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

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

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

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

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

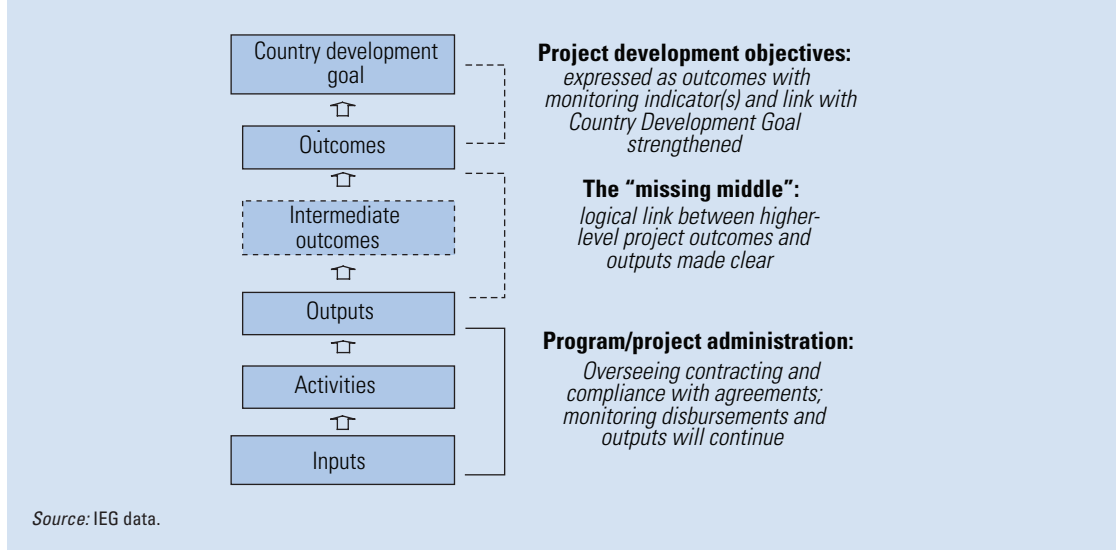
tracking inputs (disbursements), and possibly project outputs, while evaluation is done at the end when the "results" are clear. At the project level, evaluation has meant the ICR, which is then validated by IEG.

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

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

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

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

Box B.1: Implications of Results for Bank Operations

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

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

Monitoring

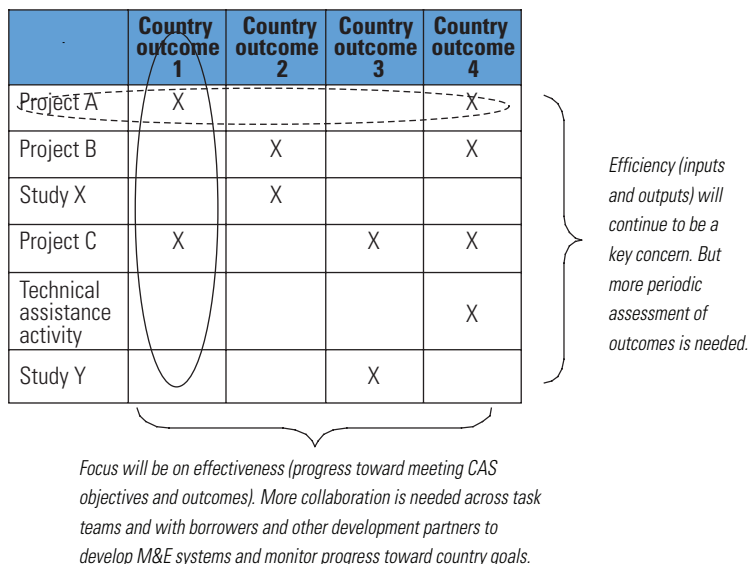
The increasing focus on results will require changes in the way the Bank approaches monitoring of its operations. At the activity level (projects and AAA), this will require an increase in the amount of information collected and monitored. Inputs (contracts, disbursements) and outputs will continue to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

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

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

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

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



Source: IEG data.

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

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

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

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

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

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

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

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

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

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

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

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

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

Box B.2: Possible Uses of Performance Information

To Communicate

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

To Manage

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

To Learn

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

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

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

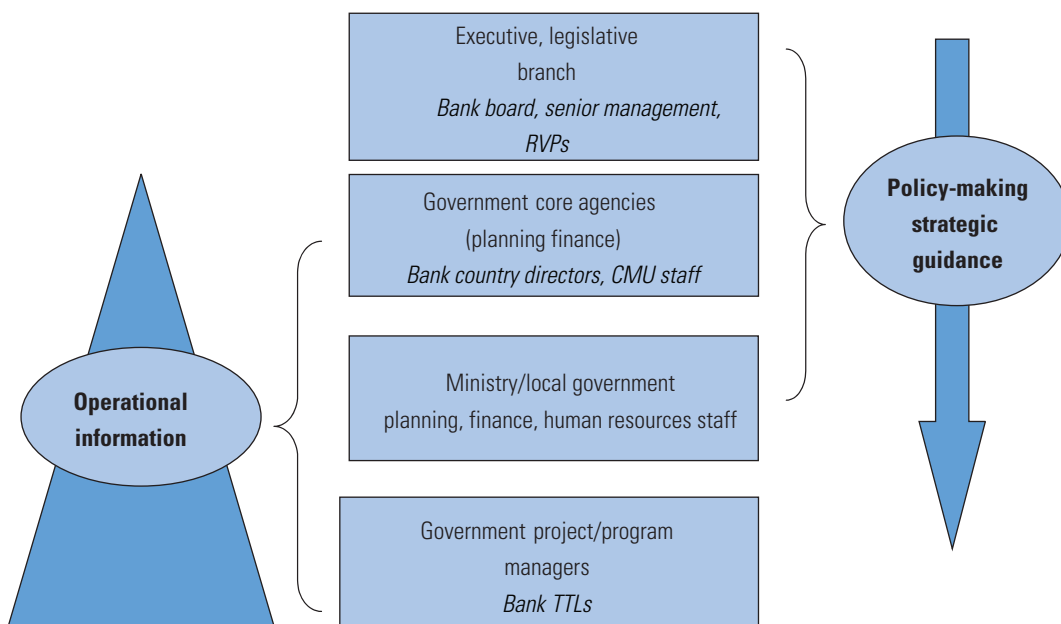
Finally, senior managers (in both the government and the Bank) would be interested in performance information, but largely to address accountability concerns of their enabling environment (the legislature, the Bank Board, and so on). Those higher up in the Bank or government hierarchy will be less concerned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

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

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

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

Figure B.2: Audiences for Performance Information



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

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

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

Evaluation

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

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

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

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

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

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

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

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

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

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

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

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

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

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