



**Office of Evaluation and Oversight, OVE
Inter-American Development Bank**

1300 New York Ave. N.W., Washington, D.C. 20577

RE-258

***Summary of OVE work on
Development Effectiveness***



Officially distributed to the Board of Executive Directors on January 9, 2002.

SUMMARY OF OVE WORK ON DEVELOPMENT EFFECTIVENESS

1. The Bank's Office of Evaluation and Oversight (OVE) has undertaken several studies during 2001 which bear on the issue of development effectiveness. Two of these, a Report on Development Effectiveness and a Report on the Results of the 2001 Evaluability Assessment Exercise, will be presented to the Board of Executive Directors early in 2002. Because both bear on issues currently before the Committee of the Board of Governors, this note is designed to summarize briefly the principal conclusions of these studies and suggest the implications of these findings for possible future changes in Bank procedures.

DEVELOPMENT EFFECTIVENESS

2. Substantial confusion surrounds the concept of "development effectiveness," primarily because the underlying concept—development itself—has undergone a major evolution during the 40 years of the Bank's existence. Development thinking has changed markedly along two different dimensions: First, there has been an evolution in thinking about what aspects of the human condition should be seen as the ultimate goal of "development". At the time of the Bank's founding, the focus was on the Alliance for Progress objectives of improving the lives and welfare of the individual. Employment and private productive activity were key to this objective, and growth in per capita GDP was seen as a useful proxy for measuring progress toward it. Concern that aggregate figures did not capture important elements of distribution led to an emphasis first on "redistribution with growth" and later on "basic human needs" and poverty reduction as the ultimate end of development activity. Recently, development thinking has focused on issues of capabilities and freedoms as the final end.
3. Second, development thinking has also evolved along a dimension dealing with the required interventions needed to pursue the evolving goal of development. Initially, the focus was on physical capital investment as the "motor" driving GDP growth. Poor results from this approach shifted attention first to human capital investment and later to issues of policy environments, markets, institutions and governance as the critical focus for intervention.
4. At the end of this process of evolution of concepts and mechanisms, the IDB and other MDBs find themselves with a huge range of potential goals and plausible interventions. The IDB-8 Agreement, for example, contains over 200 specific areas of possible action for the Bank. Yet there are few commonly accepted means of measuring achievement or weighing one intervention against another on a common scale of merit.
5. Such a weighing was possible (though not always practiced) during the initial phase of Bank operations. When the goal was growth in per capita GDP and the principal intervention was physical capital investment, then all projects could theoretically be

measured by a common standard: the contribution to GDP measured by a project's economic rate of return. Future flows of benefits could be projected, and compared with costs, so the Bank and the borrower could choose the project mix with the highest return relative to costs.

6. While never fully implemented in this ideal sense, cost benefit analysis became an important tool for the analysis of real benefit flows, particularly in the Bank's core business area of infrastructure finance. As development thinking evolved and the Bank was asked to undertake projects in other areas, however, it became harder to forecast future benefit flows, and the types of benefits captured in cost-benefit calculations appeared less relevant to the kinds of objectives which ought to be pursued. In 1981, the Bank modified its project preparation guidelines, noting *"In certain fields such as health, education, electricity and rural potable water, it is not always feasible to estimate the value of economic benefits. As a consequence, the economic efficiency analysis may be limited to a "least cost" or "cost-effectiveness" approach. This methodology compares the present value of the economic efficiency cost of alternative projects to attain given objectives, so as to select the alternatives with the lowest cost."* (OP-302-3, March 10, 1981)
7. This switch to pursuing "given objectives" allowed the Bank to pursue the broadest possible array of interventions, without firm reference to some external and comparable measure of value. This change had two very significant effects. First, it placed all of the burden for judgement regarding the development effectiveness of an intervention on the design of that particular intervention. The value of any intervention came to depend upon the quality of its objectives, since a project would be deemed "effective" if it met its own, self-defined development objectives.
8. Second, it rendered it impossible to compare one intervention against another, or one instrument against another in terms of a common standard. Choice of project or instrument was always a one-off affair, and the entire collection of Bank operations could no longer be summed and reported as progress toward a common objective of "development."
9. These aspects of project design have implications for project assessment, the raw material from which judgements about effectiveness are constructed. Current project assessment practice both in the IDB and the World Bank reflects this model. IDB projects are rated by Management in terms of a four point "probability of achieving development objectives" (highly probable, probable, low probability, improbable), supplemented with separate indicators of implementation progress, continued validity of assumptions, and at risk potential. The World Bank's Operations Evaluation Department provides a rating independent of Management and uses both on a six-point "outcome" scale (highly satisfactory, satisfactory, marginally satisfactory, marginally unsatisfactory, unsatisfactory, highly unsatisfactory.) and a 12 point Aggregate Project Performance Index which adds judgements regarding sustainability and institutional development impact to the basic "outcome" measure.

10. Despite many differences in methodology, both institutions base their judgement of development effectiveness on the achievement of objectives explicitly defined in the project. Both institutions find that the vast majority of completed projects receive satisfactory ratings in terms of outcome: in 2000 the World Bank's independent evaluation unit rated 76 percent of completed projects "satisfactory" or better, while self-evaluation by IDB Management rated 88% of projects in execution as "likely or highly likely" to achieve their development objectives.
11. The risk of such a methodology is that projects may define limited and trivial objectives, allowing the institution to claim success for the project without any meaningful impact on the welfare of the borrowing country. To guard against this risk, the Operations Evaluation Department of the World Bank applies a "relevance test" to determine whether a project's objectives were consistent with the development needs of the country before rendering a final judgement regarding outcomes, and, where possible, seeks to estimate the likely economic rate of return of completed projects. The IDB system of self-evaluation does not apply similar tests to the rating of its projects, although it is very likely that judgements about relevance are made informally by Management throughout the design and review process.
12. To gain some perspective on the documentation of the qualitative judgements made in assessment of Bank operations, OVE examined the written record on all completed projects which met two tests: they were rated "highly probable" of achieving their development objectives, and they had an available PCR. 47 projects met these two tests, and in the overwhelming majority of cases, the projects were considered successful simply because they produced their planned outputs. Although the Bank's PCR guidelines request information on both project outputs and the use made of the outputs to improve conditions in borrowing countries (outcomes), only a small minority of projects provided any evidence whatsoever regarding outcomes. Even projects which had established ex-ante rates of return provided no ex-post recalculation of actually-obtained rates of return, despite clear instructions in the PCR Guidelines that such re-estimates were required where feasible.
13. In interpreting these findings, it is important to bear in mind that this exercise only looked at the formal documentary record which underlay project assessments. From this data, we only know that outcomes and impact are not documented, not that the projects had no outcomes or impact. Management has established two Loan Committee working groups to explore improvements in both the concepts and processes used to design and evaluate projects, and their work may help to redress some of these concerns in the future.

EVALUABILITY

14. To gain a clearer understanding of how effectiveness issues are incorporated into current project design, OVE undertook, with the active support of Management, a review of all projects sent through Management's Loan Committee and on to the Board of Directors for approval during 2001. The purpose of the exercise was to look

at the extent to which the projects were designed to demonstrate their effectiveness in addressing the development challenges of borrowing member countries. Projects designed to demonstrate development effectiveness are easy to evaluate, since they are explicit as to goals and incorporate mechanisms for monitoring, so OVE developed an “evaluability assessment tool” as a mechanism for measuring design quality. The results of individual project assessments were provided to Management at the time of Loan Committee consideration as a way of improving the review of this dimension of projects.

15. OVE’s evaluability assessment measures 9 aspects of a project, which can be summarized along three different dimensions. First, the dimension of analysis, rates the project’s diagnosis of a problem, its understanding of the causes, the logic which connects the project to the problem, and the quality of its assessment of the risks involved. Second, the dimension of verifiability rates the project’s goals in terms of indicators, provision of baseline or starting-point data, and the establishment of target objectives in terms of both planned outputs and desired outcomes. Third, the dimension of monitoring rates the projects provision for data collection and feedback on project achievements during the course of executions, including provisions for both mid-term and at-completion evaluation.
16. The ratings for the 2001 group of projects show considerable shortfalls in documentation along all three dimensions. Few projects were considered to have articulated a strong analytical foundation, with the area of assumptions and risks being the weakest element. A somewhat larger number were considered to have adequate provisions for feedback and monitoring built into the design, although these elements were almost always something to be added in the course of execution, and were specific to the individual operation. It was rare to encounter attempts to link specific operations to more general systems of evaluation in the borrowing member country (where such systems exist).
17. Verifiability, however, was an issue in most of the 2001 crop of projects. Very few projects defined meaningful indicators of performance, while very few had either baseline data or explicit targets and milestones. Where such empirical indicators and targets did exist, they were applied to project outputs only, not to the outcomes anticipated to be realized in the country once the project had produced its outputs. A number of projects indicated their intent to establish clear baselines, targets, benchmarks and indicators prior to first disbursement, but it is too soon to tell if these expectations were realized for the projects approved in 2001.
18. These findings suggest that Bank projects are currently not being designed or monitored so as to transparently demonstrate development results. While a serious concern, this is not a problem unique to the IDB. A recent review of attempts to introduce results-oriented systems in development cooperation agencies¹ found that

¹ **Results Based Management in the Development Cooperation Agencies** <http://accsubs.unsystem.org/ccaqfb-intranet/references/dacrbm.pdf>

they all were struggling with common problems of how to institute effective processes and practices for measuring their performance. In particular, development cooperation agencies share with most other public institutions a strong ingrained tendency to focus on outputs—things under the direct control of the agency—rather than on outcomes. A renewed focus on results, however, requires a break with this pattern. As the report noted:

Results are usually measured at three levels -- immediate outputs, intermediate outcomes and long-term impacts. Whereas traditionally the development agencies focused mostly on implementation concerns, as they embrace results based management their focus is increasingly on measurement of results. Moreover, emphasis is shifting from immediate results (outputs) to medium and long-term results (outcomes, impacts).

19. This is the same challenge facing the IDB. At the 2001 Annual Meeting of the Board of Governors, President Iglesias noted: *"The objective of making the institution more flexible is inextricably tied to the goal of increasing the rigor and depth of the diagnostic assessment, analysis, and evaluation of the results of the Bank's action"*, To meet this challenge, the Bank will need to re-examine a number of practices and procedures relating to the design, execution, supervision and evaluation of its activities.

DESIGNING FOR EFFECTIVENESS

20. Re-focusing the work of the Bank on development effectiveness requires new initiatives in four areas: locating projects within a developmentally meaningful framework at the country level; improving the rigor of project design; experimenting with results-driven project models; and developing cooperative monitoring and evaluation systems with executing agencies and borrowing governments.
21. **Locating projects within the developmental challenges of their country** is a task which falls to programming. A recent OVE study (WP-02/01) has shown that, in general, Bank country program documents do not develop strategic frameworks, nor define measurable and evaluable goals for Bank intervention. This finding raises concerns about the Bank's capacity to locate projects in areas of high developmental significance for individual borrowing member countries. At the present time, the Board of Executive Directors is discussing the OVE study along with Management's proposed revision to the country paper guidelines. The thrust of these discussions has been toward a more disciplined and results-focused country programming process, from which should arise a significant improvement in the ability of the Bank to locate its projects within a country-defined results framework. This increased focus on concrete goals and measurable results is compatible with the recent acceptance of goal-oriented approaches to development, such as those embodied in the Millennium Development Goals and the Poverty Reduction Strategy Paper exercise within HIPIC initiative.

22. A second mechanism for reinforcing the results orientation to country programming would be to expand the annual portfolio review exercise to include discussion of progress toward achievement of development outcomes. The Annual Review of Portfolio Execution produced by ROS could contain specific sections dealing with outcome effectiveness, and each annual programming review update could provide the same sort of information at the country level. This would help subsequent programming exercises, including the drafting of Country Papers, to incorporate the lessons learned from portfolio implementation.
23. **Improving the rigor of project design** probably requires some re-engineering of the Bank's project review process. OVE's evaluability assessment exercise found that the current review process does not consistently generate projects which are designed to transparently demonstrate the developmental outcomes of Bank action. The exercise also found that comments made to the Loan Committee came too late in the review process to generate meaningful improvement in project design. This observation suggests that the best opportunities for improving the results-focus of projects lies somewhere "upstream" from the Loan Committee, either at the Programming Committee, when a project is initially approved for inclusion in the pipeline, or in subsequent phases of the review process.
24. The Bank's project review process has had a long evolution, driven largely by attempts to balance the two objectives of timely approval and developmental quality. During the 1970s, the Bank managed the quality control process by using a "two key" system for project approval. The Bank's Operations Department held one key, while the Project Analysis Department held the other. Both Departments reported to the EVP, and both had to certify that a project was ready before it could be considered for approval. During the financial crises of the 1980s, the Bank established an expedited one-key approach for the design of sector loans, giving the Department of Plans and Programs exclusive jurisdiction over this instrument. The reorganization of 1994 moved further in the direction of expedited approval, moving sector lending to the Regional Operations Departments while at the same time allowing them to incorporate the technical units formerly reporting to the Project Analysis Department.
25. This consolidation of project preparation has probably increased the flexibility and responsiveness of the Bank, but at the cost of tipping the balance away from extensive documentation of anticipated results. After 1994, quality control was the responsibility of the Loan Committee, assisted primarily by the Office of the Controller. The institutional emphasis on responsiveness and origination, however, led to a decline of this system of checks and balances, a process which culminated in the abolition of the Controller's Office in 1999. OVEs 2001 evaluability exercise suggests strongly that this process provides an inadequate focus on building explicit, transparent, documented and evaluable results frameworks into Bank operations.
26. A review of practice in other financial institutions, both multilateral and private, finds a broad recognition that the pressures and incentives created within units involved in loan origination creates what the Bank's own TAPOMA Report called an "approvals

culture.” To guard against excessive approvals pressure, other financial institutions have established independent units to manage quality control functions.

27. In the World Bank, quality control has been addressed through the creation in 1996 of a Quality Assurance Group (QAG), an independent body heavily reliant on outside technical experts, to rate the “quality at entry” of projects, as well as providing quality of supervision reviews for a sample of ongoing projects and quality reviews of economic and sector work. The existence of the QAG has had an impact on project quality via direct review of projects and via changes in the institutional incentive system. While the QAG only reviews a sample of projects, and provides “quality at entry” ratings only after project approval, the ratings are published and the potential for QAG review improves the incentives for rigorous project design.
28. In private financial institutions, risk is the critical dimension of quality, and it is considered good practice to have risk assessment of all projects managed by a separate department, completely independent of loan origination. Disputes between originators and risk assessors are brought to loan committees for resolution. The recent External Review Group on the IDBs Private Sector Operations recommended a similar structure for addressing risk management issues in the Bank’s lending to the private sector without government guarantee.
29. Just as private sector projects have financial risk, all of the Bank’s projects also have development risk—the possibility that countries may borrow money and get a flow of future benefits which is significantly smaller than the future debt service costs. Unlike financial risk, development risk is a risk to the borrower that the project may not yield sufficient results, regardless of the risk to the lender that the loan would be repaid. Development risk assessment could therefore be a potent tool for improving development effectiveness of project design, and, if applied uniformly across the entire range of Bank instruments, would provide some mechanism to assess the relative developmental risks posed by each different type of instrument.
30. The creation of a separate Development Effectiveness Assessment Unit, with a direct reporting line to the EVP, would subject all projects (not merely a sample as with the QAG) to an independent review on development effectiveness issues as part of the approval process. Conflicts regarding project design would be brought to the Loan Committee for resolution. Potentially, such a unit would also review Country Papers with a view to improving the results focus of country programming. (See paragraph 20 above), and review the quality of supervision in bringing about improved results during the execution process. Such a unit would help Management hold project teams accountable for the design and execution of projects which demonstrate measurable achievement of country-specific development goals.
31. **Experimenting with results-driven project models** In addition to improvements in the project review process, the Bank should more actively experiment with new lending instruments having an explicit focus on results. Management last year provided a background paper advocating the creation of such instruments as the

Performance Driven Program Loan, which would tie future disbursements to the achievement of concrete results. Such instruments clearly have potential, particularly as a way of experimenting with the design of results frameworks, and should be given an opportunity.

32. **Developing cooperative monitoring and evaluation systems.** Finally, an improved focus on results must be a shared responsibility with the borrowing member countries. Country program results frameworks need to be developed by the country, and should delineate areas of activity for the IDB and other development finance actors. IDB programs, and individual projects, should be making contributions toward the country's own development goals, and therefore should be monitored by domestic planning and evaluation units. Reinforcing such administrative structures where they exist, and developing them where they do not, should be an essential component of Bank operations. Some exploration of the complexity of the process of "evaluation capacity building" is contained in OVE's Report on Evaluation Capacity Building (RE-252) provided to the Board during 2001.