

A stylized world map in a light blue color, showing the continents and major oceans, positioned in the upper right quadrant of the cover.

EVALUATION CAPACITY DEVELOPMENT

ECD WORKING PAPER SERIES NO. 5: JANUARY 1999

Comparative
Insights from
Colombia, China,
and Indonesia

R. Pablo Guerrero O.



Comparative Insights from Colombia, China, and Indonesia

R. Pablo Guerrero O.

Introduction

The reasons that led to the establishment of evaluation as a tool of governance and to evaluation capacity building in industrial countries are also at the root of initiatives in developing countries. The rationale underlying the role of evaluation in public sector reform in developing countries is well articulated by Wiesner (1993, 1997). Wiesner has observed that in many developing countries, including his native Colombia, in spite of sound macroeconomic management the growth rate of the economy has stagnated. A main cause has been the constraining influence of a large and inefficient public sector. He attributes this in great measure to lack of competition in the delivery of public services. Government agencies behave as monopolies, not subject to real market tests, to a large extent unaccountable, and captured by rent-seeking special interests. While in the private sector, competition and prices largely determine whether a business thrives or fails, public sector service delivery by government agencies has not been contested. Wiesner argues that transparent public sector evaluations provide the “market” signals which would effectively show taxpayers and civil society at-large whether the public sector is delivering services efficiently and effectively.

In essence, public sector evaluations of major government programs provide signals which proxy those which make the market function. In discussing the role of evaluation, Havens (1992) has articulated a complementary point: “the more that we [public administrators] know about how our programs are functioning, the effects they are having, and at what cost, the more likely we are to search out ways of making them more efficient and more effective. To a substantial degree, this knowledge is the public-sector manager’s surrogate for the profit-and-loss statement of the business sector.”

The World Bank sees public sector evaluation as a necessary component of the increasing number of public sector reform programs and good governance initiatives it is supporting throughout developing countries. Stiglitz (1989, 1994) provides a more theoretical basis for public sector evaluation. In analyzing the role of the state, Stiglitz has shown the unique role information asymmetry plays in public failure, affecting negatively the public sector role in economic management, and by implication in economic development. Decisions based on imperfect knowledge, lead to suboptimal resource allocation. Arguably, to the extent that public sector evaluations produce valuable and accessible information about the management of public resources, their availability reduces information asymmetry among economic agents and, other things being equal, positively

affects the welfare-enhancing role of the state. Bates (1995) would argue that public sector evaluation is a non-market solution or institution which compensates for market failures. There is ample evidence based, for example, on World Bank-sponsored evaluations of public sector projects and programs, that indeed public sector evaluations can lead to improved public sector resource allocation over time. This type of evidence is at the root of the increasing interest in evaluation capacity development in developing countries today.

This paper concerns the practical aspects of evaluation capacity development in developing countries. It seeks to illustrate how selected developing countries have approached the institutionalization of evaluation in government. The demand and supply framework for evaluation capacity building provides the road map. This paper explores the context and the key issues with respect to emerging experience in three developing countries: Colombia, China, and Indonesia. These are countries where the World Bank has been supporting government evaluation efforts.

Experience in industrial countries shows that each country has unique institutional characteristics which make generalizations difficult, inadvisable, or risky. Derliens (1990) has shown that there are as many forms of government and public sector organization in industrial countries as there are types of

evaluation systems. Yet within existing systems, there are evaluation practices worth examining and perhaps emulating or adapting in developing countries.

It is not coincidence that when developing country governments have decided to move forward with institutionalizing the evaluation function, one of their first priorities has been to learn as much as possible about the development of the evaluation function in industrial countries and how it works today.

For reference, and to better understand the context of the cases discussed, it is important to note that the three subject countries have very different forms of government organization which affect the institutionalization of evaluation capacity. Colombia has a multiparty democratic system of government, with a president, congress, and senate elected by direct universal suffrage. China is governed by a single party system, with a national people's congress elected by citizens, which in turn selects a president and premier. Indonesia is a presidential republic, with a house of representatives and a people's consultative assembly, which includes the members of the house elected by universal suffrage. As will be discussed, each has pursued the institutionalization of a public sector evaluation function in a different way.

For ease of presentation, the discussion of the country cases follows three major streams. The first addresses the demand for evaluation, and covers the issues of leadership, incentives, and the institutional set-up. The second addresses issues of organization of the evaluation function and actionable steps. The third addresses supply issues – including evaluation capacity, that is, staffing and skills, financial resources, and methods. The perspectives afforded by the three cases provide a basis for approaching the institutionalization of evaluation in other countries. Table 1 summarizes where these countries stand with respect to institutionalizing evaluation.

Demand Considerations

Experience shows without effective demand, that is, demand based on real pressures on governments, any effort to institutionalize a public sector evaluation function will quickly lead to lack of interest and evaporation. What are the circumstances

that led these three countries to move forward on evaluation, and what were the first steps? First, all these countries have been facing increasingly hard resource constraints to finance ever more demanding development plans. Population growth and people's expectations about better living standards, fed in part by the worldwide information revolution, have put pressure on governments to deliver more and higher quality public services. Public sector budgets, whether internally or externally financed, have strained to keep pace with public expenditure needs. Government political promises and commitments have additionally fuelled pressures to improve public sector efficiency in resource allocation. Without appropriate mechanisms in place to do so, this has raised the prospect of funding unsustainable budget deficits. Second, the implementation of development policies, plans and programs, particularly those internally financed, has not always met its objectives. Many expensive public sector programs, including some externally financed, have failed to fully meet their objectives efficiently, and have hence generated disappointment and discontent, and pressure on public agencies to improve the use of public resources. Third, poor performance, and the realization that more productive use of resources can be achieved, has led civil society and organized political opposition to call for strengthened government accountability for the use of public funds, and more generally for improvements in governance.

In all three countries, the demand for evaluation has originated in one or more of the above circumstances. By deciding to move forward with institutionalizing public sector evaluation functions, these governments have accepted that evaluation is an essential institution of good governance. Yet in all three cases, while demand has been a necessary condition for making the decision to start evaluation capacity development programs, leadership and vision to enact action on evaluation capacity has been the linchpin. In the three countries, leadership has been exercised in vastly different ways.

In Colombia, in 1991, the government of President Cesar Gaviria commissioned a high-level study to assess how Colombia should institutionalize an evaluation function. The government had been exposed to an external evaluation of investments in the key power sector, showing that Colombia could have improved substantially the outcome and services of multi-billion dollar

investments in the sector. The high-level study was commissioned to study the feasibility of introducing evaluation to cover major public sector expenditures. While doing the study, the government quickly recognized that to institutionalize the function would require political consensus bridging the interests of all concerned parties. To do this, the government commissioned five independent evaluations of large, controversial public sector programs. These programs covered a period which had contained various political regimes, and included a hydroelectric, an urban rapid transit, a papermill, a community welfare, and an international communications project (DNP, 1992). The aim of the evaluations was to enrich dialog among the various interest groups by providing them with specific evaluation results comparing what had been achieved with what could have been done, had those evaluations been available earlier in the decisionmaking process. The results of these evaluations were discussed in a high-level conference where past and present government decisionmakers, international experts, and political interest groups participated. The conclusion of the conference led to a recommendation that the Colombian constitution, which was being redrafted at the time, should include an article making public sector evaluation mandatory. Thus, through leadership, stakeholder involvement, and opportunism, a legal foundation for evaluation was established in Colombia.

In China, the early steps were taken by the country's Auditor General who had, through his extensive relationships with auditors in industrial countries and through formal contacts such as with the International Organization of Supreme Audit Institutions, become aware of and interested in performance auditing and program evaluation. While initially the Auditor General was particularly interested in starting by enhancing the State Audit Administration evaluation capabilities, the State Council decided to involve other key ministries. This broadened the diagnosis of the institutional framework on which to develop an evaluation system. In 1992, the Auditor General, the Ministry of Finance, the State Planning Commission, and other key central agencies, decided to move forward with an institutional study which would develop alternative proposals for recommendation to the State Council. Some specifics of the process are well detailed in Hong and Rist (1997). But the salient feature of the approach followed by China was that the Ministry of Finance, as the public agency spearheading the effort, realized that progress

was going to be made only if all relevant major public agencies were part of the process. Early on, the Ministry of Finance organized a steering committee with representatives of the major public agencies. This was the first time that many of these senior Chinese officials had ever met or worked together. The steering group, with support from national and international experts, drafted a proposal with various alternatives for an evaluation system (Khan 1994). The proposals were later discussed and an approach agreed in principle at a high-level ministerial conference. While formal approval for the preferred proposal is pending in the State Council, in 1995, tacit approval was given to the State Planning Commission to draft regulations governing evaluation in public sector agencies. Various key central agencies began to institutionalize evaluation capabilities within their own organizations. Today, more than ten major public agencies have set up evaluation offices and work programs. Before the Auditor General's initiative, evaluation was only carried out on a small number of externally funded projects by few agencies and in an ad hoc manner.

In Indonesia, while the World Bank since the early 1990s had been raising the issue of evaluation capacity in the context of the normal conduct of business, leadership calling for institutionalization came from a senior advisor to the President, who persuaded the National Development Planning Agency (BAPPENAS) to carry out an institutional diagnosis, leading to a proposal for setting up an evaluation system. The rationale for taking the step was that Indonesia's new development plan—REPELITA VI—needed to be concerned much more with equity and social issues. The government, in the face of growing demand for public resources and services, realized that, evaluation could help increase their productivity and minimize waste. In 1994 an institutional diagnosis was started and in 1996 a proposal for evaluation capacity development was endorsed by BAPPENAS (Barbarie 1995). BAPPENAS assembled a multi-agency steering group to provide direction for the institutional diagnosis. The diagnosis was carried out by national and international experts. To address the legal underpinning of the evaluation function and clarify incentives, the final proposal included drafting of a ministerial decree formally outlining the responsibilities of the various agencies in the evaluation system, and more closely linking the annual resource allocation process with performance monitoring and evaluation results. The Ministerial decree was formally issued in late 1996.

**Table 1: Evaluation Capacity Development and Institutionalization
Key Issues Addressed in Colombia, China, and Indonesia**

| Issue | Colombia | China | Indonesia |
|--|---|--|---|
| Anchoring the evaluation regime. | Constitution mandates the Executive to take lead. | State Council draft resolution calls on the Central Executive Agencies to take lead. | Responsibility rests with the Executive through a Ministerial decree. |
| Positioning the evaluation functions. | Centralized in the National Planning Department (NPD). Key line agencies provide inputs. | Decentralized to key central agencies. | Centralized in the National Development Planning Agency (BAPPENAS). Line agencies provide inputs. |
| Evaluation coverage. | Public policy and major public sector programs. | Public sector projects. | Development policies, plans, programs, and projects. |
| Linking evaluation with other public sector functions. | NPD plays a key role in policy and strategy formulation and budget allocation and monitoring. | No formal links have been established. State Planning Commission involved in public resources allocation and monitoring. | BAPPENAS to link evaluation to the annual budget allocation process. |
| Using evaluation in decisionmaking. | Monitoring and evaluation information to flow to line agency heads and the NPD. | Monitoring and evaluation information to inform central agency management. | Monitoring and evaluation information to flow through line agency management to BAPPENAS. |
| Professionalizing the evaluation function. | Evaluation is a trans-discipline cutting across specific professional skills. | Evaluation is seen primarily as applied socio-economic analysis. | Evaluation is not seen as a separate profession, but a complementary discipline. |
| Resources for evaluation. | Evaluation to be mainstreamed in agencies' budgets. | Evaluation mainstreamed in central agencies' budgets. | Evaluation mainstreamed in agencies' budgets. |

Demand for evaluation does not arise spontaneously; it must be nurtured through appropriate incentives and leadership. Laws and regulations are a necessary part of the incentives framework, but not a sufficient condition for a successful institutionalization of evaluation. The three country cases described here provide evidence of a varied set of conditions and circumstances. The successful implementation of an evaluation regime requires an enabling environment, that is, one in which there are real pressures for the type of institutional change that systematic public sector evaluation represents. The pressures can be external, driven by public accountability demands, or by ac-

countability to, for example, external funding agencies. The pressures can be internal, to meet laws, rules and regulations, or managerial needs. The pressures can also arise from resource constraints. As Derlien (1990) has shown, these conditions have a clear parallel in the development of evaluation in industrial countries. But institutionalizing evaluation also requires direction and leadership, that is, a champion with power and commitment to build consensus and make the necessary decisions. As will be discussed, an evaluation system must also be capable of implementation through actionable steps by skilled staff and financial resources. Finally, the incentives

framework must include effective rewards for those who are called to produce or use evaluation, or be responsible for evaluation outcomes.

Organizing the Evaluation Function

There are at least three aspects related to the organization of evaluation. Lessons on the location of the evaluation function, drawn from industrial country experience, include whether the evaluation function in the public sector should be in the executive or in the legislature, whether it should be centralized or decentralized: and whether to develop internal evaluation expertise or use outside experts. Also important are the linkages of evaluation to other key public sector functions and its implications for the organization of the function. How have Colombia, China, and Indonesia addressed these organizational aspects?

In Colombia the 1991 Constitution calls on the executive to lead the process of institutionalizing the evaluation function in the public sector. It also calls on the Controller General to independently carry out performance and compliance audits. The National Planning Department (NPD) is directed to implement the new constitutional mandate. It is doing so by linking evaluation to the National Development Plan, involving key government agencies responsible for implementing public expenditure programs (World Bank 1997). These agencies must prepare annual plans specifying expected outputs, outcomes and performance measures, which will serve as a basis for ongoing, and ex post evaluation of public sector performance. The planning units of each agency are responsible for the annual ex post evaluation activities and for delivering such evaluations to the NPD as a prerequisite for budget allocation requests. The NPD plans to conduct selected in-depth ex post evaluations of policies, sectors, programs, or entities, which it will use as a basis for defining the subsequent expenditure program, identifying institutional needs, and reviewing strategic policy formulation. In addition, inter-agency coordination units will bring together representatives of relevant stakeholders, including ministries, implementing agencies, departments and municipalities, and the civil sector. These units will provide coordination, cross-entity planning and evaluation oversight. To address the increasing demand for accountability, the proposed system also calls for the

Controller General to check periodically on data quality and collection procedures.

In China the evaluation system is being institutionalized following a decentralized approach. While the State Planning Commission is directed by the State Council to issue the regulations and to commission evaluation of public sector programs, evaluation units have developed in key central agencies. For example, the State Development Bank, Ministry of Construction, the China Engineering Consulting Corporation (a technical arm of the State Planning Commission) and others have set up evaluation units. In all these cases, the evaluation units are being located at the highest levels of each organization to ensure independence from operational line responsibilities. The ministerial proposal for the evaluation system described earlier calls for decentralizing evaluation responsibilities to provincial public sector agencies. It also calls on the State Audit Administration to provide oversight of the evaluation system. In China there are no formal links yet established to ensure that evaluation information is fed to other key government functions, such as public expenditure allocations. However, the infrastructure for such links is being set up. At this early stage of the institutionalization processing, China is putting greater effort into using evaluation information as a learning function, and to improve agency management of public sector programs.

In Indonesia the organization of the evaluation function in the public sector is following a path similar to Colombia's. Evaluation is seen as a tool to correct policy and public expenditure programs, through more direct linkages to the National Development Plan and the resource allocation processes. Specifically, the Indonesian evaluation system relies on strengthening evaluation in all the key central and decentralized government agencies, with BAPPENAS at the apex. The proposal under implementation calls for strengthening monitoring and performance measurement in the line units within each agency, coupled with establishing units independent from the line at the top of each agency. The function of these units is to ensure that all public expenditures have appropriate monitoring mechanisms, and that the information is evaluated at the agency strategic level. These units then feed the relevant evaluation information to BAPPENAS to support the annual budget request cycle. BAPPENAS, which is a powerful resource allocation agency, will

also strengthen its own evaluation capabilities to provide cross-sectoral and strategic evaluations to the top policymakers, and to facilitate adjustment to the development plans. The Indonesian evaluation institutionalization plan calls for an eventual oversight role for the internal and external audit agencies. This will come in due course when the audit agencies have been sufficiently strengthened to meet this responsibility.

For evaluation to be successfully implemented—and it is still too early to tell in the three cases presented here—it must be located in several places within the governance structure if its to meet the demands of the various markets and stakeholders. In both Colombia and Indonesia, the organization of evaluation in the public sector has been driven by a explicit expression of information needs from the demand side. In these two cases, the demand has come from the well-expressed decision of the powerful national planning agencies to link evaluation information to resource allocation processes. However, in all three countries, the impetus for evaluation institutionalization is lodged in the executive where the relevance of government programs is unlikely to be challenged. As these systems mature, these countries are likely to experience an increasing need to strengthen the oversight role of independent agencies reporting to the legislative arm of government. Indeed, in Colombia, the role of the Controller General in evaluation is recognized in the new constitution. In China the Auditor General, who played a pivotal role in the beginning of the evaluation institutionalization process, may (as contemplated in the proposal before the State Council) eventually take up specific responsibilities regarding oversight of the performance of the evaluation system. In Indonesia the situation is similar to that of China, although the Auditor General's office participated in setting up an evaluation system, it has not played an active part in defining its own role. In this respect, industrial country experience, such as Australia's, provides useful models of interaction between the executive and legislature, with regard to the role of the independent audit agency (Mackay, 1996).

Complementing these considerations, the three governments recognize the importance of both locating evaluation with program management, tied to expenditure implementation within agencies, and locating evaluation at the corporate level in the form of evaluation units separated from operational respon-

sibilities, and reporting independently to agency management. In Colombia and Indonesia an important aim is to decentralize evaluation responsibilities to all actors with public expenditure allocation responsibilities. It is important to note in this regard that Colombia and Indonesia are implementing fiscal decentralization policies. China is de facto decentralizing, and is equally concerned with extending evaluation responsibilities over time to all relevant public sector agencies, central and decentralized. A decentralized evaluation function with clear reporting responsibilities is a natural complement to the decentralization process itself. If a key objective of decentralization is to bring decisionmaking closer to the clients of public services, only a decentralized results-based evaluation process, for example, at the local government level, will ensure that adequate signals are generated about whether welfare-enhancing results are being attained. Often the local political process itself is supposed to assure feedback and improved accountability, but without transparent information on outcomes through systematic evaluation processes, it is naive to believe that decentralized public expenditures will yield desired results.

Supply Considerations

The most relevant spectrum of evaluation activities in developing countries ranges from policy, at the strategic and sectoral levels, to public expenditure programs, and to investment projects. This range of evaluation coverage defines the questions that evaluation must answer. These questions in turn determine the supply requirements. Resources required to carry out evaluations fall into four generally interdependent categories: methods and standards, information, financial resources, and professional skills. Information resources are largely determined by the specific evaluation methodologies to be used to answer the evaluation. But the methods themselves depend on the evaluation questions to be answered, the financial resources, skills, and the time available for evaluation. Evaluation timing is also key to meeting demand (Chelimsky, 1987) which often leads to trade-offs among these supply factors. For example, the best methodology may also be the most resource-intensive and the one that would take the longest to yield the information demanded. If evaluations are not produced in time to be responsive to demand, resources committed to them will be questioned and support for the function will wane.

In all three countries, these inputs are in constrained supply but while financial and information resources are limited, the more severe shortages are in the methodological and skills areas. Bemelmans-Videc, 1992, discusses in detail the human resource development issues involved in effective evaluation practice. While evaluation methods for program and project evaluation are well developed and known in industrial countries, methods for the evaluation of public policies are continuously evolving. Newer areas such as methods for the evaluation of policy advice are beginning to be developed and put into practice (Uhr and Mackay, 1996). While in all three countries there are academic programs to train professionals in the social and technical sciences, in none are there programs in evaluation. Evaluation, to the extent that it is practiced, is carried by professionals trained in other disciplines. Indeed, evaluation is not seen as a profession: rather it is seen as a trans-discipline tapping multi-disciplinary skills. All three countries have, nevertheless, been sending prospective evaluation staff for training abroad, or inviting foreign lecturers to train local officials and professionals in evaluation methods.

All three countries have also considered setting professional evaluation standards, with China and Indonesia having drafted specific proposals. China and Indonesia are also encouraging their academic communities to get involved in evaluation. In both countries, local universities participated actively in the institutional assessments made before launching the function. China is taking this one step beyond by pursuing the establishment of an evaluation professional association bringing together the official sector, local universities and research institutions, as well as external experts.

In Indonesia the evaluation system being put in place also calls for training in evaluation for different levels. It calls for training in technical skills for evaluators, interpersonal skills to encourage a multi-disciplinary evaluation approach, and management skills, for those who are put in charge of the evaluation units in the various agencies. Indonesia, as well as China, is also considering filling the skills gap by commissioning outside parties to carry out evaluation. Indonesia proposes to encourage the use of third parties when specialized expertise is needed, and to fill gaps and enhance independence. In general, evaluation is not yet seen as a career in the public services. Evaluation is seen with caution. Those who champion evaluation, have no doubt about

the implications of developing a strong function. But public officials called upon to undertake evaluation responsibilities express concern about their professional careers: they face the risk of being seen as conveyers of bad news. This is not an untypical reaction of evaluation practitioners everywhere. In the three countries reviewed here, as well as other developing countries institutionalizing the evaluation function, it is amply recognized that evaluation as an institution of governance will succeed only when an evaluation culture has been established. For evaluation to take root as a profession, it will be necessary to do what Australia and New Zealand are aiming to accomplish in creating an evaluation culture in government at all levels. Developing an evaluation culture is a long-term process of building incentives as well as capacity.

On information requirements, Colombia and Indonesia aim to identify monitoring and performance indicators to be used in the tracking and evaluation of public expenditure programs. Yet neither has the information systems in place to do so. China has a well-developed system of information requirements, mostly designed to track the implementation of public sector programs. All three countries are taking steps to complement their existing information systems with data requirements focusing on intermediate outcomes and results. This is also a long-term process which involves building the information infrastructure in parallel with the development of the evaluation function. In none of these countries has the ready availability of information been seen as a precondition to institutionalizing evaluation.

On the financial resource side, in all three countries evaluation is seen as a mainstream function of each public agency—and not an add-on unfunded responsibility. In China and Indonesia evaluation is to be funded from the regular budgets of the public sector agencies. No separate budgets are being considered to provide financial incentives for evaluation. In Colombia the budget is to provide a more strategic support for results-oriented public resource allocations. As already indicated, in Colombia, budget submissions should be underpinned by evaluation information.

Bridging Demand and Supply

The relatively recent history of evaluation capacity development in developing countries, and the experience of industrialized

countries confirm the major findings of the World Bank task force report (1994). These findings are worth synthesizing:

- Evaluation is a long-term, difficult, and complex proposition. Nurturing genuine demand for evaluation of the effectiveness of public actions and the need to build ownership are the pivotal steps in most countries. The three country cases discussed amply illustrate this.
- Evaluation capacity development must be treated as an integral part of efforts to improve public sector reform and institutional development. Using different instruments and approaches, the three countries have sought to develop evaluation as a mainstream function, and not a separate, isolated institution.
- The focus of evaluation capacity development interventions should depend on the institutional conditions of the country in question. In all three cases, the institutional framework is different and the evaluation approaches followed have been tailored to meet particular requirements. While each country has been keen to learn about the experience of others, this has not been with the intention to transfer approaches, but to adapt best practice to the local conditions and constraints.
- Information must be adequate for effective evaluation and must be accessible to those who need it. All three countries recognize the need for adequate information and have set up the mechanisms to begin to develop responsive information systems. While the information systems are being developed, these countries are introducing ad hoc information gathering approaches.
- Where past evaluation capacity development efforts have not yielded satisfactory results, they were usually seen as externally driven, with limited country participation, and lacking the necessary critical mass. While the World Bank has taken steps to help promote the institutionalization of evaluation, the focus of the effort has not been on externally funded programs and projects, but on public expenditure at large and the effort has been driven by local needs.
- The demand for and country ownership of evaluation can best be developed if it is seen as a vehicle for learning and improving future performance, not just as an accountability instrument. The institutional assessments that have preceded the institutionalization of evaluation have led to recommendations encouraging appropriate checks and balances, not only within

public agencies, but among agencies. The trend is to strengthen line as well as apex evaluation within agencies. In all three cases, evaluation in the executive agencies of government is complemented by the oversight function of independent agencies, such as Auditor Generals or Controllers not reporting to the executive.

As seen in the three country cases, to successfully bridge the demand and supply conditions affecting evaluation capacity development interventions required in-depth diagnosis of a country's institutional framework. From this diagnosis, realistic proposals were made about institutionalizing evaluation capacity (and its pace). The diagnostic studies provided a forum to develop and nurture demand and political support for evaluation. They helped point to areas where supply capacity needed improvement, such as in building skills and information infrastructure. They also helped identify how external support could be best used.

Table 2 illustrates more specific options in evaluation capacity development in response to different demand and supply conditions, using the framework outlined in the Introduction. Since no single evaluation capacity development strategy is appropriate for all countries, flexible and opportunistic approaches may work best. The interventions listed in the table are only suggestions and are to be treated neither as an exhaustive list nor as rigid proscriptions. Indeed, in each country there are likely to be present, as in the cases reviewed in this paper, both weak and strong features on both the demand and supply side. The framework only highlights the case for analyzing each intervention against demand and supply conditions so as to match the evaluation capacity development strategy to the country context. The approach underscores the need to differentiate between countries and to adopt a graduated evaluation capacity development strategy.

The mix of interventions will depend on the priority the evaluation function is to be given in the public sector management reform in each country. The evaluation capacity development interventions derived from this analysis should be matched with a country's ongoing or proposed public sector management reforms. These reforms should provide a basis for deciding the priority, mix, and timing of the interventions. For example, in countries where both demand and supply conditions are weak, it

Figure 2: Evaluation Capacity Development: Framework for Action

| | | Demand | |
|---------------|---------------|--|---|
| | | Strong | Weak |
| Supply | Strong | <ul style="list-style-type: none"> • support evaluation of policies, programs and projects • establish links between evaluation, strategic planning, resource allocation and budgets • use expert commissions to evaluate policies • strengthen evaluation in the legislature • disseminate evaluation results to the public • organize and systematize the evaluation function • support financial and information systems | <ul style="list-style-type: none"> • disseminate evaluation methods and practices • support ongoing evaluation of programs and projects • participate in evaluations done by external funding agencies • support professional development in evaluation • support research institutions in carrying out evaluation |
| | Weak | <ul style="list-style-type: none"> • disseminate lessons of experience and best practices • set up commissions to evaluate important projects or programs • train and use private sector institutions in evaluation • support university training in evaluation • provide technical assistance to government agencies • build an evaluation network within government • promote evaluation by non-government agencies • introduce evaluation standards | <ul style="list-style-type: none"> • strengthen audit and accounting • carry out joint evaluations with funding agencies • disseminate national and international lessons of experience • support evaluation training in educational institutions • carry out country institutional and evaluation capacity assessments • promote cross-country cooperation in evaluation • raise awareness among decisionmakers |

may be premature to think about locating the evaluation function in any part of government. In developing countries with important funding from external agencies, it might be more appropriate to initiate joint evaluation of specific projects and programs as a catalyst for further development. This is not to deny the importance of a focal point for evaluation to provide national leadership. But it might be prudent to be experimental and learn rather than to be dogmatic about the evaluation location issue. In general, the evaluation capacity development process should evolve in the direction of strong demand and supply. The road to get there may be circuitous.

In Conclusion

The earlier discussion of demand and supply issues in Colombia, China, and Indonesia shows how the framework can be used to

take stock and assess the conditions in any country. Although there is no need to repeat how these countries match the framework, a few points are worth highlighting:

- In China and Indonesia it was deemed appropriate to begin the evaluation capacity development process with an institutional diagnosis;
- In all three countries, there was an expressed need to raise awareness among key decisionmakers;
- In all three cases, there was a keen interest by the local authorities in learning about international experience, particularly about promising approaches to evaluation capacity development;
- Colombia has commissioned independent evaluation of important projects while the capacity is being developed and mainstreamed;

-
- In China and Indonesia there has been great interest in participating more actively in the evaluations carried out by external funding agencies such as the World Bank;
 - In both Colombia and Indonesia there is a policy decision to link evaluation to the key functions of government (such as strategic planning and resource allocation);
 - China and Indonesia are taking steps to strengthen the capacity of independent oversight agencies outside the executive;
 - Colombia is actively promoting the dissemination of evaluation results to the public.

In short, steps are being taken in these countries to address weak demand and supply areas. Having made the decision to institu-

tionalize evaluation in government, authorities are impatient to adopt or put in place evaluation mechanisms and procedures akin to those being set up in industrial countries. These, and other developing countries that have embarked on strengthening their evaluation capacities as part of their public sector reform programs, are intent on bypassing the waves (Derlien, 1990) that have characterized the ups and downs of evaluation in industrial countries, and move to adopt more robust and sustainable approaches based on what works. Given that there are no general solutions for institutionalizing evaluation, it is for the interested reader to draw specific conclusions about what is useful or relevant to the circumstances.

References

- Barbarie, A. (1995). "Strengthening Performance Evaluation Capacity". Unpublished report prepared for the National Development Planning Agency . Republic of Indonesia. Jakarta
- Bates, R.H. (1995). "Social Dilemmas and Rational Individuals - An assessment of the new institutionalism". In J. Harriss, J. Hunter, and C.L. Lewis (eds.), *The New Institutional Economics and Third World Development*, New York: Routledge
- Bemeimans-Videc. M.L. (1992). "Institutionalization Evaluation: International Perspectives". In J. Mayne and others (eds.), *Advancing Public Policy Evaluation: Learning from International Experiences*. New York: North-Holland
- Chelimsky, E. (1987). "What Have We Learned About the Politics of Program Evaluation?". *Evaluation Practice*. Volume 8, Number 1, February, 1987. Thousand Oaks: Sage Publications
- Departamento Nacional de Planeacion (1992). *Conmol y Evaluacion de la Gestion Publica - Los Desafios de La Nueva Constitucion*. Santafe de Bogota: DNP
- Derlien, H-U (1990). "Genesis and Structure of Evaluation Efforts in Comparative Perspective". In R.C. Rist (ed.), *Program Evaluation and the Management of Government - Patterns & Prospects across Eight Nations*. New Jersey: Transactions Publishers
- Havens, Harry S. (1992). "The Erosion of Federal Program Evaluation". In C.G. Wye and R.C. Sonnichsen (eds.), *Evaluation in the Federal Government: Changes, Trends, and Opportunities*. New Directions for Program Evaluation. Number 55. Fall 1992,. Jossey-Bass Publishers
- Hong, H. and R.C. Rist (1997). "The Development of Evaluation in the People's Republic of China". In E. Chelimsky and W.R. Shadish (eds.), *Evaluation for the 21st Century - A Handbook*. Thousand Oaks: Sage Publications
- Khan. M.A. (1994). "Establishing a Performance Evaluation Capacity in the Public Sector of the People's Republic of China". Unpublished report prepared for the World Bank Department: Ministry of Finance, Beijing
- Mackay, K. (1996). "The Institutional Framework for Evaluation in the Australian Government". Paper delivered to a World Bank seminar, Department of Finance, Canberra
- Stiglitz. J.E. (1989). "On the Economic Role of the State". In A. Heertje (ed.), *The Economic Role of the State*. London: Basil Blackwell
- Stiglitz. J.E. (1994). *Whither Socialism*. Massachusetts: The MIT Press
- The World Bank (1994). *Report of the Evaluation Capacity Development Task Force*. Washington, DC: The World Bank
- The World Bank (1997). *Colombia - Paving the Way for a Results-Oriented Public Sector*. Washington, DC: The World Bank
- Uhr, J. and K. Mackay (1996). *Evaluating Policy Advice: Learning from Commonwealth Experience*. Federalism Research Center, Australian National University. Canberra

Wiesner, E.D. (1993). *From Macroeconomic Correction to Public Sector Reform - The Critical Role of Evaluation*. World Bank Discussion Paper No. 214. Washington, DC: The World Bank

Wiesner, E.D. (1997). "Evaluation. Markets and Institutions in the Reform Agenda of Developing Countries". In E. Chelimsky, and W.R. Shadish (eds.), *Evaluation for the 21st Century - A Handbook*. Thousand Oaks: Sage Publications

Other Papers in this Series

Keith Mackay, *Lessons From National Experience*.

Stephen Brushett, *Zimbabwe: Issues and Opportunities*.

Alain Barberie, *Indonesia's National Evaluation System*.

Keith Mackay, *The Development of Australia's Evaluation System*.

Keith Mackay, *Evaluation Capacity Development: A Diagnostic Guide and Action Framework*

Other Recommended Reading

1. Keith Mackay (ed.), *Public Sector Performance—The Critical Role of Evaluation*.
2. The World Bank, *Building Evaluation Capacity*, Lessons & Practices no. 4 1994.*
3. The World Bank, *Designing Project Monitoring and Evaluation*, Lessons & Practices no. 8 1996.*
4. The World Bank, *Evaluating Development Outcomes: Methods for Judging Outcomes and Impacts*, Lessons & Practices no. 10 1997.*
5. The World Bank, *Assessing Development Effectiveness*.*

* These publications are soon to be available in both French and Spanish.

Operations Evaluations Department Publications

The Operations Evaluation Department (OED), an independent evaluation unit reporting to the World Bank's Executive Directors, rates the development impact and performance of all the Bank's completed lending operations. Results and recommendations are reported to the Executive Directors and fed back into the design and implementation of new policies and projects. In addition to the individual operations and country assistance programs, OED evaluates the Bank's policies and processes.

Operations evaluation studies, World Bank discussion papers, and all other documents are available from the World Bank InfoShop.

Summaries of studies and the full text of the *Précis* and *Lessons & Practices* can be read on the Internet at <http://www.worldbank.org/html/oed/index.htm>

How To Order OED Publications

Documents listed with a stock number and price code may be obtained through the World Bank's mail order service or from its InfoShop in downtown Washington, DC. For information on all other documents contact the World Bank InfoShop.

Ordering World Bank Publications

Customers in the United States and in territories not served by any of the Bank's publication distributors may send publication orders to:

The World Bank
P.O. Box 960
Herndon, VA 20172-0960
Fax: (703) 661-1501
Telephone: (703) 661-1580

The address for the World Bank publication database on the internet is: <http://www.worldbank.org> (*select publications/project info*).

E-mail: pic@worldbank.org
Fax number: (202) 522-1500
Telephone number: (202) 458-5454

The World Bank InfoShop

The World Bank InfoShop serves walk-in customers only. The InfoShop is located at:

701 18th Street, NW
Washington, DC 20433, USA

All other customers must place their orders through their local distributors.

Ordering via E-Mail

If you have an established account with the World Bank, you may transmit your order via electronic mail on the internet to: **books@worldbank.org** Please include your account number, billing and shipping addresses, the title and order number, quantity, and unit price for each item.

OED Study Series

- Evaluation and Development: The Institutional Dimension (1998)
1997 Annual Review of Development Effectiveness (1998)
India: The Dairy Revolution (1998)
The World Bank's Experience with Post-Conflict Reconstruction (1998)
Financial Sector Reform: A Review of World Bank Assistance (1998)
Rebuilding the Mozambique Economy: Assessment of a Development Partnership (1998)
Mainstreaming Gender in World Bank Lending: An Update (1997)
Agricultural Extension and Research: Achievements and Problems in National Systems (1997)
Fiscal Management in Adjustment Lending (1997)
Reforming Agriculture: The World Bank Goes to Market (1997)
Paddy Irrigation and Water Management in Southeast Asia (1997)
Poland Country Assistance Review: Partnership in a Transition Economy (1997)
1995 Evaluation Results (1997)
Zambia Country Assistance Review: Turning an Economy Around (1997)
The Aga Khan Support Program: A Third Evaluation (1996)
Lending for Electric Power in Sub-Saharan Africa (1996)
Industrial Restructuring: World Bank Experience, Future Challenges (1996)
Social Dimensions of Adjustment: World Bank Experience, 1980-93 (1996)
1994 Evaluation Results (1996)
Ghana Country Assistance Review: A Review of Development Effectiveness (1995)
- Evaluation and Development: Proceedings of the 1994 World Bank Conference (1995)
Developing Industrial Technology: Lessons for Policy and Practice (1995)
The World Bank and Irrigation (1995)
1993 Evaluation Results (1995)
Structural and Sectoral Adjustment: World Bank Experience, 1980-82 (1995)
Gender Issues in World Bank Lending (1995)
The World Bank's Role in Human Resource Development in Sub-Saharan Africa: Education, Training, and Technical Assistance (1994)
1992 Evaluation Results (1994)
New Lessons from Old Projects: The Workings on Rural Development in Northeast Brazil (1993, contains summaries in French, Portuguese, and Spanish)
World Bank Approaches to the Environment in Brazil (1993, contains summaries in French, Portuguese, and Spanish)
Rapid Appraisal Methods (1993)
Trade Policy Reforms Under Adjustment Programs (1992)
World Bank Support for Industrialization in Korea, India, and Indonesia (1992)
Population and the World Bank: Implications from Eight Case Studies (1992)
The Aga Khan Rural Support Program in Pakistan: Second Interim Evaluation (1990)