakeholders more actively in the evaluation process. This has the potential to address three of the challenges—ensuring relevance, involving stakeholders, and communicating the results to stakeholders.

The use of public choice continues to evolve as a policy mechanism to improve performance—the contestable supply of government-funded services through vouchers or other citizen-controlled selection of services. Previously, assessment and evaluation of the results were blunt instruments. They amounted to the quantitative assessment of how citizens exercise choice. This resulted in simplified performance information on the outcomes of a program.

Broadly based personalized capture of qualitative information means that evaluation can now be a lot more nuanced, reflecting the aggregation of individual citizen evaluations. There are many examples of civil society making use of government information to enable more effective evaluation. As government information becomes more widely published as raw data, this trend will become more widespread. As a result, the evaluation of the effectiveness of government programs becomes a cooperative effort—

government working with the community to share data and information on the impacts and outcomes of individual policy initiatives. This option, of using a wider range of sources and tools, gives policy makers a richer palette from which to paint the evaluation picture.

Laurence Millar

Managing Director, Global Village Governance Limited and Former New Zealand Government Chief Information Officer and Deputy Commissioner, State Services Commission, New Zealand

Institutional Arrangements and Policies of Monitoring and Evaluation Systems to Ensure Quality, Access, and Use of Information

Introduction

Within the framework of state modernization, the government’s coordination task is a critical, central point. The government must seek to improve efficiency of state activities on a continual basis, and this has, in turn, made it necessary to seek new forms of exercising adequate strategic leadership and control mechanisms for all government activities.

Thus, interest in the allocation and use of resources has led to concern over results and, especially, impacts. It has generated a tendency to consider the budget as one of the governments’ primary operational systems for coordination, monitoring, and evaluation. As a result, the financial ministries have taken on an increasingly central role.

In Chile, the Treasury Ministry’s mission is defined as “maximizing the economy’s long-term growth potential and promoting the best use of the country’s productive resources in order to achieve sustainable economic growth that translates into a better quality of life for all Chileans and especially for the most vulnerable, left-behind sectors.” Especially noteworthy among its strategic objectives are (1) to promote efficiency in the use of public resources, developing management control techniques and instruments linked to the budget; (2) to direct the state’s financial administration with the objective of fulfilling its institutional mission, and (3) to develop the budget proposal for the public sector, and establish the norms for its implementation. Outstanding among the main products obtained are the following:

- Development of the nation's budget and its negotiation in parliament
- Control over budget implementation and the corresponding achievement of expected results
- Evaluation of government programs
- Definition, implementation, and monitoring of economic policy
- Participation in international economic negotiations
- Advancement in the process of perfecting the capitals market, and the country's international financial integration

Mechanisms for ensuring the quality of M&E indicators

To ensure the quality of the indicators in the M&E systems, governments must consider and understand the function of strategic planning and work planning. In this way, it is possible to develop better M&E indicators that will allow for more precise, assertive evaluations. In Chile, the elements considered include the government program and presidential addresses, messages and priorities, and government programming. Based on these instruments, indicators are added and defined to monitor the government's operation and to better evaluate the progress made in fulfilling what has been promised.

Government program

This is the starting point for defining the government's tasks, because it contains the proposals and issues that were revealed during the presidential campaigns and that the government has committed to achieving.

Presidential messages

In these messages, presidents give an administrative account of what has been achieved, present their perspectives on the situations in their countries, declare new commitments, and specify the issues proposed in the government program or define new policies and projects.

Presidential discourses

These are the presidents' public interventions that refer to specific issues or particular realities, and that define public policies, proposals for action, or commitments for government actions.

Presidential priorities

Based on the government program, as well as the commitments in presidential messages and addresses, a set of goals and projects in which the president has particular interest are presented. These may be legislative initiatives or matters related to government management.

Government programming

This system, operated by the Inter-Ministerial Coordination Office of the General Presidential Ministry Secretariat, makes it possible to maintain constant, total monitoring of all the institutions in the executive branch of government, with regard to progress made and fulfillment of their primary annual commitments. It assumes that, for each year, with the Government Program in mind, the various institutions establish their primary commitments in relation to the goals for their actions. This programming task is important for the development of the budget proposal for each institution and should be subsequently adjusted to what is established in budget legislation.

The task of monitoring allows the government to obtain an overall perspective of its priorities within each part of the administration. It makes it possible to determine the degree to which each concrete plan is being carried out; to what degree the results meet the objectives; and the commitments assumed by the president. The primary objective of this monitoring process, in addition to organizing the government's activities, is to anticipate and contribute toward resolving the obstacles and difficulties that the state apparatus may be confronting in achieving the objectives. These difficulties may refer to matters of a legal, budgetary, or management nature, or to insufficiencies in human resources aspects.

In particular, government programming involves an ongoing process that is initiated with the definition of annual priority tasks by each ministry. These tasks must be in accordance with the definition of policies in each institution. The process later incorporates evaluation stages, with monitoring and evaluation (M&E) at the end of each quarter. The information generated in this entire process is turned over to the president, ministries, and assistant minister's offices. Along with this quantitative evaluation, the Inter-Ministerial Coordination Office drafts a quarterly report of a qualitative nature for each ministry, and this report emphasizes the main achievements or delays in the quarter and suggests some lines of action.

Tools for optimizing quality, reliability, access, and use of M&E information

The tools for optimizing quality, reliability, access, and use of information produced and used by M&E systems are important instruments for verifying progress of planned activities, and if corrective measures are needed. The following tools are available:

Governmental programming

Governmental programming makes it possible to maintain constant, total monitoring of all the institutions in the executive branch of government, with regard to progress made and fulfillment of their primary annual commitments.

Monitoring of presidential priorities

Every month a document is developed with an analysis of the current status for each one of the presidential priorities. The information is recorded in an ad hoc document. This monitoring process makes it possible to carry out appropriate actions to fulfill the identified priorities.

Control over budget use

One of the elements in state modernization corresponds to understanding the budget in relation to the desired results and impacts from the implementation of public policies—more than in terms of the use of inputs and control over such use. National budget development and budgetary control actions increasingly have been advancing toward a conceptualization of a results-based budget. Thus, the task of the Treasury Ministry's Budget Office is fundamental and complementary to the work of monitoring governmental programming and therefore to monitoring the achievement of the government's tasks and objectives. In addition a fundamental role is played by the Planning Ministry in the social evaluation of investment projects, a task that facilitates and complements the monitoring of governmental programming.

Evaluation of government programs

Evaluation is directed by the Budget Office and focused on verifying impacts from the various public programs.

Analysis and monitoring

Analysis and monitoring of public policies is carried out by presidential advisors in collaboration with the

General Presidential Ministry Secretariat and the Budget Office.

The Law on Access to Public Information (No. 20.285, dated April 20, 2009)

This law recognizes the right of individuals to access public information, with the appropriate legal exemptions. This law also obliges state entities to maintain a body of relevant information for supervision over their performance permanently available to the public through their respective Web sites, and establishes sanctions for authorities who deny access to information without justification.

Most of these instruments are characterized by their simplicity and ease of use, and contribute to the functioning of the entire model and to the creation of synergies among them. The production and processing of associated indicators is assisted by available modern information technologies, particularly statistical programs, tools for databases and treatment of information, and, above all, Web pages for disseminating associated information and the corresponding access available to citizens.

The main tool is the System of Management Control and Results-Based Budgeting. Through this tool, a link is provided between the use of allocated resources and the consequent decisions regarding allocation of funds, by obtaining information on performance and introducing best practices for improving the quality of expenditures. The basis for the model is the traditional budgetary cycle: the development of the budget, its discussion and approval, its implementation, and its evaluation. During the various stages of the cycle, various performance-linked instruments are used, including performance indicators and evaluation.

Performance indicators and evaluations are instruments for collecting information regarding performance, and even when both have common elements and complement each other; they are different in terms of methodological aspects and use. Performance indicators are measurements that contain algorithms for measuring results. Evaluations, on the other hand, are studies and, consequently, more comprehensive than measurements, although they include performance indicators.

Other instruments cover another part of the decision-making process and are included in the model. One of these instruments is the presentation of budget initiatives,

which has its own methodological and conceptual elements, and is focused on reviewing new initiatives or on expanding those already existing for a budget presentation. In addition, the comprehensive management analyses (CMA), which are annual reports prepared by the institutions and coordinated by the Treasury Ministry, cover a set of financial and nonfinancial performance topics. Finally, the management improvement programs (MIPs) have a less direct link with budget decision making and include monetary incentives for the public functionaries involved, in accordance with the fulfillment of certain objectives in carrying out the management system.

Challenges in ensuring the quality of information for decision making

One ongoing challenge to ensuring the quality of information for public policy decisions continues to be to identify the best size of the system for the best administration. This objective is pursued by moving toward those indicators identified as the most useful and the most relevant, and with the least difficulty or time involved. This may require having to focus on intermediate indicators of results, which are simpler to construct but are not less assertive, and which may be evaluated in shorter time periods during the implementation of programs.

Another challenge is to move toward products indicators focused on aspects of quality, such as user satisfaction with respect to the public services they receive, and the opportunity for attention, as well as other indicators linked to quality, such as beneficiaries' perceptions and the technical quality of goods and services delivered.

With regard to evaluations, it is necessary to improve the use of the results in public institutions. Although statistics on use of results point to some progress, they also show that much work is left to do to generalize their use, especially in the case of performance information generated by M&E systems. Matters also are pending in the area of improving the information available within institutions for better applying evaluation methodologies and thus precisely improving evaluation processes.

Difficulties, advances, and challenges in policies and laws for access to public information

The recently enacted Law on Access to Public Information recognizes the right of individuals to access public infor-

mation. According to this law, all information "prepared with public funds and any other information possessed by Administration entities, no matter what its format, medium, date generated, origin, classification or processing" is deemed public information.⁶¹ The only reasons for information to be reserved or maintained in secrecy, causing access to be partially or totally denied, are the following:

- When the public availability, communication, or knowledge of such information may affect the necessary due performance of the entity's functions.
- When the public availability, communication, or knowledge of such information may affect the rights of persons, particularly when their safety, health, private life, or commercial or economic rights are involved.
- When the public availability, communication, or knowledge of such information may affect the nation's security, especially when national defense, maintaining public order, or public security are involved.
- When the public availability, communication, or knowledge of such information may affect national interest, especially when public health, international relations, or the country's economic or commercial interests are involved.
- When involving documents, data, or information that a "qualified quorum" has declared reserved or secret (in which case administration entities are required to make and maintain an updated index public).⁶²

While the enactment of this law has meant a major step forward in guaranteeing citizens' access to information generated by M&E systems, its implementation has not been without difficulties. The main difficulty has involved the delayed dissemination of information within the public sector regarding the new institutionalism, as well as the insufficient training of key public functionaries prior to the law entering into effect. This has hindered the law's implementation, and the necessary elements have not been in place since the beginning for the various public services to be prepared. Thus, they have not been able to provide an adequate response to citizens' requests for information. Also, the limited time period established for its implemen-

tation has not been sufficient for departments to improve their archive systems and the organization of their information in a timely manner. The greatest challenge, however, has involved replacing the tradition of secrecy—which continues to prevail in the public sector—with a new tradition of transparency.

Nevertheless, despite all the difficulties, a preliminary evaluation gives reason for optimism, because significant advances have been noted, resulting from efforts to promote the new legislation. This means it will be increasingly possible to count on more public functionaries to contribute to creating a public tradition of transparency.

Other considerations

These considerations clearly contribute to improving access to information generated by M&E systems, thus balancing the technical and political difficulties involved in disseminating M&E information, given the right to access to public information. With regard to the practices and mechanisms generated by the M&E systems for providing information to civil society, the first element that should be emphasized is a self-imposed requirement that all information produced by these systems is of a public nature. Thus, such information is not only sent to Congress, but also directly reported to the entities involved and also published in full on their respective Web sites, in addition to the Budget Office Web site,⁶³ thereby made permanently available to all the country's political and social actors having a need for such information. In addition, the information must fulfill certain technical requirements and, to this end, there is a set of support documents, methodologies, and procedures that also are available to the public.

Producing timely results is an important requirement, since they should be used in analysis and decision-making processes. If information is late in becoming available, it is no longer relevant and becomes obsolete. This makes it necessary, therefore, to have rigorous procedures, to ensure that time limits are observed, and to guarantee both the quality and the timeliness of the corresponding information.

With regard to the impacts from the dissemination of and access to M&E information, specifically in terms of improvements in public management, the first point to underscore is the positive effect on the public departments that have been subjected to M&E systems. This has helped

these departments prepare adequately for making all necessary information available. This does not always happen, however, because although public entities have large volumes of information, they do not necessarily administer or organize it well. And the organization they have developed is not necessarily useful for the purposes of internal administrative control. In addition, the application of these systems has been advantageous: these entities now are concerned about and paying attention to the need for better coordination. Lastly, the greater dissemination of M&E information has led to a favorable public perception in terms of greater transparency in public management. With the recently enacted Law on Access to Public Information, this transparency undoubtedly will be increasingly recognized.

Fernando Barraza

International Consultant, Former Subdirector of Information Technology for the Service of Internal Taxes, Chile

Session 5 Notes

⁴³ The Centre for Good Governance, Hyderabad, conducted the study for the Department of Administrative Reforms and Public Grievances, Government of India, in 2008.

⁴⁴ These tools were developed by the Centre for Good Governance, Hyderabad, India in a World Bank Institute-funded initiative.

⁴⁵ See the New Zealand government Web site, www.ssc.govt.nz.

⁴⁶ This approach is described at the State Services Commission Web sites, <http://www.ssc.govt.nz/display/document.asp?DocID=7121>, <http://www.ssc.govt.nz/display/document.asp?DocID=3530>, and <http://www.ssc.govt.nz/display/document.asp?DocID=3581>.

⁴⁷ See the Web sites, <http://www.ssc.govt.nz/display/document.asp?navid=346>, <http://www.tbs-sct.gc.ca/maf-crg/index-eng.asp>, http://www.whitehouse.gov/omb/assets/egov_docs/2009_Expanding_E-Gov_Report.pdf, and <http://www.whitehouse.gov/omb/expectmore/>

⁴⁸ See the Web site, <http://www.oecdbookshop.org/oecd/display.asp?sf1=identifiers&st1=5ksm37l6svjk&CID=&LANG=EN>

⁴⁹ The entire book is available at <http://yupnet.org/benkler/>

⁵⁰ See the Web sites, <http://www.slideshare.net/missroguel/government-20-architecting-for-collaboration> and <http://www.e.govt.nz/resources/research/progress/transformation/chapter7.html>.

⁵¹ See the Web site, <http://www.policeact.govt.nz/wikil>.

⁵² See the Web site, <http://www.mixedink.com/opengov/>.

⁵³ See the Web site, <http://www.katelundy.com.au/2009/05/29/public-sphere-2-open-government-policy-and-practice/>.

- ⁵⁴ See the Web site, <http://www.whitehouse.gov/blog/2009/12/09/why-open-government-matters>.
- ⁵⁵ See the Web site, <http://www.youtube.com/v/BPr8EITQMlgu>.
- ⁵⁶ Examples can be found at the Web sites, www.data.gov, www.open.org.nz, www.opengov.se, and www.opengov.fi.
- ⁵⁷ For full details, see the Web site, <http://www.ero.govt.nz/ero/publishing.nsf/Content/ERO%27s+Role>.
- ⁵⁸ See the Web sites, <http://www.telegraph.co.uk/education/6193787/Scrap-school-league-tables-say-experts.html>, <http://www.yomiuri.co.jp/dy/editorial/20091018TDY04304.htm>, and <http://www.parra.catholic.edu.au/News---Events/Latest-News/Latest-News.aspx/Alliance-of-school-communities-unite-to-oppose-school-league-tables.aspx>.
- ⁵⁹ See World Bank 2005. Institutionalizing M&E systems in Latin American and Caribbean Countries. PREM Notes 102, available at the Web site, <http://www1.worldbank.org/prem/PREMNotes/premnote102.pdf>.
- ⁶⁰ See the Web site, www.hacienda.cl/ministerio.
- ⁶¹ Article 5.
- ⁶² Article 21.
- ⁶³ See the Web site, www.dipres.cl/fr_control.html.

Select Presentations from the Fourth Meeting of the Latin American and Caribbean Monitoring and Evaluation Network

Belo Horizonte, Brazil, November 25, 2008

Institutionalizing Impact Evaluation: Lessons from Latin America and Other Regions⁶⁶

This paper describes the progress being made in the transition from individual impact evaluation (IE) studies to building a systematic approach to identifying, implementing, and using evaluations at sector and national levels. When this is achieved, the benefits of a regular program of IE as a tool for budgetary planning, policy formulation, management, and accountability will begin to be appreciated. To date, the institutionalization of IE has been achieved only in a few developing countries, mainly in Latin America, but others have started the process. This chapter draws lessons on the benefits of an institutionalized approach to IE, the conditions that favor it, the challenges limiting progress, and the steps in developing such an approach. The institutionalization of IE can be achieved only within the framework of a comprehensive monitoring and evaluation (M&E) system that provides a menu of evaluations to cover all the information needs of managers, planners, and policy makers.

The importance of impact evaluation for official development assistance

To assess the effectiveness of official development assistance (ODA) programs in contributing to poverty reduction, economic growth, and sustainable development, it is important to conduct systematic analysis of development effectiveness. Many assessments of effectiveness are based either on data collected from project beneficiaries *after* the project has been implemented, or on the use of M&E or results-based management to measure changes that have tak-

en place in the target population over the life of the project. In all these cases, data are generated only on the target population and no information is collected on sectors that do not benefit from the project or that, in some cases, may even be worse off as a result of the project.

With all of these approaches, there is a tendency for the evaluation to overestimate the true benefits or effects produced by the project (Bamberger 2009b). Typically, only project beneficiaries are interviewed. Families or communities that do not benefit are not interviewed and the evaluation does not present any information on the experiences or opinions of these nonbeneficiary groups. Many of these evaluations are methodologically weak and often biased and consequently there is a serious risk that ODA agencies may continue to fund programs that are producing smaller impacts than are reported and that may even be producing negative consequences for some sectors of the target population.

Although no hard statistics are available, it is quite likely that rigorous IE designs are used in only 10 percent of ODA impact evaluations. Given the widespread recognition by ODA agencies of the importance of rigorous IEs, why are so few conducted? Reasons include the limited evaluation budgets of many agencies. Most evaluations are not commissioned until late in the project cycle and consultants are given a short time to conduct data collection. Also, many government agencies see evaluation as a threat or something that will demand a lot of management time without producing useful findings. Because many evaluations produce a positive bias that show programs in a positive light, many agencies do not feel the need for more

rigorous (as well as more expensive and time-consuming) evaluation methodologies. One of the challenges for the institutionalization of IE is to convince both ODA agencies and host country governments that rigorous and objective impact evaluations can become valuable budgetary, policy, and management tools.

Institutionalizing impact evaluation

Institutionalization of IE at the sector or national level occurs when (1) the evaluation process is country led and managed by a central government ministry or sector agency, (2) buy-in from key stakeholders is strong, (3) procedures for selecting, implementing, and using IEs are well defined, (4) IE is integrated into national M&E systems that generate much of the evaluation data, (5) IE is integrated into national budget formulation and development planning, and (6) the focus is on evaluation capacity development (ECD). Institutionalization is a process, and at any given point, it will have advanced further in some sectors than others. How IE is institutionalized will also vary from country to country, reflecting different political and administrative systems and historical factors such as donor support for particular sectors.

Many IEs have not been able to achieve their potential contributions to program management, budget planning, and policy making because evaluations were selected and funded in an ad hoc and opportunistic way determined by the interests of donor agencies or individual ministries rather than by national planning priorities. The value of IE as a policy-making tool can be greatly enhanced once selection, dissemination, and use of the evaluations becomes part of a national IE system. This requires an annual plan for selection of the government's priority programs on which decisions have to be made concerning continuation, modification, or termination and for which the evaluation framework permits the comparison of alternative interventions in terms of cost-effectiveness and contributions to national development goals.

The three pathways

No single strategy is always successful in the institutionalization of IE. Countries that have made progress have built on existing evaluation experience, political and administrative traditions, and the interest and capacity of individual ministries, national evaluation champions, or donor agen-

cies. Figure 27 identifies three alternative pathways to the institutionalization of IE.

The first pathway

The ad hoc or opportunistic approach evolves from individual evaluations that took advantage of available funds and the interest of a senior government official or a particular donor. The approaches were gradually systematized as experience was gained in selection criteria, effective methodologies, and how to achieve both quality and utilization. A central government agency, usually finance or planning, either is involved from the beginning or becomes involved as the focus moves toward a national system. Colombia's national M&E system illustrates this pathway with the Ministry of Planning responsible for managing the National System for Evaluation of Public Sector Performance (SINERGIA). Although selection of evaluations initially was somewhat ad hoc and strongly influenced by donor interests, as the program of IE evolved, the range of methodologies was broadened and technical criteria in the selection of programs to be evaluated were formalized through policy documents (with more demand-side involvement from the agencies managing the programs being evaluated) and in how the findings are used. Most of the IEs use rigorous econometric evaluation techniques (Mackay 2007, 31–36; Bamberger 2009a).

The second pathway

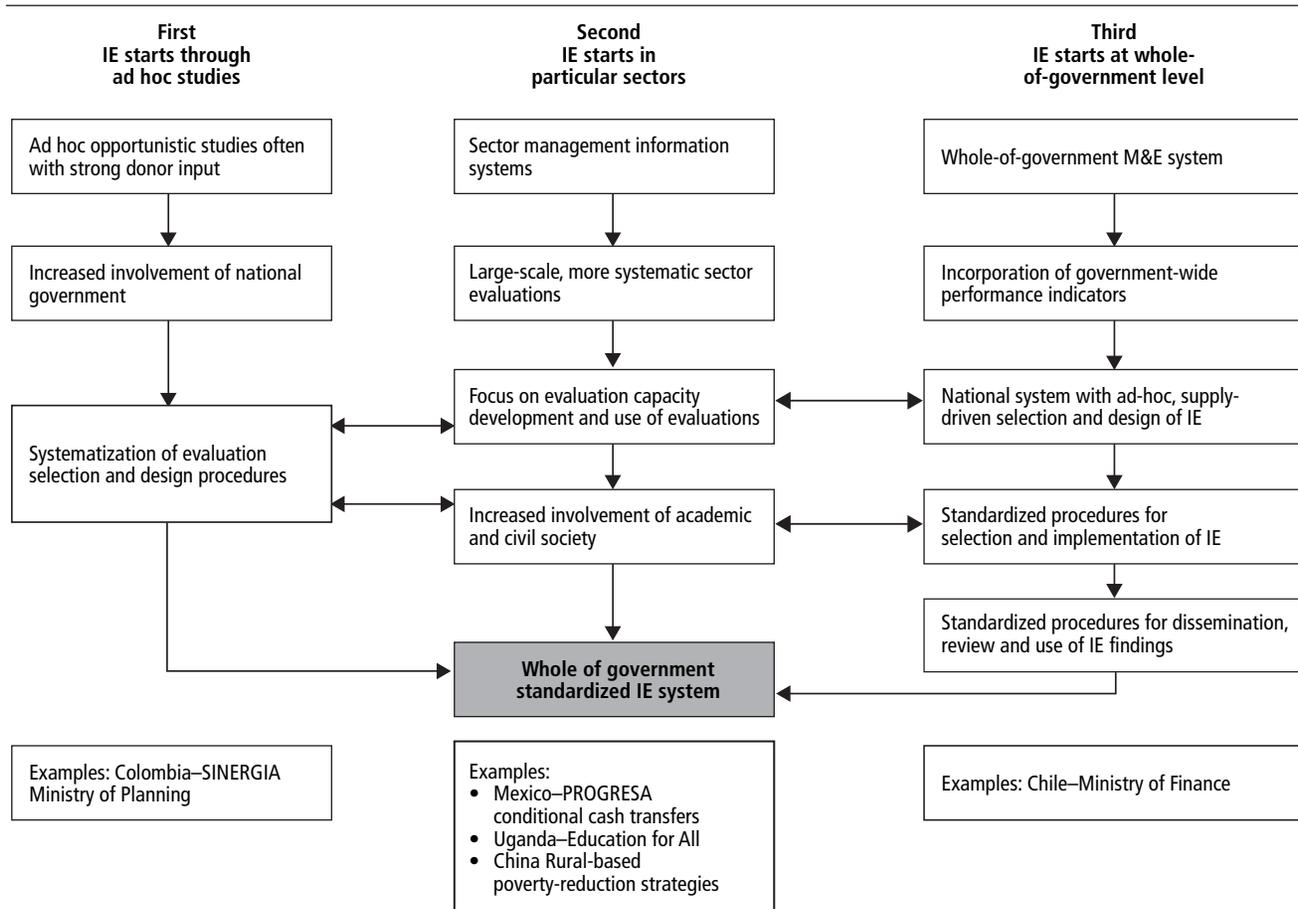
The second pathway is where IE expertise is developed in a priority sector supported by a dynamic government agency and strong donor support, and for which there are important policy questions to be addressed. Once the operational and policy value of these evaluations has been demonstrated, this becomes a catalyst for developing a national system.

Mexico example

The evaluations of the Education, Health, and Nutrition Program (*Programa de Educación, Salud y Alimentación*, PROGRESA) conditional cash transfer programs in Mexico are an example of this approach (Mackay 2007, 56; Bamberger 2008). A series of rigorous evaluations of the PROGRESA programs conducted over a number of years convincingly demonstrated the effectiveness of conditional cash transfers as a way to improve the welfare (particularly health and education) of large numbers of low-income

FIGURE 27:

Three Pathways for the Successful Institutionalization of IE



Source: Author.

families. The evaluations are considered to have been a major contributing factor in convincing the new government that came to power in 2002 to continue these programs started by the previous administration. The evaluations also convinced policy makers of the technical feasibility and policy value of rigorous IEs and contributed to the passage of a law by Congress in 2007 mandating the evaluation of all social programs. This law also created the National Commission for the Evaluation of Social Programs, which was assigned the responsibility for regulating the development of M&E functions in the social sectors.

Africa example

The World Bank Africa Impact Evaluation Initiative (AIM) is a regional initiative to help governments strengthen their overall M&E capability and systems. The initia-

tive currently supports some 90 experimental and quasi-experimental IEs in 20 African countries. At least 40 countries in Asia, Latin America, and the Middle East are now implementing sectoral approaches to IE with World Bank support. A number of international evaluation initiatives are being promoted through collaborative organizations such as the Network of Networks for Impact Evaluation (NONIE) and the International Initiative for Impact Evaluation (3IE).

The third pathway

In the third pathway, a program of IEs was developed as one component of a whole-of-government system, managed and championed by a strong central government agency, usually the ministry of finance or planning. Chile is a good example of a national M&E system with clearly

defined criteria and guidelines for the selection of programs to be evaluated, their conduct and methodology, and how the findings will be used (Mackay 2007, 25–30; Bamberg-er 2008). Starting in 1994, a system of performance indicators was developed; rapid evaluations of government programs were incorporated in 1996; and, in 2001, a program of rigorous impact evaluations was introduced. There are two clearly defined IE products. The first are rapid ex post evaluations that follow a clearly defined and rapid commissioning process, where the evaluations are used by the ministry of finance as part of the annual budget process. The second are more comprehensive evaluations that can take up to 18 months and cost \$88,000 on average. The strength of the system is that it has defined and cost-effective procedures for commissioning, conducting, and reporting of IEs; a defined audience (the Ministry of Finance); and a clearly understood use (the preparation of the annual budget). The disadvantages are that the studies cover only issues of interest to the Ministry of Finance and the buy-in from the agencies being implemented is typically low.

Integrating IE into sector and national M&E and other data-collection systems

The successful institutionalization of IE requires that selection, design, and use are integrated into national M&E systems and data-collection programs. This is critical for several reasons. First, much of the data required for an IE can be obtained efficiently and economically from program M&E systems. Second, IE findings that are widely disseminated provide an incentive for agencies to improve the quality of M&E data they collect and report. For example, the Ministry of Education in Uganda reported that the wide dissemination of the Education for All evaluations contributed to improvements in the quality of monitoring reporting. Third, national household survey programs provide valuable sources of secondary data for strengthening methodological rigor of IE design and analysis (for example, the use of propensity score matching to reduce sample selection bias). Some IEs have cooperated with national statistical offices to piggyback information required for the IE onto an ongoing household survey or to add a special module. Creating demand for IE

Efforts to strengthen IE and other kinds of M&E systems often are viewed as technical fixes—mainly involving better data systems and the conduct of quality evaluations.

Although the creation of evaluation capacity to provide high-quality evaluation services is important, these supply-side interventions will have little effect unless there is sufficient demand for quality IE. This requires that quality IEs are seen as important policy and management tools in at least one of the following areas:

- Budget resource allocation
- Policy formulation and analytical work
- Management and delivery of government services
- Accountability

Creating demand requires powerful incentives within government to commission IE, create a good level of quality, and use IE information intensively. A key factor is a public sector environment supportive of the use of evaluation findings as a policy and management tool. To generate demand for IE among potential users, MacKay (2007, Table 11.1) suggests some such positive incentives as building ownership and financing to build better IE systems; sanctions and threats, such as laws and decrees mandating the planning, conduct, and reporting of IE; and positive messages from key figures, such as high-level statements of endorsement and conferences on good IE practice.

Capacity development for IE

There has been an increasing awareness of the importance of evaluation capacity development. This is largely due to the recognition that many past efforts were overambitious and had disappointing results, often because they focused too narrowly on technical issues without understanding the institutional and other resource constraints faced by many countries. ODA agencies now recognize that institutionalization of IE will require an evaluation capacity development plan to strengthen the capacity of key stakeholders to fund, commission, design, conduct, disseminate, and use IE. This involves strengthening the *supply* of resource persons and agencies able to deliver high-quality and operationally relevant IEs and developing the infrastructure for generating secondary data to complement or replace expensive primary data collection.

An evaluation capacity development strategy must target (1) agencies that commission, fund, and disseminate IEs; (2) evaluation practitioners who design, implement, and analyze IEs; (3) evaluation users and groups affected by the programs being evaluated; and (4) public opin-

ion (Bamberger 2006). The active involvement of leading national universities and research institutions is critical as these institutions can mobilize leading national researchers, and they have the resources and incentives to work on refining existing and developing new research methodologies. Through teaching, publications, conferences, and consulting, they can strengthen the capacity of policy makers to identify the need for evaluation and to commission, disseminate, and use findings.

An important role of evaluation capacity development is to help agencies design “evaluation-ready” programs and policies. Many programs generate monitoring and other administrative data that could be used to complement or replace the collection of survey data, or to provide proxy baseline data in cases in which an evaluation started too late to have been able to conduct baseline studies. Often, however, the data are not collected or archived in a way that makes it easy to use for evaluation purposes. Cooperation between program staff and evaluators can greatly enhance the utility of project data for the evaluation. In other cases, changes in project design or implementation could strengthen the evaluation design. For example, in some cases, a randomized control trial could have been used, but evaluators were not involved until it was too late.

Promoting IE

There is widespread concern that even for evaluations that are methodologically sound, the use of evaluation findings is disappointingly limited and the process of IE is not institutionalized (Bamberger, Mackay, and Ooi 2004). In addition to factors that are external to the evaluation, the reasons many include (1) lack of ownership, (2) lack of understanding of the purpose and benefits of IE, (3) bad timing, (4) lack of flexibility and responsiveness to the information needs of stakeholders, (5) wrong question and irrelevant findings, (6) weak methodology, (7) lack of local expertise to conduct, review, and use evaluations, (8) communication problems, and (9) lack of a supportive organizational environment.

Program evaluations can be influential in many different ways, not all of which were intended. Some of the different kinds of influence that IEs can include are (1) alerting public service agencies to problems of which they were not fully aware or had not considered important; (2) providing objective, quantitative data that civil society can use

to pressure agencies to improve services; (3) alerting agencies to the need to incorporate vulnerable groups and address community conflicts caused by limited access to services; (4) providing political cover to government to take a politically sensitive decision, and showing how to mitigate negative consequences for influential “losers” from the policy change; (5) developing methodologies to systematically document problems in the use of public funds that were widely suspected but that have not been possible to document; and (6) providing empirical evidence to convince new administrations to continue major programs initiated by previous administrations (Bamberger 2008).

Bamberger (2008) suggests various ways to strengthen the use of the findings of impact evaluations:

- Understand the political context within which the evaluation is conducted.
- Time the launch and completion of the evaluation appropriately.
- Decide what to evaluate. A successful evaluation usually focuses on a limited number of critical issues and hypotheses based on a clear understanding of the clients’ information needs and of how the evaluation findings will be used.
- Base the evaluation on a program theory model so that evaluators, clients, and stakeholders share a common understanding of the problem the program is addressing, the objectives, how they will be achieved, and what criteria the clients will use in assessing success.
- Create ownership of the evaluation.
- Include an ECD strategy.
- Develop a follow-up action plan to monitor whether agreed decisions are implemented.

One of the factors contributing to evaluation underutilization may be client disagreement with, or lack of understanding of, the evaluation methodology. Therefore, IE should be combined with process analysis and formative evaluation. Even when the primary objective is to assess results, it is important to help the client understand how outcomes and impacts are affected by what happens during program implementation. Process and formative evaluation also provide rapid feedback to the client, increase ownership, and improve perception that the evaluation is useful.

Conclusion

ODA agencies are facing increasing demands to account for the effectiveness of the resources they have invested in development interventions. This has led to an increased interest in more rigorous evaluations of the programs and policies these agencies support, and an increasing number of methodologically rigorous impact evaluations have been conducted, particularly in Latin America. However, less progress has been made toward institutionalizing the selection, design, and use of IE. Consequently, many evaluations have not achieved their full potential as instruments for budget planning and development policy formulation.

Successful institutionalization of IE requires that the approach be adapted to the political and administrative systems of each country. This paper identifies three quite distinct pathways toward institutionalization. Conducting a number of rigorous IEs in a particular country does not guarantee that agencies will automatically increase their demand for evaluations. As a result, a strategy is needed to create demand. While it is essential to strengthen the *supply* of evaluation specialists and agencies able to implement evaluations, experience suggests that creating the *demand* for evaluations is equally if not more important. This requires a combination of incentives, sanctions, and positive messages from key figures, and that IE be seen as an important policy and management tool in such areas as resource allocation, policy formulation, and analytical work, delivery of government services, and underpinning accountability.

Evaluation capacity development is a critical component of IE institutionalization. Some groups require the capacity to design and implement IE; others need to understand when an evaluation is needed and how to commission and manage it, whereas others must know how to disseminate and use the evaluation findings. Many well-designed and potentially valuable IEs are underutilized because of the lack of ownership by stakeholders, bad timing, failure to address client information needs, lack of follow-up on agreed actions, and poor communication and dissemination. A strategy to promote utilization is essential for IE institutionalization.

ODA agencies will continue to play a major role in promoting the institutionalization of IE. This requires a commitment to move from ad hoc support of individual IEs to a genuine commitment to helping countries develop an IE system that serves the interests of national policy mak-

ers and line ministries. This also requires a commitment to evaluate capacity development, and the development of a wider range of IE methodologies that are designed to respond more directly to country needs rather than to impose methodologically rigorous evaluation designs that often are of more interest to ODA research institutions than to developing country governments.

For their part, developing country governments must invest the necessary time and resources to ensure that they fully understand the potential benefits and limitations of IE and the alternative methodologies that can be used. Usually one or more ODA agencies will be willing to assist governments wishing to acquire this understanding, but governments must seek their own independent advisors so that they do not depend exclusively on the advice of a particular donor who may have an interest in promoting only certain types of IE methodologies.

Michael Bamberger

Social Development and Program Evaluation Consultant

Building a Monitoring and Evaluation System Based on Results for Social Development

Introduction

Results-based management is designed to improve the performance of public institutions seeking greater efficiency and effective use of resources. At least three objectives are pursued through results-based management. The first is to direct government actions in the provision of goods and services to achieve measurable goals efficiently. The second is to promote evidence-based decision making. The third is to try to meet the demands of society regarding transparency and accountability.

A number of member countries of the Organisation for Economic Co-operation and Development (OECD) have used this approach. Since 2000, more than 75 percent of OECD member countries have reported using performance information in budget processes (ILPES-CEPAL 2003). Countries with the most experience include the United Kingdom, with the introduction in 1982 of the Financial Management Initiative, and New Zealand, with the introduction in 1984 of a system of purchase agreements between ministers and autonomous executive bodies

that contain terms of reference expressed as results indicators. Subsequently, in 1993, the United States approved the Government Performance and Results Act, which seeks to install a budget system based on performance and managerial flexibility, and which concludes with a programmed reporting process that is delivered to Congress to support the budget discussion (Shack Yalta 2003).

More recently, Argentina, Colombia, and Chile have made similar efforts (World Bank and OECD 2006; May et al. 2006). Chile's effort is one of the most important Latin American examples. Its origins date from the early 1990s, when a pilot scheme of the Plan for the Modernization of Public Management was launched to phase in a results-based budgeting system. The Chilean model now has a management system linked to the budget with such tools as performance indicators linked to targets, an evaluation system for programs and institutions, a management improvement program linked to performance incentive payments to public officials, and a competitive fund linked to the logical framework matrix, which, among other things, is used for ministries (and programs) to compete for funding (Guzmán 2004).

In Mexico, in 1999, the Budget for Federal Spending established the obligation to perform annual external evaluations of all programs subject to operating rules. Since then, the Social Development Secretariat (*Secretaría de Desarrollo Social*, SEDESOL) has striven to incorporate a management approach geared toward achieving results. The approach emphasizes the development and strengthening of the functions and tools for M&E. Initially, this effort focused on the annual external evaluation of the programs subject to operating rules. In addition, since 2004, the secretariat has promoted the construction of a results-based M&E system. The system is designed to strengthen the evaluation component with rigorous measurement of the impact of the programs and to develop a pilot monitoring device based on results. SEDESOL is now ready to begin a second phase of the system under a new model with greater scope. To that end, it has identified important cross-sectional challenges in the institutional, technical-operational, and data-processing spheres.

This article describes SEDESOL's efforts at implementing results-based management. It is divided into four sections that detail (1) the first phase of the results-based monitoring and evaluation (M&E) system, (2) the change in

the institutional and regulatory framework, (3) the new model for the expansion and consolidation phase and, finally, (4) the conclusions.

Building a results-based monitoring and evaluation system, 2004–07

For many years, social programs lacked the mechanisms to ensure accountability and provide evidence of effectiveness. The first efforts, still few and far between, to quantify rigorously the benefits of government intervention include an impact assessment of the Education, Health, and Nutrition Program (*Programa de Educación, Salud y Alimentación*, PROGRESA) between 1997 and 2000. The obligation to evaluate programs annually subject to operating rules and quarterly reporting of indicators subsequently has been introduced.

SEDESOL asked the World Bank to undertake a diagnosis of existing information systems and the ministry's institutional capacity, and to develop an M&E system that would facilitate progress toward results-based management. The World Bank's diagnosis concluded that SEDESOL did indeed have the necessary foundation for moving toward results-based management. However, it was necessary to modify the following aspects:

- Incipient linkage of delivery of program products and services to the measurement of results
- Abundant production of operational information (products), not necessarily complemented by information on results (outcomes, impacts)
- Multiplicity of administrative, operational, and management data systems
- Many reports on indicators to different external institutions without any clear use for managing the programs
- Design of systems for reporting activities, but with limited information and usage for feedback to those responsible for decision making
- Disconnect between strategic, operational, and budgetary planning
- Need to create incentives at midhierarchical levels to promote results orientation

Based on this diagnosis, an M&E system was designed to support results-oriented management, considering three

main components: (1) evaluation of annual results, (2) impact assessment, and (3) pilot implementation of a system of indicators for monitoring focused on results. The conceptualization of the three components in a results-based monitoring and evaluation system (RBMES) shifts the emphasis of traditional management from input to achieving results. Figure 28 shows the interaction of components and their scope, all of which focus on results-based management.

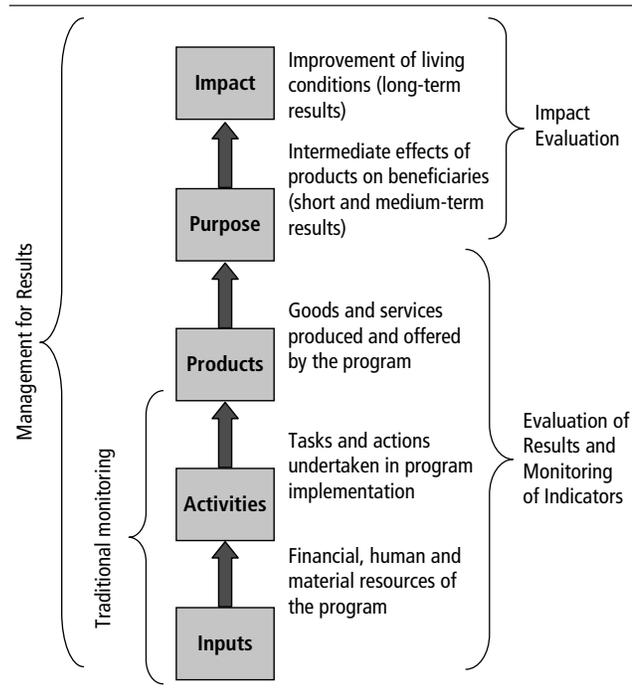
Annual evaluation of results

In keeping with current regulations, an evaluation report should be prepared each year by an external evaluator for programs subject to operating rules (see Table 20). The evaluation approach was primarily for accountability in the fiscal year concerned with meeting the program objectives, operating rules, and physical and financial targets. Additionally, an analysis of economic and social benefits was requested whose scope was limited by the schedule for the delivery of evaluation reports. A feedback of evaluation results was promoted through the external evaluations tracking system that had a work plan set up to address the recommendations from the evaluations.

Multiyear evaluation of impact

To complement these findings, various medium- and long-term evaluations were conducted to quantify the causal relationships between actions implemented by the programs and different dimensions in changes to recipients' well-being. The impact assessments attempted to provide robust evidence of the effectiveness of various interventions in child nutrition, school attendance, family income, access to basic services, and the like. These evaluations require more investment in terms of budget, a higher level of expertise for their analysis, and a longer maturation time for interventions to observe the expected impacts. Consequently, they were performed selectively for strategic-type programs and activities such as Opportunities, Habitat, Your Home,

FIGURE 28:
Results-Based M&E System, 2004–07



Source: Author.

Food Support Program, Fortified Liconsa Milk, Micro-regions Strategy, and Cement Floors. The results of these evaluations are public and have been used to inform decisions about adjustments to the design, expansion of coverage, and continuity of programs.

Results-based monitoring

Although periodic reports on indicators of the physical and financial progress of the programs were produced, they were not systematic or focused on achieving results. Based on the diagnosis undertaken by the World Bank, we began by implementing a pilot phase of the results-based monitoring component. The pilot phase included four programs (Productive Options, Social Co-investment, Habitat, and Food Support) with different degrees of institutional progress in

TABLE 20:
Annual External Evaluations, 2004–07

	2004	2005	2006	2007
SEDESOL-sponsored programs	7	8	8	11
Programs operated by decentralized bodies or sector entities	6	8	9	10
Total	13	16	17	21

Source: Author.

the collection of indicators and data systems. This phase included extensive training in the methodology of the logical framework matrix (LFM), the construction of matrices, the conceptual design of the system, and the development of a computer platform. LFMs are public management tools for logically and systematically presenting the objectives of a program, the causal relationships of its components, and factors outside the program that can influence its objectives. LFMs were developed through participatory workshops with the leaders of the design, operation, and evaluation of programs. These workshops (1) reached consensus on the overall goal of each program and its corresponding logical model (string of results: activities, outputs, outcomes, and impacts); (2) identified program indicators and goals; and (3) determined the sources of information for monitoring indicators. The introduction of the LFMs led to an improvement in the relevance and accuracy of the indicators. Once the indicators had been identified and targets set, these were concentrated in a control panel. The panel showed data using traffic lights to indicate the progress of the indicators with regard to established targets.

Since 2007, SEDESOL has moved toward the definition and use of LFMs in 21 programs subject to operating rules. This has meant a further step toward the improvement and standardization of the formulation of indicators and targets, and establishing a more solid foundation for performance measurement of its programs.

Changes in the regulatory and institutional framework

The two main reasons behind the need for a second phase of the RBMES involve the reestablishment of the Ministry's results-oriented model. The first relates to the experience gained from pilot implementation of the monitoring component. The second has to do with the need to adapt the RBMES to a new regulatory and institutional context marked by the reforms of the national budget, the creation of the National Social Policy Evaluation Council (*Consejo Nacional de Evaluación de la Política de Desarrollo Social*, CONEVAL), and the need to adopt the guidelines of federal initiatives currently under development as the performance evaluation system and the Results-based Budget (RBB).

By mid-2007, clear guidance from the Ministry of Finance and Public Credit (*Secretaría de Hacienda y Crédito Público*, SHCP), the Ministry of Public Administration

(*Secretaría de la Función Pública*, SFP), and CONEVAL was introduced to move toward the establishment of a link (at national and sectoral levels) between planning, budgeting, and the implementation of federal programs using M&E tools. In particular, this new approach is embodied in the federal administration through the creation of the performance evaluation system and based on the National Development Plan (NDP) and sector programs.

The guidelines for evaluation and budgeting raise the need for a framework of reference in management that considers the strategic objectives of the country and sectors and their alignment and articulation. Likewise, the guidelines recommend that the agencies responsible for programming the budget use performance data, and promote institutional improvement based on performance incentives. Added to these guidelines is the agencies' increased accountability through more and better performance data for budget discussions.

In this context, as of late 2007, the SHCP coordinated an effort to implement results-based management from the perspective of the performance evaluation system whose objectives include the following:

- Verify and monitor fulfillment of goals and objectives, based on strategic and management indicators that show the results of programs
- Link planning, programming, budgeting, monitoring, use of resources and evaluation of public policies and budget programs

This resulted in the general guidelines for moving toward a results-based budgeting and performance evaluation system. A matrix of indicators (or an LFM) is implanted in this system as a tool for strategically planning budget programs. In addition, an exercise is carried out to link the matrix indicators of all programs subject to the federal budget's rules of operation.

All the matrix indicators are concentrated through a computer platform, from the operating level to impact levels (and their connection with the strategic objectives of the agencies and the NDP), which are periodically monitored and reported to ascertain the degree of progress in meeting associated goals.

To ensure the strategic management of information, the SHCP, in coordination with each agency, selected from the matrices a subset of key indicators to monitor program

budgets. The President's Office also tracks the indicators and targets of sector programs. This new institutional and regulatory framework established a different context for the development of the RBMES. This context is in stark contrast to the previous one based on an individualized focus on the programs, particularly their operation. Previously, a broader frame of reference to link strategic sector planning and the budget was lacking, as were strong mechanisms to stimulate demand and the use of performance data by the presidency, the SHCP, SFP, and CONEVAL.

A new RBMES model for the expansion and consolidation phase, 2008–12

In response to the institutional and regulatory changes, there was a rethinking of the RBMES. The new model moves from an approach based essentially on the operation of programs to a larger scheme focused on the aggregate performance of the social development sector. This scheme is based on measuring performance through the strategic dimension of the Social Development Sector program and its alignment with the NDP. Under this approach, the redesign of the RBMES is intended to operate the budgeting for results initiative and the performance evaluation system within the SEDESOL, incorporating their guidelines and purposes. This new approach aims to consolidate the progress achieved, correct any problems encountered, and complement its scope by incorporating other areas of management evaluation (including the budget).

The process of rethinking the system considered (1) results-based monitoring, (2) results-based evaluation, and (3) a results-based budget (see Figure 29). The development, strengthening, and integration of these three components provide a better balance in the process of generating information about performance and facilitate its use within the decision-making processes. The first step will bind the budget to adequate planning of sector targets and outcomes. Monitoring will seek to be a continuous function to strengthen the operation. And, evaluation will be a selective activity based on the criteria of priority policies and a search for areas to improve.

The monitoring component makes it easier to track the operation of programs and their results, a task that focuses mainly on the area of the Social Development Sector's strategies and plans. With the evaluation component, progress is being made in the consolidation of methodologies for the

analysis of causality between major programs operated and coordinated by SEDESOL and its effects, as well as incorporating new tools such as evaluations of consistency and results and specific performance details.

To improve the quality of spending, the results-based budgeting component introduces mechanisms to ensure adequate planning and feedback of budget decisions with performance data from M&E. The balance in the joint operation of these components will facilitate and encourage the monitoring of the ministry's aggregate results, by creating incentives for continuous improvement. Additionally, this approach strengthens strategic planning within social policy management and improves accountability of results, both internally and externally.

Objectives of RBMES

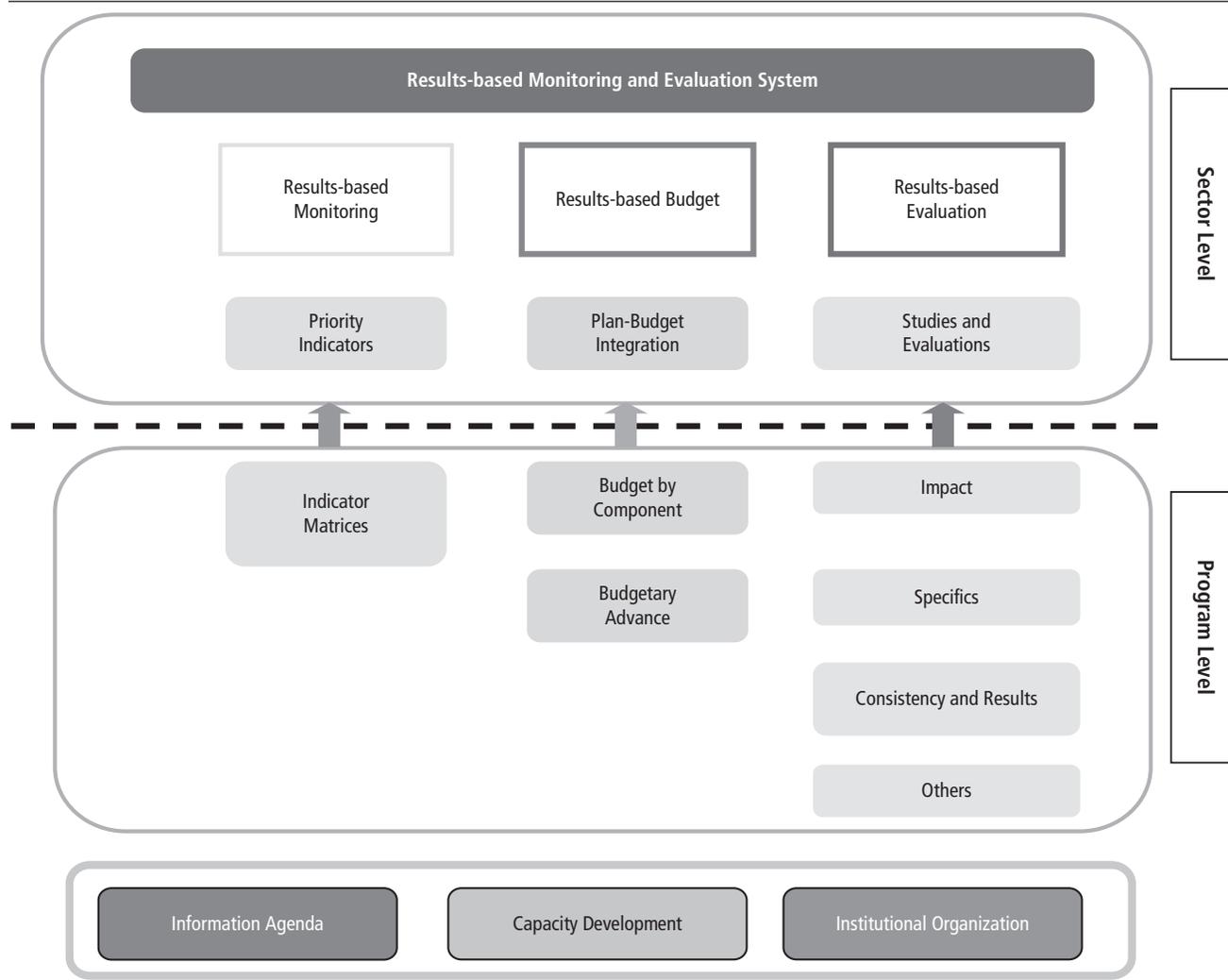
The overall goal is to make operative within the social development sector the RBB initiative and performance evaluation, under an approach that seeks the achievement of sector targets, through an RBMES. It aims to advance the implementation and consolidation of the field as a tool of governance that will (1) monitor the integrated physical and financial progress of programs; (2) establish a unified framework, based on priority criteria for the development of annual and impact assessments to determine their effect on the improvement of the beneficiaries' quality of life; (3) provide feedback for decision making about the operation, resource allocation, and program design; and (4) facilitate and encourage the management of the secretariat in relation to sector and national planning.

The aim is to advance the implementation and consolidation of the RBMES as a tool of governance that will permit (1) integrated tracking of the physical and financial progress of the programs; (2) the establishment of a unified framework, based on priority criteria for the development of annual evaluations and impact assessments to determine their effect on the improvement of the quality of life of beneficiaries; (3) feedback for decision making about the operation, resource allocation, and design of programs; and (4) facilitating and stimulating the ministry's management to plan at the sector and national levels.

Specific objectives include the following:

- Establish a permanent, progressive link between planning, budgeting, implementation, and M&E

FIGURE 29:

Results-Based M&E System, 2008-12

Source: Author.

processes, based on selected output, outcome, and impact indicators, and sector and social development goals, to improve their alignment with the NDP and Sector Program.

- Improve the quality, availability, and timeliness of M&E data that is useful for decision making at the strategic and management levels.
- Coordinate and balance the functions of monitoring, evaluation, and accountability, to guide the management of the SHCP toward achieving results, by improving the design and operation of social programs and the quality of spending.
- Strengthen the institutional quality of the system, by specifying the roles and responsibilities of the

various structures, as well as the coordination and decision mechanisms required.

- Encourage a new organizational dynamic within the SHCP and coordinated entities, linking the areas responsible for planning, programming, and budgeting, as well as M&E, with the programs' operating units.
- Build the M&E capacity of the programs operated and entities coordinated by SEDESOL to produce reliable, timely information for decision making on programming and budgeting.

Conceptual design

The conceptual design of the RBMES is structured around the first three pillars of (1) monitoring, (2) evaluation, and

(3) RBB as well as the system's transversal aspects of (4) institutional structure required to support the RBMES, (5) institutional capacity building for M&E, and (6) the information agenda.

Results-based monitoring

A set of 16 strategic indicators were selected to monitor progress in meeting goals set by SEDESOL in the Social Development Sector program and the NDP. The selection took into account the different levels of information (monthly, quarterly, yearly, and five yearly) and objectives (results, components, and budget). These indicators were selected strategically from the Social Development Sector program and the result-based matrix indicators to satisfy the demand for information that is relevant to the aggregate of individual decision makers in a timely manner. The main user of the indicators report at this level is the President's Office, through its strategic indicators tracking system.

At the program level, as in the rest of federal government, the LFM provides a sound methodology for the selection of indicators for management oriented to program results. The actor requesting monitoring and reporting of indicators of the results base matrix indicators is the SHCP, through its Web portal.

Currently, there are matrices for 19 programs subject to operating rules and about 500 indicators: 112 are linked to results, 170 are related to the quantity, efficiency, and quality of goods and services provided by programs, and the rest are activities necessary for production. Of these indicators, 40 were identified as priorities for the development of the 2010 draft expense budget and budget monitoring. These indicators are mainly at the component level and reflect the priority goods and services provided by social programs and directly related to the availability of funds for each program.

Results-based evaluation

Considering the importance of the Sector Program and the NDP, evaluations are planned on the basis of (1) the amount of public funds involved, (2) the size and characteristics of the target population, (3) the importance to the sector, (4) the newness of the program or policy, and (5) the potential for application and replication of the lessons learned.

This planning defines evaluation measures for social programs in the short and medium term, considering dif-

ferent types of evaluation, such as (1) evaluations of consistency and results or design, (2) specific performance evaluations, (3) impact evaluations for strategic programs, (4) evaluations of processes, and (5) specific evaluations of perception of beneficiaries and monitoring of public works.

Areas for improvement that show up in the evaluation results are tracked so that they effectively contribute to improving social programs. An institutional work plan and specific work plans for each program set out the measures, persons and areas responsible, dates, and deliverables to address the areas for improvement identified in the external evaluations.

Results-based budget

The purpose of an RBB is for the definition of budgetary programs to emerge from a sequential planning process: scheduling and setting goals, indicators, and targets geared toward the achievement of sector goals. This purpose requires the evidence provided by M&E to make the RBMES consistent with the performance evaluation system and RBB guidelines. This involves establishing congruity between outputs and outcomes, on the one hand, with the budget to establish a link between the latter, sector program priorities, and evidence of results. The procedure is an attempt to identify where SEDESOL's expenditure is directed and, based on M&E, to learn whether the expected results are being obtained with the budget allocations for the goals set.

The initial phase of RBB is defined as presentation. At this stage, the budget relates to goods and services provided by various social programs, allowing the alignment of programs and their budgets for the objectives of the sector to be checked. Specifically, the RBB was linked to the other two pillars of the RBMES through the draft 2010 budget. The latter took into account the 40 priority indicators (linked to monitoring) and the evidence provided by external evaluations that the goods and services provided contribute to the achievement of expected results, that is, to the purpose and end (linked to evaluation).

The connection with the budget must be contextualized by considering the difference between the time frame for the budget activity, on the one hand, and the length of time required for the goods and services to materialize and for the expected effects to mature and be measured, on the other hand. While the budget period is almost instanta-

neous, building a water network, for example, can take several months, and measuring the change in the incidence of diarrheic diseases can take even longer. Additionally, consider that the budget-results relationship is neither direct nor automatic, so that the RBMES should be regarded as a tool that supports rather than determines decision making.

Information agenda

To stimulate continuous improvement, one of the prerequisites of the RBMES is the strategic use of information on spending, accountability, and management feedback. This implies the availability of quality, basic, and accessible information for all programs, and it facilitates learning the management results and informed decision making.

It therefore is important to promote an information agenda that includes regulatory bases and standards to generate, collect, and process performance data, as well as quality-control mechanisms. The information agenda should seek to determine the parameters of interoperability of existing technology tools and the connectivity of the offices, entities, and delegations.

Institutional organization

It is important to consider the organizational structure to strengthen the integration of planning and budgeting processes, and to establish coordination mechanisms for institutional monitoring tasks. Likewise, it is essential to set up an institutional framework that ensures that information on performance results provides feedback for the decision-making processes and ensures the sustainability of these efforts in the long term.

Institutional capacity development

To consolidate a culture of results-based management at SEDESOL, it is necessary to build institutional capacity. To achieve this goal, attempts will be made to (1) identify methodologies, including basic manuals for the construction of indicators and the application of assessment methods; (2) disseminate the conceptual underpinnings and tools of the system; and (3) train public officials in results-based management and M&E.

Conclusion

Since 2004, M&E has evolved into systematic activities that support a results-based management system. During

the first stage, a system with a greater emphasis on evaluation was conceptualized, with a focus on the management of individual programs.

In response to recent changes in the regulatory and institutional framework and lessons learned in the first stage, the model was redesigned. This model seeks to incorporate a vision of sector performance, linking M&E with the planning and budgeting process and using explicit critical transversal components, including an information agenda and institutional organization and capacity.

Building a results-based management system is a continuous medium- and long-term process that requires a combination of technical and political elements. It means generating relevant, timely, and quality information to be used in decision making. It requires shaping the institutional bases to allow connections among planning, programming and budgeting, implementation, and evaluation activities. An incentive scheme aimed at achieving results and mechanisms is important to ensure the sustainability of the initiative.

Gerardo Franco

Secretaría de Desarrollo Social, Sedesol, Mexico

Impact Assessment of Regional Development and Peace Programs in Colombia

Building an Evaluation Approach for Assessing Change in Social Capital and Conflict Management

Introduction

Policy makers increasingly are faced with the challenge of evaluating initiatives that attempt to bring about social change in “intangible” aspects. Designing social interventions is complex, since their objectives often go beyond the classic concepts of poverty and well-being. To achieve true change, social interventions must bear in mind and cope with the complexity of the social context and the sociocultural dimensions of the agents. Thus, in addition to reducing poverty, social programs increasingly seek to produce effects in values, aptitudes, perceptions, and motivations of its beneficiaries to modify the structure of the social fabric and to improve relations within communities, between

the agents and the market and between the agents and the state. To a certain extent, this requires assuming a (neo)institutional approach to the interpretation of development processes, in which the modification of the formal and informal rules of the game that define how agents interact with each other is as important as their objective conditions. The renewed popularity of concepts such as social capital, social networks, capabilities, and empowerment is proof of the fact that these dimensions of development processes play a key role in the political discourse and the design of social programs, and should therefore also be included in their evaluation.

In this paper, I will develop some ideas on the integration between quantitative and qualitative methods for evaluation, on the basis of the experience of the impact assessment of the Regional Development and Peace Programs (RDPPs) in Colombia.⁶⁹ The title refers to the urgent need to adjust our evaluation tools to record the effects of programs on sociocultural dimensions, what in Colombia we defined as relationship patterns or “modes of relationship.” An evaluation approach based on classic multidimensional indicators of poverty runs the risk of omitting these essential dimensions. It is here that the integration of evaluation methods becomes essential, as complementing qualitative and quantitative approaches enables one to reflect the complexity of the processes of social change under way.

From a quantitative perspective, it is possible to advance toward the development of new instruments to evaluate sociocultural dimensions, without diverging from the classical framework of rigorous impact assessment. In this presentation, I will concentrate on certain aspects of the progress achieved precisely for this purpose in evaluating the RDPPs in Colombia.

Regional development and peace programs in Colombia

The RDPPs are networks of organizations that seek to create conditions of peace and development in the regions most affected by violence and poverty. They posit that peace is the result of the consolidation of fair and inclusive development models, and they defend the principles of respect for life, human dignity, solidarity, and participatory democracy.

The RDPPs play a similar role to that of “social funds,” whereby a stock of resources is used to finance a heteroge-

neous set of projects in the areas of (1) productive development, (2) institution building, (3) human rights, and (4) culture of peace. Using a bottom-up participatory model, the projects are promoted and implemented by grass-roots social organizations and express the views, priorities, and identities of local communities (“the life we want”). Beneficiaries are usually members of the same grass-roots social organizations. In practice, the RDPPs promote the creation of a coordinated structure of social networks among community members, among organizations, and between organizations and local public institutions.

The RDPPs’ work has been carried out with financing from various sources. In particular, Peace and Development and Peace Laboratories represent joint a joint effort by Colombia’s national government and international cooperation to support the RDPPs. Peace and Development functions as a result of a loan granted to Colombia by the World Bank, and the Peace Laboratories are a contribution by the European Union to these initiatives. The two currently support six RDPPs in Magdalena Medio, Norte de Santander, Oriente Antioqueño, Macizo Colombiano and Alto Patía, Meta, and Montes de María.⁷⁰ In these regions, the RDPPs finance nearly 1,400 projects, implemented by approximately 800 grass-roots organizations in 120 municipalities, with an estimated population of 180,000 beneficiaries.

How to evaluate the regional development and peace programs?

To encourage debate and reflection on the contributions and achievements of the RDPPs, in 2005 the Colombian government, through the National Planning Department, began implementing an ambitious evaluation plan.⁷¹ The evaluation was intended to apply a rigorous methodology and use quantitative indicators to assess the programs’ impact.

The initial challenge for evaluating the programs involved developing a suitable theoretical approximation for evaluating the expected changes. The complexity of the task was due to the peculiarity of the intervention model. The RDPPs have ambitious, long-term objectives, embedded in the concepts of integral development and sustainable peace. Their impact strategy is both integral and multifaceted. In particular, the RDPPs seek to have successive levels of incidence, since they focus on a limited number of direct beneficiaries, and expect to modify territorial con-

ditions (municipalities and regions) through propagation mechanisms. Finally, the programs promote intangible processes that are difficult to measure.

The first step in undertaking the evaluation involved drawing up and reaching agreement on a “change theory” model of the programs that would make it possible to outline the main assumptions, logical stages, and channels of action of the intervention (Figure 30).

This conceptual reference framework posits that the conditions of violence and poverty characterizing the regions targeted by the RDPPs are directly linked to negative “modes of relationship” (distrust, clientelism, and individualism, among others), creating a vicious circle. Within this framework, the RDPPs’ proposal consists of supporting specific projects that encourage new “modes of relationship” based on positive values among direct beneficiaries (for example, trust, reciprocity, and collective action, among others). It is expected that these new relationship patterns, based on the dynamics of a virtuous circle, will lead to positive changes in the socioeconomic and security conditions of the participants in the direction of building peace and development. The model assumes that this virtuous circle would extend slowly and sustainably, spreading from the projects’ beneficiaries to their communities and local institutions, positively affecting an increasingly large population, and achieving a regional impact.

Based on this conceptual framework, the evaluation methodology focuses on establishing the programs’ contribution at the level of direct beneficiaries, and attempts to

find answers to the following question: How are the RDPPs modifying the living conditions of their direct beneficiaries and their “modes of relationship” with the community, toward the conflicts and toward public institutions? This question is thought to deal appropriately with the RDPPs’ transversal aims, regardless of the heterogeneity of the projects financed.

The next step in the construction of the “change theory” model involved designing an evaluation strategy and a battery of instruments that would permit the measurement of the changes achieved in the dimensions concerning conditions and “modes of relationship.” Whereas the first issue can be dealt with a classic evaluation approach based on quantitative indicators, it is more difficult to evaluate the effects on relationship patterns from a rigorous perspective. To that end, two complementary methodological approaches were used. On the one hand, the traditional framework of quantitative impact assessment was extended by using a set of “novel” measurement instruments. On the other hand, this main strategy was complemented by the qualitative methods with which it was combined.

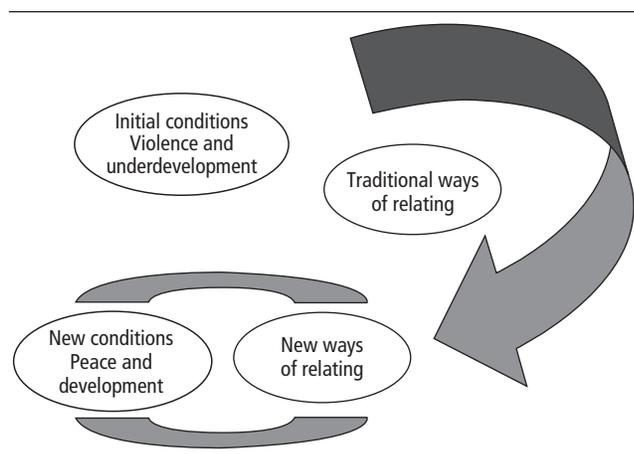
The quantitative strategy is based on a quasi-experimental design that permits a comparison of the outcomes of interest between a group of households benefiting from the RDPPs (treatment group) and a group of nonbeneficiaries (control group) in the same municipalities. The plan includes two data collection rounds: a base line and a follow-up. The first data collection occurred between November 2006 and February 2007, while the follow-up will take place in 2010. The design enables effects of the RDPPs to be estimated by using difference-in-difference models, and by comparing beneficiaries among themselves through intensity models. In addition to a structured survey, the innovative element included two extra instruments used to measure aspects associated with “modes of relationship:” a set of experimental games and a social dilemma. In what follows I describe these instruments and discuss how they can become a tool to measure the impacts of social interventions on sociocultural dimensions.

Social dilemma

Social dilemmas are information-gathering exercises in which participants are asked to describe options in a story narrated in the third person. They allow an exploration of the categories of right and wrong, or the processes of moral

FIGURE 30:

RDPP Change Theory



Source: DNP 2009.

deliberation on topics for which using direct survey questions may yield biased answers. Within the RDPPs' evaluation framework, a social dilemma was used to measure changes obtained in the modes of relationship toward conflicts (conflict management). For this end, the story of Juana, a housewife and farm worker was developed to illustrate a dilemma. Boxes 6 and 7 review the story and the dilemma.

At the end of the story, Juana, faces a dilemma in which she has to choose from a series of alternatives to deal with a direct threat to her life that she receives from an illegal group. Participants listened to a recording of the story and were asked to express their opinions on Juana's context and situation and evaluate various choices of action for her (see Box 7).

The purpose of the exercise is to help the participants identify with the situation, or at least relate the story to their individual context. When the participants express their opinions on what Juana should do, they reveal their processes of deliberation and unravel their views on what would be the fairest or best way to act in a situation such as the described. Threats by illegal armed groups (guerrillas and paramilitary groups) are common in the regions and

BOX 6:

Juana's story

Juana was "a housewife, an ordinary farm worker, until the day some violent men killed Jacinto, her husband. . ."

"the neighbor and her husband, who were vegetable fiber growers, convinced Juana to start working with them . . ."

"her reputation grew until one day, Pedraza, a well-known councilor and politician, suggested that she run as candidate for the Council."

"however, before she took office, anonymous messages began to arrive, addressed to Juana, with death threats from the same armed group that had killed her husband . . ."

Source: Econometría, SEI-IFS. Informe Metodológico para la Evaluación de Laboratorios de Paz y Paz y Desarrollo (2006).

BOX 7:

Alternative responses to the social dilemma

1. "Juana should meet the illegal group on her own, to find out for herself why they want to kill her."
2. "If the community speaks to those making the threats, Juana can maintain her peace of mind and continue her political activity."
3. "Juana should go to the authorities to accuse Pedraza of having links with illegal groups."
4. "The main result of Juana's becoming a community leader was the increase in the risk of her being killed."

Source: Econometría, SEI-IFS. Informe Metodológico para la Evaluación de Laboratorios de Paz y Paz y Desarrollo (2006).

communities where the RDPPs operate. Indeed, the programs aim at modifying participants' attitudes toward conflict management by increasing their capacity to act collectively through social networks.

Juana's alternatives, from which participants are asked to choose, illustrate the main types of conflict management of (1) individual conflict management, (2) collective conflict management, and (3) institutional conflict management. It is also possible to elicit other features of participants' aptitudes regarding sensitive issues, such as (4) the risk associated with leadership in a context of violence (letters refer to the answer options in Box 7).

A preliminary analysis of the answers to "Juana's Dilemma" based on intensity models shows that the RDPPs are encouraging community and institutional forms of conflict management.⁷² In particular, it reveals an ideal continuity among the answers encouraged by the programs, because initial community-based responses tend to lead in the longer run to institutional conflict management options. This confirms the impact of the initiatives on the improvement of democratic governance. Moreover, participation in the RDPPs reduces the perception of risk associated with leadership.

The experience of applying the social dilemma opens a wide-ranging scenario to measure aptitudes and capabilities

related to sociocultural dimensions of development processes. Within this perspective, it is useful to consider some key elements of the debate about the scope and relevance of the instrument. The main issue concerns the existing relationship between the answers given regarding a hypothetical scenario and the agents' actual behavior. For instance, participants tend to respond positively to choices involving institutional conflict management, while conflicts actually may not be handled this way. It is possible that answers given in a hypothetical scenario reflect views about socially acceptable (correct) conflict management solutions and depend partly on criteria of social desirability. This does not necessarily translate into real actions when people face the actual constraints, costs, and benefits of their decisions. In this respect, it is important to complement information on the processes of moral deliberation with questions on actual behavior. As a corollary to this general point, there is still a need to understand the role played by certain features of the design of social dilemmas: the fact that the story is written in the third person, the balance between generality and accuracy in the description of the context and the events, and the opportunity to use multiple choice answers. Lessons can be drawn from the previous use of similar tools in social psychology.

Experimental exercises

To measure attitudes and behaviors in the domain of the "modes of relationship" toward the community, two experimental exercises were carried out as part of the evaluation of the RDPPs. Experimental exercises constitute a possible approximation to the topic of "social capital" from a motivational (preference based) perspective and are based on the theoretical framework of experimental economics and game theory.

The instrument is based on the fact that people tend to react differently when it comes to answering direct questions on issues such as trust or contribution to the public good, compared with how they would act in actual situations. The tool therefore consists of simulating scenarios in which participants face real dilemmas. Players' decisions are turned into monetary payoffs according to the "rules of the game" of the exercise, and the exercises are carried out in a "controlled" fashion so that the decisions mainly involve consequences of an exclusively "economic" nature. An analysis of the decisions made in these simulated sce-

narios enables one to draw conclusions about participants' motivations, values, and attitudes. Two experimental exercises were carried out within the framework of the RDPP evaluation: a trust game and a public good game.

The public good game is carried out through a group dynamic. Each participant receives an initial monetary endowment and has to decide whether to invest it in a private or a public account. Every "peso" invested in the public account produces equal benefits for every member of the group whereas every "peso" invested in the private account only benefits the owner of the account. The dilemma stems from the fact that there is a conflict between self-interest and optimum social outcome. If everyone invests in the public account, the total earnings are greater than they would be if everyone invested in his own account (Pareto superior social outcome). However, the optimal strategy for the individual is always to invest in a private account (self-regarding Nash equilibrium).

Beneficiaries and nonbeneficiaries participate in the public good exercise in mixed groups of 40 people distributed equally. The decision is taken in two rounds, first confidentially and then after 10 minutes of free interaction between the participants. This makes it possible to measure the effect of communication. By sharing information and establishing agreements on the game strategy, participants may be able to more effectively achieve the optimal social outcome or converge on the Nash equilibrium.

The decision dilemma raised in the public good exercise is particularly relevant to the RDPPs because they precisely (re)create conditions that enable people to act collectively in focused regions by rebuilding bonds of trust in the community. An analysis of the behavior observed indicates that the level of contribution to the public good of the municipalities where the RDPPs operate is much higher than that recorded in other municipalities in the country with comparable characteristics. Moreover, the effects have a high degree of "propagation" and spread within municipalities through spillover mechanisms, in line with the assumptions of the Programs' "change theory."⁷³

Although experimental exercises traditionally have been used in a laboratory context, their use in the field has been spreading rapidly. The potential to evaluate the effect of social interventions on sociocultural domains is conditional to understanding the type of motivations identified

with the instrument. As for the social dilemma, this in turn points to the need to examine the relationship between the behavior adopted in the simulated scenario and the agents' actual decisions.

In its more simplified version, economic theory assumes that people act according to the calculation of the economic consequences of their actions, where the expected payoffs of decisions only count insofar as they affect the decider (individualism). Experimental exercises emerge as an attempt to measure deviations from this theoretical approach, testing the extent to which agents diverge from the paradigm of *homo economicus*. To achieve this aim, the exercises are designed in such a way that the consequences of decisions are limited exclusively to the monetary spectrum (through anonymity and the definition of the pay structure), and certain institutions (rules of the game) are established, according to which there is a trade-off between individual and collective monetary payoffs.

If the observed results are different to what economic theory predicts, we can conclude that people value the monetary consequences of actions in a non individualistic manner (*other regarding preferences*). However, there are certain limitations to the interpretation of agents' behavior in experimental exercises. First, the exercises purposely exclude noneconomic consequences of actions (such as the sociocultural payoff of cooperation), whereas the consequences of actual decisions are not only economic and may involve cultural or social dimensions (the effects of reputation, the esteem of peers, the importance of enjoying these in the future, and so on). Second, unless their design improves, the exercises do not allow differentiating between extrinsic motivation (based on the consequences of actions) and intrinsic motivation (based on obedience to principles—one's own criteria on what is fair or unfair, good or bad—regardless of the results of actions). Third, it is not clear whether the monetary payoffs used in the exercises constitute a big enough incentive to mimic the economic consequences of actual decisions.

This discussion indicates the need to advance toward improving the design of the experimental exercises to capture and differentiate between a broader spectrum of motivations. At the same time, it complements experiments with other data collection tools that provide information on the occurrence of positive "modes of relationship," such as trust or cooperation in the actual context.

Conclusion: The need to integrate methods

Social dilemmas and experimental exercises constitute two promising lines of work for applying a quantitative framework to the appraisal of the effects of complex programs on the sociocultural dimensions of development processes. However, to have a broader and deeper understanding of the transformations under way, it is essential to ensure that they are complemented by qualitative methods. Using a sequenced combination of quantitative and qualitative approaches, and ensuring effective feedback between the two, will make it possible to develop and refine hypotheses on the channels of impact of the interventions, adapt the design of the instruments to the hypotheses of change, assess the scope of instruments, contextualize the interpretation of results, and triangulate information in depth.

In the case of the evaluation of the RDPPs, developing the change theory model has been the result of a participatory and essentially qualitative process. Moreover, the second phase of evaluation reveals the need to integrate the principal strategy with qualitative methods to (1) validate hypotheses on the programs' effect on relationship patterns; (2) enhance understanding of the scope, potential, and limitations of experimental exercises and the social dilemma; (3) explore the negative and unexpected impacts; and (4) analyze the effects of horizontal propagation of the programs. The Colombian government, through the National Planning Department, is committed to this task to contribute to the search for, identification of, and dissemination of best practices in evaluation.

Luca Pellerano

Departamento Nacional de Planeación-Colombia, Oxford Policy Management and Institute for Fiscal Studies, England.

Progress and Challenges of Monitoring and Evaluation Systems in Latin America and the Caribbean

Introduction

This paper examines recent progress in the creation or strengthening of monitoring and evaluation (M&E) systems in Latin America and the Caribbean. M&E systems are a fundamental pillar of Management for Development Results (MfDR). This analysis uses two parameters:

(1) the work conducted by the Inter-American Development Bank's (IDB) Program to Implement the External Pillar of the Medium Term Action Plan for Development Effectiveness (PRODEV) and the PRODEV Evaluation Tool (PET), and (2) a survey by PRODEV conducted in late 2009.⁷⁵ The PET provides a diagnosis of M&E systems while the survey reveals the trend observed in late 2009.

MfDR and M&E systems

Management for Results (MfR) is one of the instruments and approaches that emerged from the new public management efforts to strengthen the government's capacity to promote development. MR is defined here as a frame of reference aimed to support public organizations to effectively and comprehensively improve the creation of public value (results), ensuring maximum efficacy, efficiency, and effectiveness in their performance, the achievement of government objectives, and the continuous improvement of their institutions (IDB and CLAD 2007).

In the Organisation for Economic Co-operation and Development (OECD) countries, MR is used to guide government actions toward maintaining the level of development already attained. In contrast, in the countries of Latin American and the Caribbean, the emphasis is on accelerating development, hence the term Management for Development Results (MfDR).

MfDR revolves around the creation of public value, which includes observable and measurable social changes the government introduces in response to social needs or demands that emerge through the democratic process. As a result, these changes "make sense" to the people. The ultimate goal of MfDR is to endow its organizations with the capacity to manage public value creation and achieve the government program's objectives.

In MfDR, results relate to the social changes resulting from government actions and are not restricted to the activities or outputs that contribute to such changes, which often taken as evaluating parameters for government actions. For example, the results of the Ministry of Education will not be measured by the number of schools built, teachers trained, or graduating students, but by the learning results actually achieved and, in the long run, the quality of the job offered based on the education received.

Social changes do not only refer to changes in social policies like health and education, but include other pub-

lic policies such as citizen security, agriculture, energy, employment, and justice. All of these policy changes create public value and induce social changes.

From the point of view of public value creation, market regulatory functions are in the same category as the provision of public goods. Thus, the establishment of competitive markets allowing for the provision of goods in a competitive scenario that makes those goods available at competitive prices also contributes to the creation of public value.

As a management strategy, MfDR requires decisions to be based on reliable information about the effects of government actions on society. Its implementation will therefore depend on how the changes resulting from such actions are measured, that is, the existence of instruments to capture these variations, systems to process the information and procedures to include data analysis in decision making. But these elements are not always found in government administration.

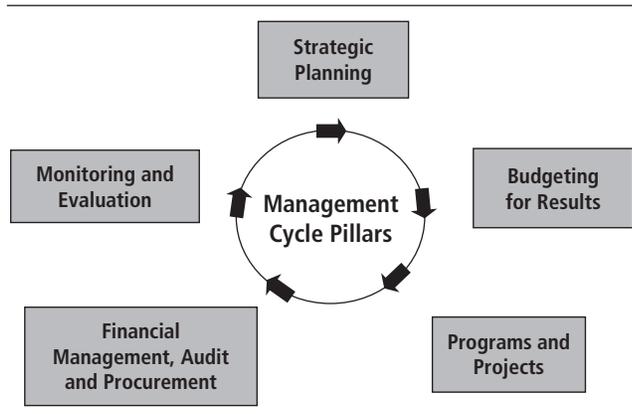
MfDR encourages a sense of responsibility of officers and institutions to achieve results and not for just performing their duties. This promotes stronger and more substantial verifiable evidence-based accountability. This is a major change from traditional management and, consequently, requires explicit political will from the highest government level and a unit with enough power to promote change throughout the public sector. Although MfDR is primarily a frame of reference for management, and its implementation should reside with public authorities and officers, the gradual appropriation by mid-ranking officials and those responsible for service provision is key.

Finally, results must come first. In other words, the process states the results to be achieved and assigns resources and processes accordingly. This is important and contrasts with the bureaucratic approach, which starts with available inputs and known processes to identify potential results.

PET methodology

PET assesses the institutional capacity of countries in the region to implement MfDR. The management cycle pillars (see Figure 31) include the elements that must be part of the cycle to align it with MfDR. These elements fall into three categories: (1) components (16), (2) indicators (37), and (3) minimum or variable requirements (141). Each element is rated from zero to five. The average scores indicate a country's MfDR capacity.

FIGURE 31:

Management Cycle Pillars

Source: Author.

This instrument analyzes the management cycle at national central government level only, not at subnational level. Neither does it look at autonomous entities or public companies. The sample included 25 countries from the Latin American and Caribbean region (see Table 21).

Results reveal great differences among the various countries of the region, which were aggregated into three groups according to their MfDR score. The first group includes countries with a score of 3 or higher, the second includes countries with scores between 1.5 and 3, and the third is made up of countries with scores lower than 1.5. More than half of the countries (16) are in the medium group. The average MfDR score for the region is 2 out of 5, indicating that it is in the early stages of MfDR implementation.

There are noticeable differences in the degree of progress between the different countries and also between MfDR pillars. A comparison of the pillars shows that financial management, auditing, and procurement (FMAP) is the most developed. This is a direct consequence of the efforts deployed by governments in the region and the in-

ternational cooperation to improve public financial management during the last decade. Initially, these efforts were part of the government reform measures of the 1990s that aimed to overcome the economic crisis. These measures had a strong government bias and did little to improve managing capacity. Today, these efforts stem from the conviction that efficient management of public resources is essential to a country's development.

Strategic planning is also in a better position than the other pillars because most countries, after abandoning it in the 1980s and 1990s, have once again embraced planning as an instrument of development that targets strategic priorities with greater emphasis on operational efficiency and participation.

The programs and projects management pillar still shows poor performance, despite the fact that over the last two decades, countries in the region have implemented a wide range of social programs to fight poverty and other social ills. These programs generally have been financed by international organizations and were aimed at improving the quality of education and health services. Even though externally funded programs are backed by appropriate technical, economic, and financial feasibility studies, it seems that they have failed to change the management culture of the line ministries that produce activities and outputs without much attention to results.

The pillars with the lowest scores are budgeting for results (BfR) and M&E. Both pillars are closely related, as BfR is based on the information provided by M&E systems that follow up on government actions. Only those countries in the advanced group show good capacity in these pillars.

The high-scoring countries—Brazil, Chile, Colombia, and Mexico—have implemented innovative management. They have a big lead over other countries in all pil-

TABLE 21:

Country Ranking by Degree of MfDR Development

MfDR development	Countries	MfDR Score
Advanced	Brazil, Chile, Colombia, Mexico	>3
Medium	Argentina, Barbados, Bolivia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Honduras, Jamaica, Nicaragua, Panama, Peru, Trinidad and Tobago, Uruguay	>1.5 to <3
Low	Bahamas, Belize, Haiti, Paraguay, Suriname	>1.5

Source: Author.

TABLE 22:

MfDR Pillars by Country Group

Pillar	Average	MfDR development		
		Advanced	Medium	Low
Strategic planning	2.4	3.5	2.3	1.9
Budgeting for results	1.5	3.1	1.3	0.5
Financial management, auditing and procurement	2.6	4.0	2.5	1.3
Programs and projects management	2.0	3.1	2.0	0.8
Monitoring and evaluation	1.7	3.8	1.3	0.8
MfDR Score	2.0	3.5	1.9	1.1

Source: Author.

lars, particularly BfR and M&E, without which no MfDR implementation can take place. In this group, pillars are also more balanced, thus confirming the integrality of pillars as systems become more developed (Figure 32).

As a whole, the region shows heterogeneous development of MfDR capacities. At one end a small group of countries are implementing measures that gradually lead them to systems for which results will be at the heart of public sector management. At the other end are countries with basic systems, some even inserted in a traditional, bureaucratic management framework. Most countries in the region are between these two extremes, sharing what may be called an incipient phase of MfDR development. Their financial management systems show greater progress,

whereas M&E and BfR are the least developed. Chile leads the pack in the region thanks to its perseverance in pursuing management innovation.

Monitoring and evaluation systems

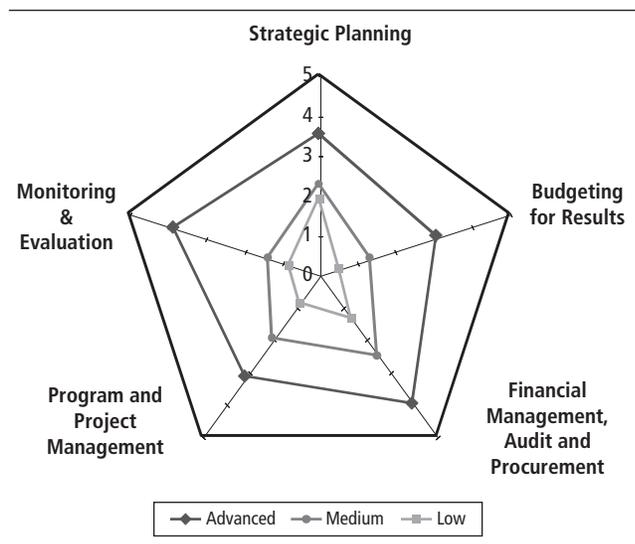
Monitoring is the “continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indicators of the extent and achievement of objectives and progress in the use of allocated funds” (OECD 2002). The goal is to ascertain progress in the achievement of government objectives and goals as stated in its policies, programs, and projects. The monitoring function is thus closely related to the functions of planning and budgeting. In MfDR, the monitoring system manages the public sector by means of indicators that verify the achievement of objectives and their quantitative goals. It analyzes both the achievement of objectives and the resources used in achieving them.

A traditional follow-up system uses indicators to inform on inputs, activities, and outputs, without considering whether they contribute to fulfilling the objectives (results). In contrast, a results-based monitoring system uses indicators that measure the results obtained through programs and projects.

Monitoring is an overarching function of the management cycle because each pillar has a role to play in its execution. For example, the planning pillar seeks to establish management objectives. BfR takes part in goal setting and public financial management provides information and data analyses. For its part, the monitoring system provides information on the different levels of public policy execution,

FIGURE 32:

MfDR Score by Pillar and Country Group



Source: Author.

such as services, programs, institutions, and policies. An overarching and integrated monitoring system imposes significant challenges to the articulation of the monitoring system, which may be overcome only through a continuous improvement process.

Evaluation is the

systematic and objective assessment of an on-going or completed project, program or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors. (OECD 2002)

The main difference between monitoring and evaluation is that monitoring provides information on the state of achievement of the objectives and effects of a policy, program, or project, while evaluation explains why these objectives or effects are being (or have been) fulfilled, or not. Evaluation uses systematic data collection and relevant information analysis processes to explain the causes and reasons for the results, examine unintended results, study the process used to obtain them, and provide recommendations for future actions.

Monitoring systems

Of all the countries studied, only three do not have a dedicated unit to follow up on government objectives and goals. Monitoring units fall under the Presidency of the Republic or the Ministry of Planning and, to a lesser degree, the Ministry of Finance. The trend to place this function under the Presidency of the Republic is a reflection of the growing interest of government officials to count on a set of accountability and/or management instruments. The monitoring systems that report to the ministries of finance are found in countries that have either consolidated (Chile) or initiated (Peru and Mexico) their results-based budgeting systems.

Although most countries have an entity that is formally responsible for monitoring, the system is currently operational in only 13. Over the past decade, most have

completed the process or are still in the experimenting, designing, or implementing stages. This reflects the sparse institutional development of the monitoring function in the “medium” group of countries. Another significant shortcoming of such systems in these countries is the lack of results indicators. In the best of cases, there are output indicators, but this limits MfDR implementation because it does not show whether these outputs have had the expected results. For example, building schools and delivering school books (outputs) do not necessarily lead to an improvement in children’s reading or writing skills (result) measured through standard learning tests.

The data collected reveal the wide gap in monitoring between industrial countries and the rest. In fact, the former have a much higher score than the latter and exhibit less of a difference between the institutional frame of the monitoring function and the use of information, a clear indication of their advanced stage. Nevertheless, in all cases, the indicator for information use and dissemination had the lowest score, showing that the mere implementation of a monitoring system does not necessarily entail the use of information with management purposes. In this sense, it appears that several countries have stressed the preparation of indicators and the construction of data applications but have failed to create regulations and build institutional capacity to apply these tools in public management analysis and conduction.

With the exception of Chile, no other country has implemented institutional criteria and procedures to analyze and manage government actions based on the information provided by the monitoring system. This does not mean that other countries ignore the information, but they use it in a noninstitutionalized manner, without formally integrating it to the different stages of the management cycle. For example, in those countries where the monitoring system falls under a unit of the Presidency of the Republic, senior officials often base their decisions on the information provided. However, this use is discretionary and depends exclusively on the will of the current administration. Similarly, in countries that feature public investment systems, the use of project monitoring information tends to be regulated, although these systems are restricted to investment projects that are only a portion of the government plan. Finally, those countries that only recently began developing a results-based budgeting framework still lack perfor-

mance indicator systems that are sufficiently robust to be used formally.

One important weakness present in most monitoring systems is the absence of vertical (government-organization-operation unit) and horizontal (planning-budgeting-execution-M&E) articulations. Prior studies have determined that most monitoring systems, especially those connected to the presidents' offices, are used primarily by high-ranking authorities and only marginally by public officials to manage their organizations (Cunill and Ospina 2008). Monitoring systems are also weakly connected to the budget process and to program and project management. This suggests that many countries see their monitoring systems more as an accountability mechanism at the top of the government hierarchy than as a management instrument that permeates the entire institutional tissue.

Evaluation systems

Of all the countries studied, 13 have instituted laws that in one way or another make public management evaluation mandatory. But only five (Brazil, Chile, Colombia, Mexico, and Peru) have a legal frame that specifically requires policies, programs, and projects to be evaluated. The remaining countries have evaluation regulations that do not look at the results of government actions in the sense discussed here. In most cases, these regulations refer to financial evaluation or analysis of achievement of the projects' physical goals.

Since 2003, Chile (2003), Colombia (2004), Brazil (2006), and Mexico (2007) have introduced specific

norms on the evaluation of policy, program, and project results. Only Peru lacks specifically designed institutionalized mechanisms because the regulations were established only recently and the system is still in the beginning stages. In the other four, evaluation mechanisms and methodologies are formally operating together with an evaluation program with a budget of its own. This is a clear indication that systematic ex post evaluation of policies, programs, and projects is a new phenomenon in the region's public policy landscape.

Also worth mentioning is the fact that the four countries have placed a high priority on the evaluation of social programs and projects. Moreover, the evaluation systems of Brazil and Mexico fall under the social sector and, until now, only evaluate social actions targeting social development and poverty fighting. Even though the Colombian and Chilean systems reside in the National Planning Department and the Ministry of Finance respectively, they also focus on social programs.

A number of differences also appear among these four systems. In Chile, the articulation capacity of this system is both vertical (government, organization, personal) and horizontal (planning, budget, execution, and M&E), but the rest are still affected by bottlenecks. The Chilean system is therefore capable of evaluating policies as well as programs and projects, and of systematically considering evaluation results in budget preparation and service management.

The M&E and the Budgeting for Results pillars are the least developed in the Latin America and Caribbean region.

TABLE 23:

M&E Indicators by Country Groups

M&E indicators	Average	MfDR development		
		Advanced	Medium	Low
Monitoring institutions	2.2	4.5	1.9	0.8
Extent of program and project monitoring	1.5	4.0	0.9	1.0
Use and dissemination of monitoring information	1.3	3.7	0.9	0.2
Statistical information systems	2.8	4.1	2.8	1.8
Legal and institutional frameworks for evaluation	1.7	3.9	1.4	0.8
Extent and articulation of the evaluation system	0.7	2.8	0.3	0.1
Actions derived from failure to achieve goals	0.9	3.1	0.4	0.3
Dissemination of evaluation results	1.4	4.6	0.8	0.4
Total	1.7	3.8	1.3	0.8

Source: Author.

Even though most countries have implemented systems to monitor the achievement of government objectives, these are still in the early stages of development. Their greatest limitations to effectively contribute to MfDR are (1) the lack of formally established technical standards and work methodologies, (2) the lack of coordination between goal-objective achievement information and resource investment information, and (3) the absence of result indicators. Interestingly, evaluation systems appear to be less developed than their monitoring counterparts, for only a few countries have institutionalized ex post evaluation of programs and projects. With the exception of the four countries that feature advanced MfDR implementation, countries lack in evaluation capacity. This is a great challenge for MfDR capacity in the region because the absence of a system that evaluates government management results constitutes a major hurdle in the process of strengthening the strategic planning, budgeting for results, and program and

project management pillars. Without evaluation, it is not possible to guild planning and results-based budgeting or to generate a continuous improvement culture in the provision of public services.

Results of the PRODEV survey

A recent survey by PRODEV describes a number of initiatives in the region with positive results in strengthening and consolidating M&E systems. The first concern was to determine whether expenditure monitoring had increased. Out of the sample, 71 percent of respondents believe that expenditure monitoring had increased as a percentage of total expenditure and only 7 percent believed it had decreased.

The survey also asked whether the number of corrective actions to address the failure to meet goals or achieve results has increased. This information is important because it provides information on how evaluations are used. In

BOX 8:

Information dissemination

The Internet is an effective tool to disseminate the relevant information the government produces. This increases public management transparency and helps promote the involvement of civil society. Countries produce large amounts of information that they make available to the public through the portals of the various ministries. The most widely disseminated information has to do with the budget, financial statements, government procurement, audit reports, and the National Medium-Term Plan. There is less information on monitoring, management reports and evaluations partly because countries generate little information in this area.

Information available through the Internet	MfDR development			
	Average	Advanced	Medium	Low
Information on the budget	3.3	5.0	3.4	1.3
Information on financial statements	3.3	4.5	3.2	2.3
Information on bids, proposals and awards	3.3	4.3	3.3	2.3
Audit reports	2.7	4.8	2.6	0.8
National Medium-term Plan	2.4	4.5	2.1	1.3
Monitoring system	1.5	4.0	1.1	0.0
Reports on sector management results	1.4	3.2	1.4	0.0
Reports on ex post evaluation	1.4	4.5	0.9	0.0
Reports on ex ante evaluation	1.0	2.0	0.9	0.0
Total	2.3	4.1	2.1	0.9

Source: Author.

this case, 55 percent of respondents believed that corrective actions had increased, whereas 35 percent perceived no changes, and 10 percent thought they had decreased.

As to whether access to evaluation reports by the legislature or civil society had improved, 73 percent of the sample believed that use by legislators and activists had grown,

BOX 9:

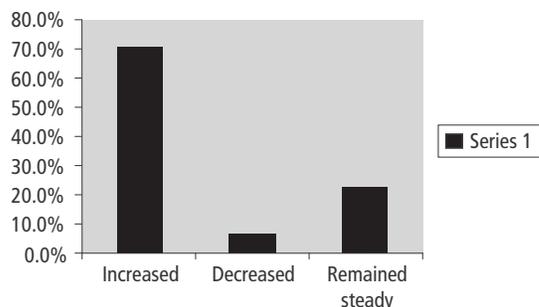
PRODEV survey results

PRODEV Survey 2009

The percentage of expenditure monitored in relation to total expenditure

Increased	70.9%
Decreased	6.8%
Remained steady	22.4%

The percentage of expenditure monitored in relation to total expenditure

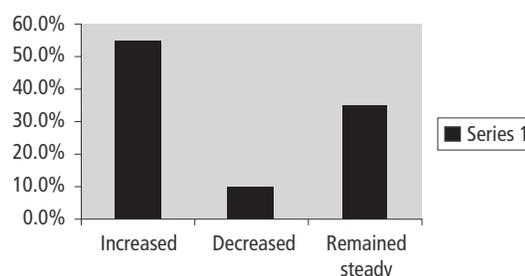


PRODEV Survey 2009

The technical, administrative and financial actions introduced to correct the failure to meet goals and objectives detected during the evaluations

Increased	55.0%
Decreased	10.0%
Remained steady	35.1%

The technical, administrative and financial actions introduced to correct the failure to meet goals and objectives detected during the evaluations

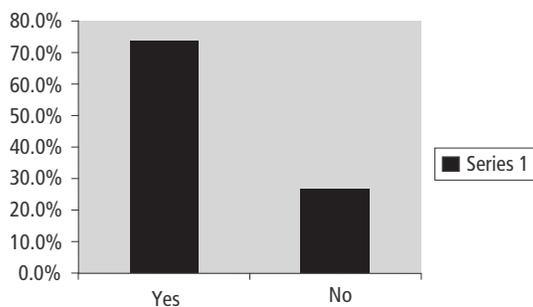


PRODEV Survey 2009

Do the Legislature and Civil Society have greater access to evaluation reports?

Yes	73.5%
No	26.5%

Do the Legislature and Civil Society have greater access to evaluation reports?

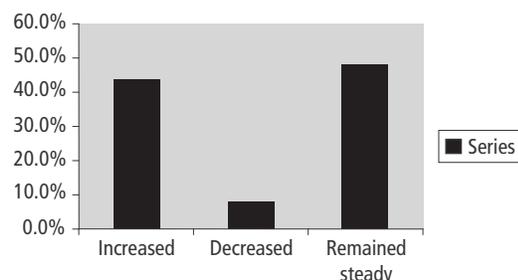


PRODEV Survey 2009

Human resources availability and capacity in Monitoring and Evaluation

Increased	43.8%
Decreased	8.0%
Remained steady	48.2%

Human resources availability and capacity in Monitoring and Evaluation



Survey: Author.

while 26.5 percent believed it had not. In addition, human resources are not being assigned in sufficient numbers and quality to perform M&E functions properly. Among the respondents, 48.2 percent felt that the assignment of public sector human resources to M&E tasks had not changed and 8 percent felt that the number had declined.

Finally, the survey showed improved coordination between executing agencies and evaluation institutions: 66 percent of respondents believe articulation between these institutions has increased.

Conclusion

The Latin America and Caribbean region is experiencing some good momentum for the institutionalization and strengthening of M&E systems. The region must use this momentum to good advantage and continue facing, with perseverance and continuity, the many challenges ahead. To that end, the following are some conclusions that may be drawn from this analysis.

First, M&E systems remain one of the weakest pillars of MfDR, therefore, strengthening and consolidating them is crucial to the process of MfDR implementation. In this sense, there is a growing interest in building and developing M&E systems in the region. Within these systems, monitoring is at a more advanced stage than evaluation and the proportion of expenditure monitored has grown.

Second, in those countries that have implemented M&E systems, the information generated is barely used for decision-making purposes in the public sector, although the tendency is to increase its use. Efforts should be made to further involve political actors to expand the use of the information generated by M&E systems. At the same time, it is important that nonexperts can access this information in a friendly, accessible, and timely manner.

Third, dissemination of M&E results to congresses and civil society will improve coordination among executing bodies and agencies and the institutions responsible for M&E. Dissemination of information to congresses and civil society must be continuous and permanent. Also important is to disseminate information in a timely manner to make it an input for analysis and decision making as well as for the generation of proposals that may replace or supplement current public policies and programs.

Finally, improvements in the number and quality of human resources assigned to M&E tasks are not appar-

ent. Training for human resources assigned to M&E tasks and the financial resources of agencies responsible for these functions should be reinforced.

Roberto García López

Prodev Coordinator, Inter-American Development Bank

The Brazilian Monitoring and Evaluation Network: A Report on the Creation and Development Process⁷⁷

This paper discusses the background that led to the creation of the Brazilian Monitoring and Evaluation Network and some of its recent advances and future challenges. The public management evaluation process began in Brazil in the 1990s. The Master Plan of the State Apparatus Reform, issued in September 1995, described management as a core tool to consolidate economic stability and ensure sustainable economic development. It reinforced the concepts and reforms needed for a more efficient government. Results-oriented management emerged in Brazil soon after the development of the Master Plan and at the same time as efforts to control inflation and stabilize the economy (which improved the market's capacity to plan its investments and future projects). This movement resulted in important effects on the Brazilian federal government. It influenced the work of the core and sector-oriented areas, and of internal and external controls, in addition to the activities performed by the international organizations and civil society. The adoption of these new public management procedures was anchored in the 1988 Constitution's provisions (Art. 165) which set forth the rules to govern the federal budgeting process by introducing (1) the Multiyear Plan (PPA), (2) the Budgeting Guidelines Law (LDO), and (3) the Annual Budgeting Law (LOA). The state constitutions and municipal organic laws reproduced the Federal Constitution's provisions. As an input to new public management methods, the PPA represents a coming together of planning and budgeting by (1) implementing the medium- and long-term government planning, (2) coordinating government actions, and (3) setting guidelines, objectives, and goals to the public administration to guide the allocation of public resources.

The Ministry of Planning, Budgeting, and Management (MPOG) has been the main actor in this effort. For

the last 10 years, the MPOG has worked to incorporate a results-oriented focus into sector-oriented public policies. The effort started with the 2000–03 PPA through three major innovations. The first was to make investments based on the land planning studies, resulting in a portfolio of 952 projects with a budget of Brazilian reais (R\$) 317 billion. Classified by macro sectors (Social Development, Economic Infrastructure, Environment, and Information and Knowledge), it considered the integration of activities, liability for results, and allocation of resources by the private sector. The second innovation was to include indicators on implementation and outcome, promoting the effective emergence of a results-oriented management culture and concepts of monitoring and evaluation (M&E). The third was to assign a manager to the program and introduce the practice of results-oriented responsibility. With that, the MPOG promoted an entrepreneurial culture in the federal government, introducing programs and project management tools, and training the public managers in the use of such tools. The publication of the Methodological Guide to program indicators stands out as an example of the program's progress (SPI/MPOG 2010).

The sector-oriented ministries have begun efforts toward M&E policies in the fields of education, health, and social development. Created in 2003, the Ministry of Social Development and Hunger (MDS) was the first ministry to assemble a specialized bureau in information management. The bureau has produced studies, promoted training, and implemented systematic methods to monitor policies on cash transfers, social assistance, and food security (Vaitsman, Rodrigues, and Paes-Sousa 2006). The Information Evaluation and Management Bureau is responsible for the bulk of studies, baseline research, beneficiary surveys, and impact assessments produced to support the social policies (Wenceslau and Hobbs 2008).

The federal government's internal control, in turn, has undergone an institutional strengthening process, which started in 2002 upon the establishment of the Office of the General Comptroller (CGU). The CGU resulted from the merger of the Federal Internal Control Bureau (SFC) and the Internal Control Coordination Commission (CC-CI), with the status of ministry. The CGU's work has fostered public information transparency and quality, focusing on prevention, inspection, and control of federal resources. The CGU's tools include (1) inspection of municipalities

by lottery, (2) annual financial auditing, (3) investigation of cases, and the (4) federal government general balance. Ferraz and Finan (2008) have shown that Brazilian electors, when well informed, reject those politicians who have been charged of corruption in favor of the honest candidates. The publication of the CGU auditing results in the municipalities where many cases of corruption have been found reduced by 20 percent the chances for mayors to be reelected.

Through the Federal Government Auditing Court (TCU), the legislative power began measuring the performance of government policies with a focus on external control, in addition to the traditional focus on compliance and the law. In 1999, the TCU began implementing operational audits, incorporating case study techniques, surveys, and focus groups to assess the programs' implementation. The inclusion of performance audits among those carried out by the TCU resulted from the technical assistance provided by the U.K. Department of International Development (DfID) as well as from exchanges of experiences with the U.K. National Audit Office (NAO) and the U.S. Government Accountability Office (GAO) (Flint 2004).

In the last few years, some states and municipalities have also shown increasing interest in the M&E agenda. Among these, the state of Minas Gerais, through its Planning Bureau, State for Results Unite, and the João Pinheiro Foundation (JPF), has developed and promoted a set of innovative practices in public management, with a strong emphasis on the use of indicators and goals to provide focus and direction to public administration (SEPLAG/MG 2008, 2009).

Among Brazil's state institutions, the work developed by the SEADE Foundation,⁷⁸ is also worth mentioning. In the last decade, the Foundation has specialized in creating or improving methodologies and tools for the formulation, monitoring, and evaluation of public policies and programs, notably in the social area. The foundation's activities provided inputs to different public management stages, by assisting in outlining programs and policies, and facilitating eventual adjustments and path corrections. To support the agencies that deliver or coordinate the government activities, these activities include assessment studies of specific public programs; construction of monitoring methodologies, defining indicators, to implement the policies' or public agencies' monitoring; and, with similar objective, the structuring of broad systems on management information.

Brazilian development banks, notably the *Banco do Nordeste do Brasil* (BNB), are also working in the field of M&E. Among some outstanding initiatives are the Crediamigo program evaluation, carried out in partnership with the World Bank, and the assessment of the Northeast Financing Constitutional Fund (FNE). The FNE program includes an assessment system that uses primary and secondary data. It has assessed the FNE-Agroindustrial, the FNE-Comércio-Serviços e Industrial, the FNE-Rural subprograms, and global fund assessments.

The Brazilian private sector has performed an important role in strengthening this agenda. Initiatives, like the Competitive Brazil Movement (*Movimento Brasil Competitivo*, MBC), established in November 2001, are mainly focused on enabling projects aimed at increasing the organizations' competitiveness and the population's living conditions, including the public sector management model. This initiative has always promoted the adoption of results-based management practices and already has served many Brazilian ministries, states, and municipalities.

By contracting studies, organizing seminars, and requesting information, many civil society organizations and movements, like the Itaú Social Foundation, Unibanco Institute, and the Futura Channel, have worked in the M&E of programs.

Finally, as an example of initiatives developed by international organizations, the Brazil Evaluation (*Brasil Avaliação*, BRAVA) program, developed by the World Bank as a technical assistance project, not bound to loan operations, provided the federal government assistance in the development of M&E systems and activities. The BRAVA was launched in 2005 and concluded in 2009. The main partner of this initiative was the Technical Monitoring and Evaluation Chamber (*Câmara Técnica de Monitoramento e Avaliação*, CTMA), a committee made up of the main federal government bodies interested in the matter. The BRAVA's objectives were as follows:

- Assist the federal government and its ministries in the development of M&E methodologies and policies.
- Disseminate the M&E culture.
- Prepare results-oriented management systems (GPR) to—
 - clarify the institutional rules to the GPR coordination and implementation;

- establish the main incentives to enable the ministries to carry out the M&E in their programs;
- develop M&E-oriented quality standards; and
- promote the use and dissemination of the information generated by the M&E.⁷⁹

The Brazilian M&E network

Brazil's redemocratization process has increasingly enhanced citizens' roles, notably, as participants in the process of designing, monitoring, evaluating, and controlling public policy implementation. Participation in the decision-making process and the control currently exercised by citizens regarding the governors' and the elected representatives' actions—in forums and councils established specifically to that purpose—have increased the public policies' legitimacy. Moreover, the new ways that the private sector participates in the implementation of policies, like the public-private partnerships and the public-private joint ventures, enhance the need for stronger M&E systems.

As the previous sections suggests, Brazil has broad knowledge and practice in M&E of programs and projects. Most of this experience is stored in research and education centers, or in the consultant's personal collection. Hence, existing knowledge about programs and policies evaluation is spread out and not systematized. Therefore, access to this knowledge depends on the stakeholders' research on the topic, and on the dissemination capacity of those producing the given knowledge. This situation has given rise to the idea among the experts in this area of promoting an initiative that could bring together the institutions that have accumulated experiences in this field and provide forums to collaborate and share information.

The organizations' networks seek an environment to promote debate and strengthen collaboration among similar interests through the exchange of information. In addition, they seek to promote congresses, legal support, and development and provision of databases related to the participating institutions' interests. The organization of information and records, and the socialization of the information by its constituent elements can create a system that enables innovation and interaction with the external environment. The dissemination of organized and structured information to the external environment enhances visibil-

ity while promoting new elements to reconfigure and improve the current system (Lopes 1996).

Knowledge management is clearly an important factor for the network to pursue new linkages, expand the possibility of working in a dynamic context, and ensure maintenance and survival as in the creation of internal control mechanisms capable of promoting better use of existing resources.

The Brazilian M&E network creation process

In this section, we explain the building process, objective, institutional arrangement, technological tool, and results achieved in the first year of the Brazilian Monitoring and Evaluation Network.

Objectives and strategies

Some of the founding principles of the Brazilian chapter of the Latin America and the Caribbean Monitoring and Evaluation Network were decentralization, openness, and the low cost of failure. In its first year, the Brazilian chapter had more than one institutional champion. A number of institutions, including the State Statistics and Economics Foundations, public banks, and international organizations supported the initiative. Generally speaking, the representatives of these institutions in the network, in addition to serving as focal points, also represented individuals committed to these issues in their respective institutions.

A second important principle for the network is its openness. In an effort to maximize the possibilities of cooperation among different stakeholders, the network offers membership to individuals or institutions and to other networks, which translate into the membership's freedom (either personal, institutional, or of other networks) to maximize the possibilities of cooperation among different stakeholders. Finally, a third and important principle was the effort to reduce losses to a minimum if the initiative failed. The initiative's existence is justified only if it is useful and capable of showing its value, which has happened through the growing membership, increased number of visitors, and growing engagement.

The Brazilian Monitoring and Evaluation Network was launched as a chapter of the Latin America and the Caribbean Monitoring and Evaluation Network, under the leadership of the JPF, the Inter-American Development Bank (IDB), and the World Bank. Its launch was a recognition

of the existence of a considerable critical mass for this topic, the lack of forums to exchange and organize experiences, and the mobilizing effect of the Fourth International Conference of the Regional Network, which took place in 2008 and was hosted in Belo Horizonte. It used the gathering in Brazil of different actors in the topic to mobilize a side event during the conference, which led to the foundation of the Brazilian chapter of the Latin America and Caribbean Network.

In addition to the JPF, the IDB, and the World Bank, the inaugural meeting was attended by various representatives of international and national institutions interested in the topic. These institutions included the United Nations Development Programme (UNDP), SEADE Foundation, the *Banco do Nordeste to Brasil*, Itaú Social Foundation, Unibanco Institute, MDS, TCU, and Ceará State Planning Bureau.

The discussions' findings assisted in outlining the core network objectives, such as:

Promote the interchange of experiences and knowledge through the organization of an annual seminar and the social network platform, and disseminate M&E-related information through the publicizing of courses and events, with the mission of strengthening the culture of monitoring and evaluation in Brazil.

It targets the public and private actors working in the field of M&E of organizations, policies, programs, and projects, including managers, researchers, consultants, professors, students, academics, as well as in the field of financing, promoting, and assessing organizations.

Institutional arrangement

In June 2009, during the First Network National Seminar carried out in Belo Horizonte, a second meeting was held and was attended by representatives of the World Bank; JPF; the Sao Paulo state government (SEADE Foundation), Ceará (SEPLAG), Rio Grande do Sul (Planning and Management Bureau, Revenue Bureau and Economics and Statistics Foundation), Bahia (SEPLAG), Espírito Santo; BNB; TCU; UNDP; Ministry of Agrarian Development (Land Development Secretariat); MDS (Information Evaluation and Management Bureau); Ministry of Education

(Continuing Learning, Literacy and Diversity Bureau); Itaú Social Foundation; Ecofuturo Institute and Unibanco Institute; Brazilian Evaluation Agency; Herkenhoff Prates–Minas Gerais Technology and Development Institute; Ayrton Senna Institute; Federal University of Viçosa; Cecaps; and A&M Consultancy Services. Some decisions regarding the institutional arrangements were made during the meeting. These arrangements included (1) the creation of a secretariat made up of representatives from the JPF, the World Bank, BNB, SEADE Foundation, and the Brazilian Evaluation Agency; and (2) the definition of the member types, including institutional and individual members, as well as other networks. The participants reaffirmed the need to publish the contents produced by the network, and other relevant information on a Web site.

National seminar

One of the network's initial objectives was to organize an annual event at which the stakeholders could describe their projects, exchange experiences, and identify best practices in the field of M&E of public policies. The network's first meeting was organized around different kinds of assessment in the public policy cycle: diagnosis for public policy design, the logical framework, ex ante evaluation, administrative records, and ex post evaluation. Another principle employed to organize the sections was the emphasis on how these tools should be implemented and used, rather than on the results from these processes. This organizing principle should be maintained for the network's second meeting, with the addition of some topics that have aroused interest among public managers, like the selection of indicators and goals to design and the management of results agreements. The current structure to organize the network seminars has two blocks:

Block 1: design and institutional strengthening of M&E systems for public policies and programs

This block is mainly targeted at debating the current perspectives of incorporating M&E practices into the public policies process. These perspectives include (1) the creation of legal and managerial mechanisms, (2) the implementation of institutional units or public organizations devoted to monitoring and evaluating the policies, and (3) the appropriation and use of the M&E system's results to feedback the policies and programs.

Block 2: M&E methodologies and practices

This block focuses on (1) evaluation plans, (2) the logical framework, (3) selection of outcome indicators and goals of public management, (4) implementation and process evaluations, (5) impact evaluations, (6) ex ante evaluations, and (7) design of a management information system.

Web 2.0 platform

Another important tool that contributes to the knowledge management of this community of practice was the design of a collaborative platform using Web 2.0 principles. The selected platform resulted in the Brazilian M&E Web site (<http://redebrasileirademea.ning.com>) created soon after the second meeting in June 2009 (see Figure 33). The main functionalities of this platform are as follows:

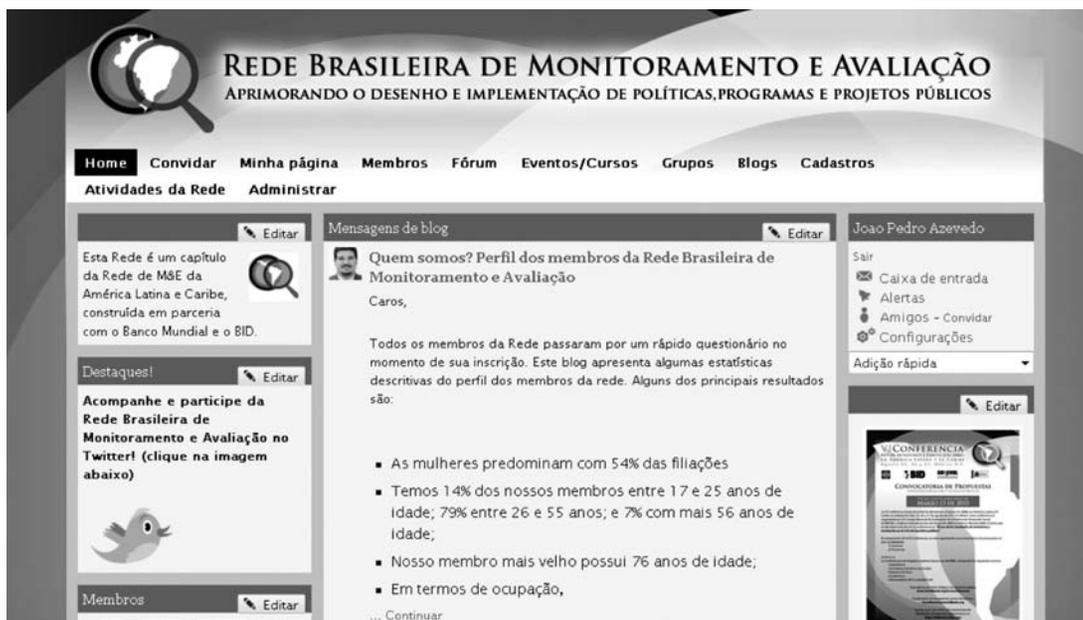
- Publication of news on a blog
- Creation of discussion forums
- Creation of thematic groups
- Publication of an events calendar
- Creation of individual profiles of the members

A crucial aspect in initiatives of this nature is strategic planning for the launch, dissemination, and communication. To be better prepared, one should pay attention to the following basic aspects:

- Before launching the Web page, provide relevant information to the target population (no one likes to become a member of a ghost city).
- Launch the Web site with a small number of functionalities (blog, videos, and events) and add other functionalities as the traffic increases (forums, groups).
- Take advantage of on-site events to disseminate your network.⁸⁰

When it was first launched, the platform was relatively closed. Few members were invited to administer the page. All members had to be approved, and the Web site content was available only to the members. The network gradually became flexible and open by removing many of these restrictions, which, in practice, proved to be extremely bureaucratic. Moreover, an effort was made to promote more members to administrator. The network now has eight ad-

FIGURE 33:

Brazilian M&E Web Site

Source: <http://redebrasileirademea.ning.com/>.

ministrators. The most recent change was the full and open access of the network's content. Membership is required only to post messages and events, and to participate in the thematic groups and discussion forums.

Outcomes

A strategic aspect to consolidating the networks is to monitor results by constructing and updating the members' profiles, and monitoring and analyzing traffic to the Web page.

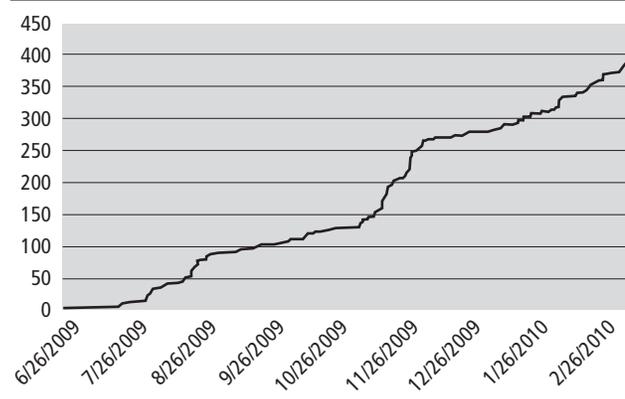
Members

As of March 2010, the Brazilian M&E Network has more than 400 members from 18 states in Brazil and 14 countries. Figure 34 shows the development of the network membership in the last eight months. The discontinuities observed are clearly associated to specific events like (1) the first network meeting in Belo Horizonte (June 2009), (2) the second seminar on M&E in Rio Grande do Sul (November 2009), (3) the ANIPS national Meeting (November 2009), and (4) the CONSAD 2010 meeting (March 2010). The network members who were directly or indirectly involved in the event's organization disseminated information about the network. The resulting increased membership is remarkable.

Network members are mainly professionals from public and private institutions that are interested in M&E activities. Following are some of the main features of the network's membership:⁸¹

- Women account for 54 percent of membership.
- 14 percent of members are between 17 and 25 years of age; 79 percent between 26 and 55 years of age; and 7 percent are older than 56.

FIGURE 34:

Growth of Membership in the Brazilian M&E Network

Source: Author.

- Our oldest member is 76 years old.
- 58 percent are civil servants, 18 percent are private employees, 16 percent are consultants, and 8 percent are students.
- 39 percent work for state governments; 10 percent for municipal governments; 14 percent for the federal government; 19 percent for nongovernmental organizations (NGOs); 10 percent for private institutions; and the remaining 8 percent for education and international institutions.
- 48 percent of the members heard about the network from a colleague; 25 percent learned about it in some event where information about the network was disseminated.

Monitoring visitors to the Web page

Another important aspect of the network is monitoring and analyzing its use. Currently, many free tools facilitate this purpose. These tools allow for the real-time follow-up on the network performance, collecting different use statistics, including the number of hits, countries and cities, and Web site of origin. Moreover, tracking visitors also enables building dashboards (see Figure 35) that compare the network traffic performance over time, identify Web sites of the same nature, and automatically send reports to the members. This kind of monitoring allows the network sec-

retariat to stay informed about the site's performance and to improve the initiative's mobilization activities and planning.

User engagement

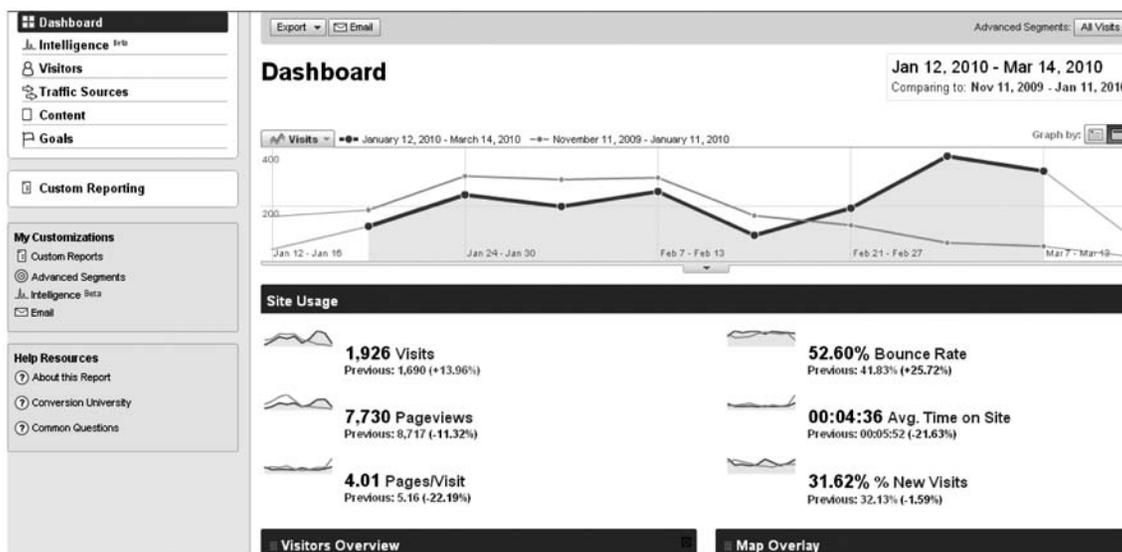
Another interesting metric has to do with the level of content publication in the platform. Until recently, network members have posted more than 100 blogs, about 60 events, and 13 bibliographic references. Moreover, the number of comments among participants has also increased, suggesting deeper engagement of the members.

Final remarks and recommendations

The Brazilian Monitoring and Evaluation Network stands for a partnership of different organizations with different natures. Among some of the success indicators, the capacity of attracting and retaining members is outstanding. However, ensuring sustainability is one of the network's main challenges. This sustainability will result from the network's capacity to make the contents relevant to its members, rotate leadership in the secretariat, and attract new facilitators and moderators to the network, notably to the thematic groups. Another important challenge is the geographic distribution of the members and content presented, making the network space useful to members and institutions in different stages of experience. Finally, another concern is to maintain and expand the network's openness and avoid creating a space

FIGURE 35:

Dashboard of the Brazilian M&E Network



Source: Author.

captured by a small group of interests. Some of the Web platform's important challenges are as following:

- The proper use of “virtual spaces”
 - Many members have found it difficult to understand the difference among blogs, forums, thematic groups, and fields of comments. Thus, they ignore the conceptual differences among these, and post their contents in a disorganized way.
- Lack of reference links
 - Another difficulty is the posting of content (blogs and comments) with no reference links on the topic.
- The visual presentation of blogs
 - Many members have difficulty with the visual presentation of their blogs and ignore many editing functions provided by the tool. The page layout becomes confused.

Finally, a network needs (1) a critical mass of individuals, (2) a concerted effort to attract and sustain a plurality of institutions to protect the network from the political cycle, and (3) a group of champions to facilitate and moderate the content of the collaborative platform. The champions' duties should include the following:

- Attract new members
- Get to know and analyze the members' profiles (individuals and institutions)
- Input relevant contents to the Web site, meeting the members' expectations
- Promote communication among members
- Prepare regular reports of the activities
- Connect to similar activities

Afonso Henriques

President, João Pinheiro Foundation

Jânia Pinho

Manager to the Studies, Surveys and Assessment Environment of the Banco do Nordeste do Brasil

João Pedro Azevedo

Economist, The World Bank

John L. Newman

Lead Economist, The World Bank

Juliana Wenceslau

Graduate student in Public Management, Administration Post-Graduation Program (PPGA), University of Brasilia (UnB)

Márcia Paterno Joppert

General Director, Brazilian Evaluation Agency

Sonia Nahas de Carvalho

Technical Advisor to the Deputy Board of Information Analysis and Dissemination, SEADE Foundation

Session 5 Notes

⁶⁴ This presentation is adapted from Bamberger (2009) and draws extensively on MacKay (2007).

⁶⁵ For more information on the Africa Impact Evaluation Initiative, see the Web site, <http://worldbank.org/afi/impact>.

⁶⁶ See the Web site, <http://www.worldbank.org/ieg/nonie/index.html>.

⁶⁷ See the Web site, <http://www.3ieimpact.org>.

⁶⁸ This paper is based on the author's experience as consultant for the Office of Public Policy Evaluation of the National Planning Department, which runs the Colombian evaluation and monitoring system SINERGIA.

⁶⁹ For more information on the other RDPPS in other parts of the country, see <http://www.redprodepaz.org.co> and <http://www.accionsocial.gov.co/portal/default.aspx>.

⁷⁰ For results from the first phase of the evaluation see DNP (2009), available at http://www.dnp.gov.co/PortalWeb/Portals/0/archivos/documentos/DEPP/Evaluacion_Politiclas_Publicas/Serie_EPP11_Laboratorios_de_paz.pdf.

⁷¹ The initial design of the evaluation and the collection of baseline information were carried out by Unión Temporal Econometría-SEI-IFS, under the supervision of the NPD. Baseline information was used to produce a descriptive analysis of the Program beneficiaries and certain preliminary impacts were estimated. The results have been recorded in DNP (2009, *ibid.*)

⁷² For an in-depth analysis, see DNP (2009, *ibid.*)

⁷³ For an in-depth analysis, see DNP (2009) and Attanasio, Pellerano, and Phillips (2009).

⁷⁴ The survey of a sample of 2,288 people took place in September 2009. The response rate was 48.7 percent and covered 24 countries of the region.

⁷⁵ Argentina, Bahamas, Barbados, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Suriname, Trinidad and Tobago, and Uruguay.

⁷⁶ The authors thank the other network members for their comments and suggestions and participation in the V International Conference of the Latin American Monitoring and Evaluation Network, notably Eder Campos and Lycia Lima. The

authors also thank the initiative by Ricardo Santiago and the support of Ricardo Paes de Barros, Eduardo Rios-Neto, André Urani, Wanda Engles, and Pedro Olinto to the launching and development of the first seminar of the Brazilian Monitoring and Evaluation Network.

⁷⁷ See Web site, <http://www.seade.gov.br/>.

⁷⁸ See Web site, <http://www.bnb.gov.br/>.

⁷⁹ In its five years of existence, the BRAVA delivered assistance, carried out studies, and promoted events with the federal government and state and municipal governments. These have included (1) workshops on assessment systems from Australia, Canada, and Chile; (2) outlining the methodology to re-

view the logical framework of the Multiyear Plan programs; (3) institutional assessment of the PPA and MDS assessment system; (4) participation in the outlining of program assessments (like Bolsa Família, Brasil Alfabetizado, Crediamigo, Quilombolas, and Saúde Indígena); (5) training of the TCU auditors in impact assessment through an IDF grant; and (6) support to the state of Minas Gerais in the outlining of the first household survey (Pequisa por Amostra de Domicílios, PAD/MG) methodology of impact assessments, and indicators benchmarking.

⁸⁰ See the Web site, http://help.ning.com/cgi-bin/ning.cfg/php/en-duser/workshop.php?p_sid=LDXLDYWj.

⁸¹ The Network's Web page provides a detailed profile of the network members.

References and Additional Bibliography

- Anastasia, Antonio Augusto J. 2006. "Introduction." In *O choque de gestão em Minas Gerais: políticas da gestão pública para o desenvolvimento*, Renata Vilhena, Humberto Falcão Martins, Caio Marini, and Tadeu Barreto Guimarães (organizers), 13–17. Belo Horizonte: Editora UFMG.
- Arnold, Mary. 2006. "Developing Evaluation Capacity in Extension of 4-H Field Faculty." *American Journal of Evaluation* 27 (2): 257–69.
- ADB (Asian Development Bank). 2004. "Evaluation of Capacity Development Assistance to Lao PDR." Available at <http://www.adb.org/>.
- . 2003. "Evaluation of Capacity Development for Managing Water Utilities." Available at <http://www.adb.org/>.
- . 2002. "Evaluation of Country Assistance Program Evaluation: Mongolia." Available at <http://www.adb.org/>.
- . 1999. "Evaluation of Technical Assistance for Public Expenditure Management to Bhutan, India, Kiribati and Lao PDR." Available at <http://www.adb.org/>.
- . 1997. "Evaluation of Effectiveness of Training in Education Projects in Indonesia." Available at <http://www.adb.org/>.
- Attanasio, O. Pellerano, L. and Phillips. 2009. "Peace and Goodwill. Using an Experimental Game to Analyse Paz y Desarrollo." Working Paper W09/20, Institute for Fiscal Studies, London.
- AusAID (Australian Government Overseas Aid Program). 2009. "A Staged Approach to Assess, Plan and Monitor Capacity Building." Available at <http://www.ausaid.gov.au/>.
- Baldrige National Quality Program. 2009. *Education Criteria for Performance Excellence*. Gaithersburg, MD: National Institute of Standards and Technology (NIST), U.S. Department of Commerce. Available at http://www.baldrige.nist.gov/Education_Criteria.htm.
- Bamberger, Michael. 2009. "Institutionalizing Impact Evaluation Within the Framework of a Monitoring and Evaluation System." Independent Evaluation Group, World Bank, Washington, DC. Available at http://intresources.worldbank.org/INTINFRES/Resources/2780429-1173477721799/2009_institutionalizing_ie_framework_me.pdf.
- . 2006. "Evaluation Capacity Building." In *Creating and Developing Evaluation Organizations: Lessons learned from Africa, Americas, Asia, Australasia and Europe*, ed. M. Segone. Geneva: United Nations Children Emergency Fund (UNICEF).
- Bamberger, M., K. Mackay, and E. Ooi. 2005. *Influential Evaluations: Detailed Case Studies*. Operations Evaluation Department. Washington, DC: World Bank
- . 2004. *Influential Evaluations: Evaluations that Improved Performance and Impacts of Development Programs*. Operations Evaluation Department. Washington, DC: World Bank
- Barkenbus, Jack. 1998. *Expertise and the Policy Cycle*. University of Tennessee. Available at <http://www.gdrc.org/decision/policy-cycle.pdf>.
- BDMG (Banco de Desenvolvimento de Minas Gerais). 2002. *Minas Gerais do Século XXI—transformando o*

- poder público: a busca da eficiência*. Vol. 9. Belo Horizonte: Rona Editora.
- Beck, Ulrich. 1998. *Risk-fraught Society. Towards a new modernity*. Barcelona: Editorial Paidós.
- . 2001. "Return to the Theory of the Risk-fraught Society." *Bulletin of the EGM* [[Vol.?:]] no. 30 (March).
- Boyle, Richard, and Lemaire, Donald, eds. 1999. *Building Effective Evaluation Capacity, Lessons from Practice*. New Brunswick, NJ. Transactions Publishers.
- Budget Office (Chile). 2009a. "Public Finance Report." Government Draft Budget Law for 2010, 119–61, October.
- Budget Office (Chile). 2009b. "Evaluation of Public Sector Financial Management in 2008 and Updated Projections for 2009," 25–38 and 65–75, June.
- Bureau, Sylvie. 2004. *Analyse des besoins de formation*. Développement International Desjardins, Québec. Available at http://www.lamicrofinance.org/files/17682_file_Document.pdf.
- CIDA (Canadian International Development Agency). 2000. "Capacity Development: Why, What and How." Available at <http://www.acdi-cida.gc.ca/home>.
- Capacity.org. 2005. "Exploring the Soft Side of Capacity Development." Available at <http://www.capacity.org/en>.
- . 2004a. "Meso-level Capacity Development." Available at <http://www.capacity.org/en>.
- . 2004b. "Renewing Approaches to Institutional Development." Available at <http://www.capacity.org/en>.
- . 2003. "Capacity Development: The Why's and How's." Available at <http://www.capacity.org/en>.
- . 2002. "Capacity for Development: Insights and Innovation." Available at <http://www.capacity.org/en>.
- . 2001. "Tools of the Trade: Capacity Assessment." Available at <http://www.capacity.org/en>.
- . 2000. "Informing the Capacity Debate: Operational Experiences." Available at <http://www.capacity.org/en>.
- Chelimsky, Eleanor. 2008. "A Clash of Cultures: Improving the 'Fit' between Evaluative Independence and the Political Requirements of a Democratic Society." *American Journal of Evaluation* 29 (4): 400–15.
- Corporación Chile Transparente. 2006. "About the Spanish Concept of Corruption." Internal Document, Transparency International, Santiago.
- Corporación Transparenciapor Colombia. 2009. *Índice de Transparencia Nacional Resultados 2007 y 2008*. Colección Documentos, Observatorio de Integridad No. 10, Bogotá, July.
- . 2008a. *Índice de Transparencia Municipal 2005–2006*. Colección Documentos, Observatorio de Integridad No. 9, Bogotá, May.
- . 2008b. *Índice de Transparencia Departamental 2005–2006*. Colección Documentos, Observatorio de Integridad No. 8, Bogotá, May.
- . 2006. *Índice de Transparencia Nacional – Entidades Públicas Nacionales Resultados 2004 y Resultados 2005*. Colección Documentos, Observatorio de Integridad No. 7, Bogotá, December.
- . 2006. *Índice de Transparencia Municipal Resultados 2004 – Ejerciciopiloto*. Colección Documentos, Observatorio de Integridad No. 6, Bogotá, December.
- . 2005a. *Índice de Transparencia Departamental-Resultados 2004–2005*. Colección Documentos, Observatorio de Integridad No. 5, Bogotá, December.
- . 2005b. "Transparency as a Job Category." Internal Document. Bogotá.
- Cunill Grau, Nuria. 2009. "El Mercado en el Estado." *Revista Nueva Sociedad* 221 (May–June): 34–53. Available at http://www.nuso.org/upload/articulos/3608_1.pdf.
- Danida (Danish International Development Assistance). 2005. "A Results-Oriented Approach to Capacity Change." Available at <http://www.dacp.ait.ac.th/danida.html>.
- . 2002. "Capacity Development Evaluation Step 1: Contributions to an Analytical Framework." Available at <http://www.dacp.ait.ac.th/danida.html>.
- DAC/UNEG (Development Assistance Committee/United Nations Evaluation Group). 2007. "Framework for Professional Peer Reviews. Joint Task Force on Professional Peer Reviews of Evaluation Functions in Multilateral Organizations." Available at http://www.uneval.org/documentdownload?doc_id=103&file_id=210.
- (DfID) U.K. Department for International Development. 2005. "Capacity Development and State Building: Issues, Evidence and Implications for DfID." Available at <http://www.dfid.gov.uk/>.
- ECDPM (European Centre for Development Policy Management). 2009. "Capacity Change and Performance: Capacity Development: Between Planned Interven-

- tions and Emergent Processes Implications for Development Cooperation.” Policy Management Brief No. 22, ECDPM, Maastricht. Available at [http://www.ecdpm.org/Web_ECDPM/Web/Content/Download.nsf/0/5E619EA3431DE022C12575990029E824/\\$FILE/PMB22_e_CDapproaches-capacitystudy.pdf](http://www.ecdpm.org/Web_ECDPM/Web/Content/Download.nsf/0/5E619EA3431DE022C12575990029E824/$FILE/PMB22_e_CDapproaches-capacitystudy.pdf).
- . 2008. “Capacity Change and Performance: Insights and implications for Development Cooperation” Policy Management Brief No. 21, ECDPM, Maastricht. Available at www.ecdpm.org/capacitystudy.
- . 2006a. “Study on Capacity, Change, and Performance,” Final Workshop Report, ECDPM, Maastricht. May. Available at [http://www.ecdpm.org/Web_ECDPM/Web/Content/Download.nsf/0/B03C8B92BAE9E5A1C12571C0003C144E/\\$FILE/Final%20Workshop%20Capacity%20Study_May%202006%20-%20Report.pdf](http://www.ecdpm.org/Web_ECDPM/Web/Content/Download.nsf/0/B03C8B92BAE9E5A1C12571C0003C144E/$FILE/Final%20Workshop%20Capacity%20Study_May%202006%20-%20Report.pdf).
- . 2006b. “The Concept of Capacity.” Draft Report, ECDPM, Maastricht. Available at [http://www.ecdpm.org/Web_ECDPM/Web/Content/Download.nsf/0/EF596EFCBB8A417DC1257178004AEB5B/\\$FILE/Morgan%20-%20Capacity%20-%20What%20is%20it%2010052006.pdf](http://www.ecdpm.org/Web_ECDPM/Web/Content/Download.nsf/0/EF596EFCBB8A417DC1257178004AEB5B/$FILE/Morgan%20-%20Capacity%20-%20What%20is%20it%2010052006.pdf).
- Ferraz, Claudio, and Frederico Finan. 2008. “Exposing Corrupt Politicians: The Effects of Brazil’s Publicly Released Audits on Electoral Outcomes.” *The Quarterly Journal of Economics* (MIT Press) 123 (2): 703–45.
- Flint, M. 2004. *Country Study: Brazil 1997–2003. Evaluation of DFID Country Programmes*, London: United Kingdom Department for International Development.
- Garay, Luis Jorge, et al. 2007. “Armed Conflict and Organized Crime in Corrupt Scenarios. About the Case of Colombia.” Collection of Notebooks on Transparency No. 13, Corporación Transparencia por Colombia, Bogotá.
- . 2002. “On the Role and Nature of Corruption in Colombia.” Paper presented at *Forum on Fighting Corruption, Governance and Institutional Performance*, March 21, Bogotá.
- García Lopez, Roberto. 2008. *La Gestión para Resultados y el Presupuesto para Resultados en América Latina y el Caribe*. Paper presented at the *Conferencia internacional de presupuesto por resultados*, June 8–9, México
- . 2009. *El rol del Congreso en la Gestión para Resultados del Desarrollo*. Parliamentarian Forum on MfDR, Inter-American Development Bank, Washington, DC.
- . 2009. *La sociedad civil y la GpRD, en búsqueda de una alianza estratégica*. Inter-American Development Bank, Washington, DC.
- Guimarães, Tadeu Barreto, and Bernardo Tavares de Almeida. 2006. *Da estratégia aos resultados concretos: a experiência de Minas Gerais (2003–2005)*, n. 12. CADERNOS BDMG: Belo Horizonte: BDMG.
- Guimarães, Tadeu Barreto, and Eder Sá Alves Campos. 2009. *Programa Estado para Resultados: balanço da implementação de um escritório de apoio à gestão estratégica e gestão por resultados em Minas Gerais*. II CONSAD Congress, Brasília.
- Guzmán, Marcela. 2007. “Program Evaluation: Technical Notes.” Public Management Series, ECLAC, United Nations, New York.
- . 2005. “Systems of Management and Control and Results-based Budgeting, The Chilean Experience.” Available at www.dipres.cl/fr_control.html.
- . 2003. *Systemsof Management Control and Results-Based Budgeting: The Chilean Experience*. Santiago: Ministry of Finance, Government of Chile.
- IEG (Independent Evaluation Group). 2007. *Using Training to Build Capacity for Development: An Evaluation of the World Bank’s Project-based and WBI Training*. Washington, DC: World Bank.
- ILPES-CEPAL (Instituto Latinoamericano y del Caribe de Planificación Económica y Social–Comisión Económica de América Latina). 2003. *Government Management by Results*. New York City
- IDB-CLAD (Inter-American Development Bank–Latin American Center for Development). 2006. *Modelo Abierto de Gestión para Resultados en el Sector Público*. Inter American Development Bank, Washington, DC.
- Kaplan, R. 1999. “The Balanced Scorecard for Public Organizations.” *Harvard Business Review*, Article Reprint No. B9911C, Cambridge, MA.
- Kusek, J., and R. Rist 2004. *Ten Steps to Results-Based Monitoring and Evaluation Systems*. Washington, DC: World Bank.
- Larbi, George A. 1999. *The New Public Management Approach and Crisis States*. Geneva: United Nations Research Institute for Social Development (UNRISD).

- Lopes, Sônia Aguiar. 1996. *A teia invisível: informação e contra-informação nas redes de ONGs e movimentos sociais*. Unpublished doctoral dissertation. Federal University of Rio de Janeiro, Rio de Janeiro.
- Mackay, K. 2007. "How to Build M&E Systems to Support Better Government." Independent Evaluation Group, World Bank, Washington, DC.
- Marcel, Mario, and Marcela Guzmán. 2008. Hacia un círculo virtuoso en la reforma presupuestaria en Chile. *Presupuesto y Gasto Público* 51: 305–23. Available at http://www.ief.es/default_Ing.htm.
- Marcel, Mario, Marcela Guzmán, and Marcelo Tokman. 2002. "Building a Consensus for Fiscal Reform: The Chilean Case." *OECD Journal on Budgeting* 2 (3): 35–55.
- Metzenbaum, Shelley H. 2009. "Performance Management Recommendations for the New Administration." IBM Center for the Business of Government, Minas Gerais Planning and Management Secretariat. 2007. *Plano Mineiro de Desenvolvimento Integrado 2007–2023: Estratégia de Desenvolvimento Estado para Resultados*. Belo Horizonte. Available at <http://www.almg.gov.br/ventos/pmdi/pmdi.pdf>.
- ODI (Overseas Development Institute). 2006. "Capacity Development for Policy Advocacy: Current Thinking and Approaches among Agencies Supporting Civil Society Organisations." Available at <http://www.odi.org.uk/resources/download/136.pdf>.
- OECD/DAC (Organisation for Economic Co-operation and Development/Development Assistance Committee). 2008. Third High Level Forum on Aid Effectiveness. *Accra Agenda for Action*. Available at <http://www.accrahl.net>.
- . 2006. "The Challenge Of Capacity Development, Working towards Good Practice." Available at <http://www.oecd.org>.
- . 2002. *Glossary of Key Terms in Evaluation and Results-Based Management of the Organisation for Economic Co-operation and Development*. Paris: OECD.
- . 1992. "DAC Principles for Effective Aid." Reprint available at <http://www.oecd.org/dataoecd/31/12/2755284.pdf>.
- Oliveira, Ana Clara de, Bernardo de Almeida, and Tadeu Barreto Guimarães. 2006. *A Integração Planejamento-Orçamento*. In *O Choque de Gestão em Minas Gerais: Políticas da Gestão Pública para o Desenvolvimento*, Renata Vilhena, Humberto Falcão Martins, Caio Marini, and Tadeu Barreto Guimarães (organizers), 67–76. Belo Horizonte: Editora UFMG.
- Osborne, David, and Ted Gaebler. 1992. *Reinventing Government*. Reading, MA: Addison-Wesley Publishing.
- Pitman, George K., and Osvaldo Feinstein, and Gregory Ingram, eds. 2005. *Evaluating Development Effectiveness*. Washington, DC: World Bank.
- Porto, Claudio, and José Paulo Silveira. 2006. *Implantação e Gestão da Estratégia do Governo*, n. 12. Cadernos BDMG: Belo Horizonte: BDMG.
- Rose-Ackerman, Susan. 2001. *Corruption and Governments. Causes, Consequences and Reform*. Madrid: Siglo Veintiuno Editores.
- SEPLAG/MG (Minas Gerais State Planning and Management Secretariat). 2009. *Caderno de Indicadores. Programa Estado para Resultados*. Belo Horizonte, Minas Gerais.
- . 2008. NOTA TÉCNICA DCMG/SEPLAG N° 37 /2008. Belo Horizonte: Minas Geras.
- Sánchez Albarena, Fernando. 2003. *Planificación estratégica y gestión pública por objetivos*. Santiago: ILPES.
- Segone, Marcos, and Ada Ocampo. 2006. *Creating and Developing Evaluation Organizations*. Lima: DESCO.
- Shack Yalta, Nelson. 2003. "Progress in Implementing Performance Indicators in Public Entities in Peru." In *Government Management by Results*. New York: ILPES-CEPAL.
- Shirky, Clay. 2008. *Here Comes Everybody: The Power of Organizing without Organizations*. New York: Penguin Press.
- SPI/MPOG (Ministry of Planning, Budgeting and Management). 2010. *Indicadores de programas: Guia Metodológico*. Brasília, MPOG: 128.
- Torres Fragoso, Jaime. 2008. "La influencia de las culturas nacionales en el desarrollo de la Nueva Gerencia Pública en América Latina." In *Revista Reforma y Democracia*, no. 42. Caracas: CLAD.
- UNDP (United Nations Development Program). 2007. "Capacity Development Practice Note." Available at http://www.undp.org/cpr/iasc/content/docs/UNDP_Capacity_Development.pdf.
- . 2007. *Capacity Assessment Methodology*. User's Guide. New York: UNDP. <http://europeandcis.undp.org/>

- uploads/public/File/Capacity_Development_Regional_Training/UNDP_Capacity_Assessment_Users_Guide_MAY_2007.pdf.*
- UNEG (United Nations Evaluation Group). 2007. "Ethical Guidelines for Evaluation." Available at: <http://www.uneval.org/search/index.jsp?q=ethical+guidelines>.
- . 2008. "Code of Conduct for Evaluation in the UN System." Available at <http://www.uneval.org/search/index.jsp?q=code+of+conduct>.
- . 2005a. "Assessment of 20 Capacity Development Tool." Available at <http://www.uneval.org/search/index.jsp?q=Assessment+of+20+Capacity+Development+Tools>.
- . 2005b. "Measuring Capacities: An Illustrative Catalogue to Benchmarks and Indicators." Available at <http://www.uneval.org/search/index.jsp?q=Measuring+Capacities%3A+An+Illustrative+Catalogue+to+Benchmarks+and+Indicators>.
- UNICEF (United Nations Children Emergency Fund). 2002. "Capacity Building for UNICEF Humanitarian Response in Review." Available at <http://www.unicef.org/>.
- UNIDO (United Nations Industrial Development Organization). 1990. "Report on the Evaluation of UNDP-Financed and UNIDO-Executed Projects of Technical Assistance in Metrology." New York: UNIDO.
- United Nations World Food Programme. 2008a. "Evaluation of WFP's Capacity Development Policy and Operations." Available at <http://www.wfp.org/>.
- . 2008b. "Evaluation Policy." Available at <http://www.wfp.org/>.
- Vaitsman, Jeni R., Roberto W. S. Rodrigues, and Rômulo Paes-Sousa. 2006. "The System for Evaluating and Monitoring Social Development Programs and Policies: the case of the Ministry of Social Development and the Fight against Hunger in Brazil." Policy Papers 17, UNESCO.
- Vilhena, Renata. 2006. *Rumo à segunda geração do choque de gestão*. In: Marini, Caio *et al.* *O choque de gestão em Minas Gerais: políticas da gestão pública para o desenvolvimento*, Renata Vilhena, Humberto Falcão Martins, Caio Marini, and Tadeu Barreto Guimarães (organizers), 351–56. Belo Horizonte: Editora UFMG.
- Villarreal, Julio. 2007. "Case Studies: Colombia and the Monitoring and Evaluation System (NMRS)." CLAD, World Bank, Washington, DC.
- WBI (World Bank Institute). 2007. "Capacity Development in the World Bank Group: A Review of Non-lending Approaches." Available at <http://wbi.worldbank.org/wbi/>.
- . 2006. "A Revolution in Capacity Development? Africans Ask Tough Questions." Available at <http://wbi.worldbank.org/wbi/>.
- Wenceslau, J., and J. Hobbs. 2008. "Institutional Assessment of Monitoring and Evaluation Systems (M&E) in the Ministry of Social Development and Combating Hunger (MDS)." World Bank, Washington, DC.
- World Bank. 2006. "Evaluation Capacity Development." Available at <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTCDRC/0,,menuPK:64169181-pagePK:64169192-piPK:64169180-theSitePK:489952,00.html>
- . 2005a. *Impact Evaluation Study under the Programs Evaluation Plan in Chile*. Washington, DC: World Bank.
- . 2005b. "Institutionalizing M&E systems in Latin American and Caribbean Countries." PREM Notes 102. Available at <http://www1.worldbank.org/prem/PREMNotes/premnote102.pdf>.
- . 2005c. "Capacity Building in Africa, an OED Evaluation of World Bank Support." Available at <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTCDRC/0,,menuPK:64169181-pagePK:64169192-piPK:64169180-theSitePK:489952,00.html>.
- . 2004. "Evaluation Capacity Development: OED Self-Evaluation." Operations Evaluation Department, World Bank, Washington, DC.
- . 2002. *Institutional Performance and Governance in Colombia*. Bogotá: World Bank.
- World Bank and OECD (Organisation for Economic Cooperation and Development). 2005. *Recently-Identified Good Management Practices for Development Results*. Washington, DC: World Bank.

Internet Resources

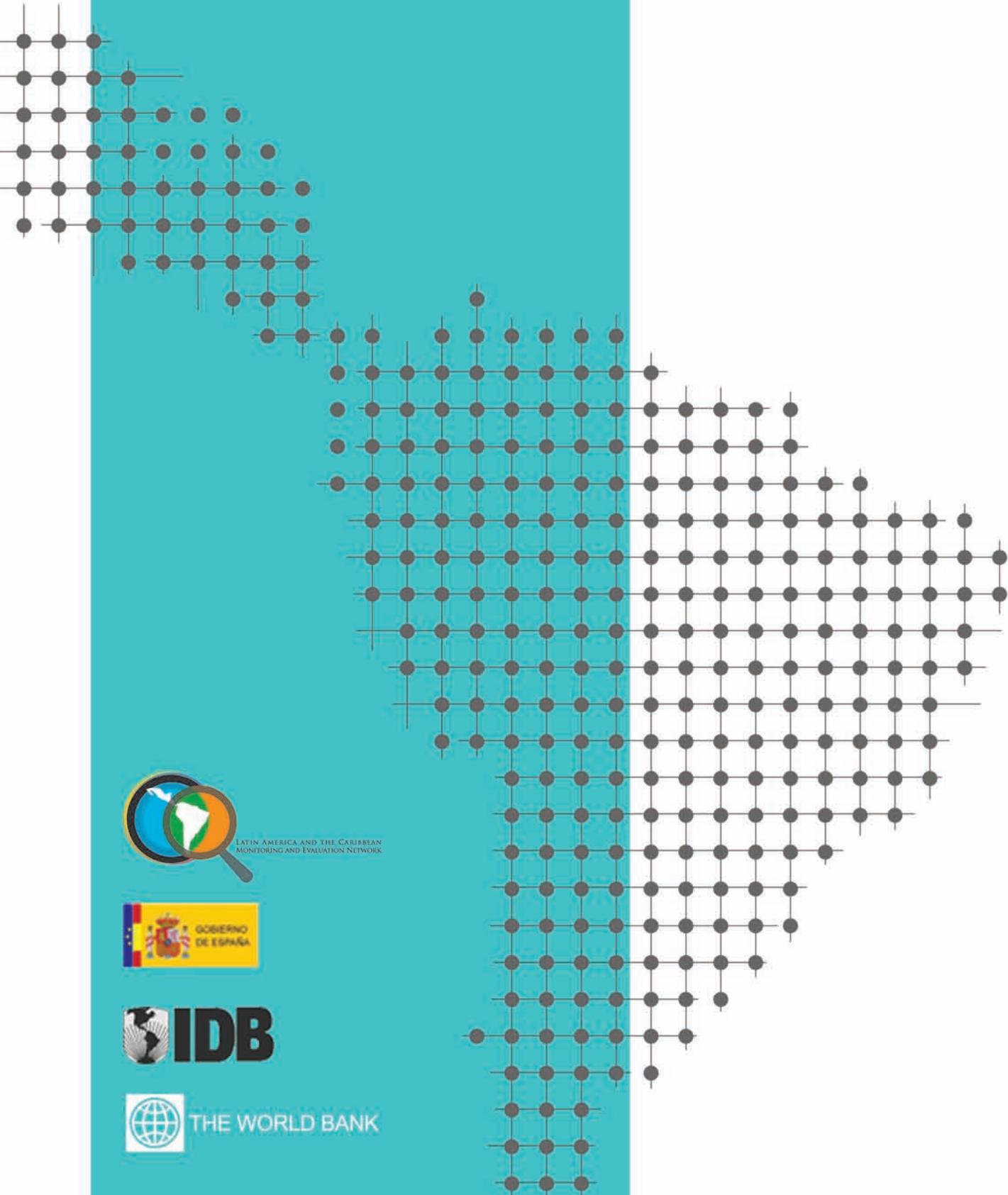
- Monitoring and Evaluation News. <http://www.mande.co.uk>.
- Mouvement Québécois de la Qualité. 2009. *Le Qualimètre*. <http://www.qualite.qc.ca/quali/qualimetre3.htm>.

National Institute of Standards and Technology. 2009. *Bal-ridge National Quality Program*. <http://www.quality.nist.gov/>.

OECD/DAC (Organisation for Economic Co-operation and Development/Development Assistance Committee). 2009. *MfDR Capacity Scan*. <http://www.mfdr.org/CAP-Scan.html>.

World Bank Independent Evaluation Group. <http://www.worldbank.org/ieg> and <http://www.worldbank.org/ieg/eed>.

World Bank Independent Evaluation Group. "Impact Evaluation." <http://www.worldbank.org/ieglie> and <http://www.worldbank.org/impacetevaluation>.



LATIN AMERICA AND THE CARIBBEAN
MONITORING AND EVALUATION NETWORK

