

Building Up the Capacity and Influence of Evaluation in the Bank and Outside, 1986-92

Yves Rovani
Former Director-General, Operations Evaluation

By the time I became DGO in 1986, the need for evaluation in the Bank had vastly increased. The operations then being completed reflected the new policies the Bank had introduced in the 1970s and 1980s, addressing poverty alleviation in general, but also energy. The first adjustment loans, a new instrument introduced in the early 1980s, with a shorter life than traditional projects, were also being completed. Moreover, evaluation was increasingly expected by the international community. This was the time when development effectiveness was being increasingly questioned, not in the streets as it is today, but by scholars who needed factual and honest answers.

But while the need for evaluation was growing, I think the evaluation capacity of the Bank was being undermined to some extent. Certainly it had been receiving much less support than under Robert McNamara, who had introduced the function and played an active role in making sure it was used. When Mervyn Weiner retired, management and the Board appointed Shiv Kapur, then the director, for an excessively short term, and recruited externally a director to be his successor.

This did not work out, and I was chosen as DGO on short notice. I happened to have been a supporter of evaluation (and self-evaluation) from its beginnings. I brought a lifetime of operational experience gained “in the trenches” as a project staff member and as a manager responsible for sector policies, strategies, and operations.

I saw my mandate at the time as twofold: to breathe new life into the Bank's evaluation system, and to encourage the further development of evaluation capacity outside the Bank. I'll discuss these two in turn.

The Bank's evaluation system

On strengthening the Bank's own evaluation system, I'll talk successively about the people, the output, and dissemination and feedback.

People, culture, management

Starting with the people, I was told as I arrived on the scene that at a staff meeting of OED, Tom Clausen, the Bank's President, had asked “What's your impact”? At the time, some in OED saw this as a rather impertinent question. They took the view that “Our job is accountability, and accountability justifies our existence—not the marketing of our

products, which in effect is in competition with our accountability function.” This made me believe we had a cultural problem to solve.

We also had a management problem, because the department as such existed only on paper. There were no central administrative functions to speak of. There was no budgeting. There was no cost information. Common services were absent. A divisional structure had been adopted about two years earlier, but had been effectively implemented only as far as one division was concerned, and that was the division of Graham Donaldson, which was responsible for agriculture and the social sectors. His division was to me the mother of what I believed OED should develop into. The rest of the department included some very good people, such as the late Carl Jayarajah, who was an outstanding fellow. It also included a bunch of prima donnas who absolutely did not want any form of structure over them.

We solved both the cultural and the management problem.

As regards the culture, my first move was to set up a staff retreat in Annapolis, putting on the table the issues of development effectiveness and policy toward the environment, and opening discussions about our own objectives and products, and the issues that were confronting support and operational staff in their day-to-day work. There was a lot of grumbling about this, but out of these discussions I think life started to spring. We put a number of task forces to work: one was to review the project completion reports issue; another was to look into support staff issues; and another was to look at office technology.

Inside of two or three years, we had the most modern office technology in the Bank, and we were well on our way to having an evaluation data system that used modern technology, and all the credit was, in effect, the staff's.

The culture I was trying to instill was a participatory style of management. No style of management, in my view, is more applicable to a unit of self-starters, such as OED. (Also—and I'll make this point later when talking about dissemination—you need high-caliber people to interact with operational staff if you want to have an impact on the quality of operations.)

On the management side, my first job was to appoint a director. I was told that I did not need one, and that the DGO-OED Director relationship had been fraught with problems. But I insisted on retaining the position and I looked for the best person I could find to meet the needs of the time. I found Ram Chopra, then an assistant director in the economics staff—who was not on the list I had been given. Three years later, when Ram left to take a job as country director in the Middle East, I chose Eberhardt Kopp to replace him. Eberhardt was a senior director with outstanding credentials in the Bank's operational policies and country operations management. The DGO-Director relationship worked extremely well throughout—clearly the source of past problems had been the appointment of the director by Bank management, rather than by the DGO!

In the first two years of my mandate, we experienced a major staff rotation of close to 50 percent—due in particular to the effects of the Bank's reorganization of 1987. We did

attract top talent, and were able over time to maintain high recruitment standards for staff as well as for managers. We also were able to adjust our staff profile to the changing needs of our programs, which our longer-range plans helped us to anticipate. As rotations continued, several staff left through promotions to other parts of the Bank, helping to dispel the widely held view that OED staff were less than welcome in the rest of the Bank.

One problem I did not have was continuity. Coming into my new job as DGO at short notice, and without a director, I found my life was greatly eased by the DGO's assistant, Ms. Rosa Maria Duncan, now the late Mrs. Rosa Maria Wiemann. Throughout my tenure, the smooth running of the evaluation function owed much to Rosa's institutional knowledge and to her professionalism. I admired and was very fond of her.

Output

Turning now to evaluation output, by 1986 we had a significant resource problem in the face of mounting demands for evaluation.

All of our friends, and the not-so friendly, were saying, drop the audit ratio, do fewer performance audits, and you will generate the resources needed to support the program. We did drop the audit ratio a little bit, setting a lower target of auditing no fewer than 40 percent of completed operations. But there was a limit to how far we could go if we were to live up to our mandate of ensuring the Bank's accountability—especially because the system of project completion reporting was decaying.

Project completion reports were declining in quality and submitted with increasing delays. A large backlog had accumulated. The reason for this was that someone had the bad idea around 1980 of shifting the responsibility for preparing project completion reports away from Bank operational staff, the authors of the operations, to the borrowers. A covenant had been introduced in loan agreements requiring borrowers to do the project completion reporting.

Of course, this idea was well meant. The borrowers were supposed to do a good job of writing the reports, and the staff were supposed to do a good job of reviewing them, and the reports would be sent to OED, and everybody would be happy. But the facts were otherwise. Anybody who has been involved in project supervision, as I have for a good part of my life, and who has seen the kind of reporting that comes from borrowers on projects, could tell you that PCRs would be handled in many cases by a clerk in a backroom, to satisfy the letter of what the Bank was asking. So we suddenly found ourselves with operational staff trying to do their best with these reports, which sometimes wasn't much, and with the backlog accumulating.

The decay meant that the whole system as conceived by McNamara—which was for comprehensive evaluation, with self-evaluation by operational staff as an essential component—was about to crumble.

I spent the next six and a half years pestering everybody about reducing the backlog and about improving the quality of PCRs. We had any number of task forces, reminders to

the Board, reminders to the president, and broken promises. After six and a half years, however, the system was back on track.

Our limited resources meant we had little left for evaluation studies. As you all do, I think that special studies are the best means for evaluators to distill and disseminate experience. I also think, like yourselves, that impact studies are absolutely vital if we're to find out the reality of what projects contribute, and if we are to understand the determinants of sustainability. Where I thought I could make a contribution here was to open the net for the selection of studies, presenting them in a two- or three-year program that would be large enough, despite the modest resources, to show some kind of balance among topics. By opening the net, we were, in effect, trying to mobilize ideas for the design of a program that would be relevant.

Of course, relevance is half of the solution to the dissemination problem. I knew situations where we were producing a study, but I was hearing from operational staff, "Well, that's very nice, but we don't do this type of operations any more." Or "This is all very fine, but it's old stuff, you know. It's old stuff." One reason they were right is that our studies took forever to produce. In the production of the study itself, you had all of the fishing for evidence built in. Instead, I was in favor of investing more time in preparation, in designing the approach to the study, so that the approach paper sent to the Board would be the end of a process rather than the beginning, and the study could be done faster. That was perhaps a dream, but it was a way to improve the quality of preparation and save time.

The studies were controversial if they had to be. If you deal with important and controversial issues, and you do a good job of it, you are present; you are not a mouse under the table, even though the inclination of management was understandably to push us down. For example, for a study on the environmental impact of projects in Brazil, I had to take several trips to Brazil to persuade the government to accept the study. I used the argument that "I can't afford not to do the study, because of the bad image it would give us." For the same reason, I told them, "You cannot afford not to allow us to do it." So the Brazilian government agreed and, because we were helping develop evaluation capacity in Brazil, they agreed to support the exercise with staff inputs. And the deal was clear from the beginning: at the end, we have our study, which we are fully responsible for, and you will have your study, if you want to, but we won't have a mixture of what you like and what we like. Otherwise, we are bound to have a diluted, a very diluted output." And that's the way it went. And I think the end result was a very good study.

We did other important studies: for example, reviews of adjustment lending; the four country studies started by my predecessors, on Sri Lanka and Pakistan, and later on Senegal and Tanzania; the rural development study, one of the major studies that shook the Bank; and the industrialization study, which looked at the early phases of industrialization in three Asian countries. This last also gave the Bank a shock. It hit the raw nerve of the Bank's ideology of the time, and so the management response to it was very negative.

As to the administration of the studies, I thought that the experience gained in managing these exercises should stay with us, so we should have our own staff be the study managers, as was the case with Gladstone Bonnicks on Tanzania.

In addition to the studies program, we were doing as many cluster audits as we possibly could; not only of series of projects, but also comparative audits of several projects. Part of the reason was to economize, but it was largely because comparison is so useful for drawing lessons of general application. Almost 60 percent of our performance audit output was in the form of clusters. That was just about the absolute limit. Many of these cluster audits were very interesting and enriching experiences.

The largest single output of OED (and the most costly) was the Annual Review of Evaluation Results. This also became the most controversial because—my luck being what it was—we began to show a decline in the performance of the Bank. And that decline continued until my last day, when the proportion of satisfactory operations had fallen to 63 percent.

When I arrived on the scene, the Annual Review was a headache because there were so many loose ends about the data that were not resolved until the very last minute of the exercise. Instead of merely massaging statistical information on projects completed in a given year, we transformed the exercise into a tool to present the important output of the department in the past year. Each year, we covered substantive topics. We then were able to treat the statistics differently, analyzing patterns by operations' year of approval, rather than by their year of completion, and thus to derive more interesting messages. We also incorporated the results of impact evaluations and introduced a focus on the sustainability of benefits.

All this made the Annual Review a much more interesting exercise. I was perusing them recently, fearing the worst, but found that they were actually pretty good pieces of work. The other thing we did was appoint the author each year, way ahead of time, to let that person manage the data retrieval and organization and also to do the conceptual work. We put some of our best people in charge of this exercise—which accounts, I think, for the good quality of the thing.

Dissemination and feedback

In the early days of evaluation the Bank was centralized, and dissemination and feedback processes could be managed effectively as part of its central systems. There were central systems for quality control, and central departments that had the responsibility to take the messages from OED, and in effect react to them and disseminate them. This was a very effective arrangement in a centralized operation such as we had then, and fine when you have a president, who is at center stage by definition, with a personal interest in seeing that evaluation results are effectively used.

But shortly after evaluation was introduced, the central quality control system of the Bank began to crumble. Operational responsibilities were decentralized, first to the Regions starting in 1972, and then to groups of countries in 1987, and the decentralized units became increasingly intolerant of any form of central oversight or advice.

This is where the participatory style of management became important: we needed the active participation and initiatives of all the staff of OED in the dissemination process. We needed them to talk at seminars, to interact with operational staff, to lecture at EDI (the precursor of the World Bank Institute), to develop courses for EDI, and to exchange views with the research people. To paraphrase Mao Tse Tung, we needed to be in operations, in EDI, and in research, as a fish is in water. Everybody had to be involved in dissemination. And we needed to back up their efforts by providing ready access to reliable evaluation data and easy retrieval of evaluation findings and lessons.

Of course, you need able people and responsible people to know how to do this selectively. The requests for help from operational staff were always there, and it was easy to find oneself doing other people's work. So we had to be selective, and this selectivity had to be administered not just by the director or division chiefs, but by the staff themselves. They had to make the most effective use of the little time they had available.

The other key thing we did in dissemination was to improve the presentation of our products. I pride myself in having reduced the Annual Review in 1986 from three volumes to one, and of a more modest size, to boot. No one missed the other two volumes! And we instituted a system where the Annual Review and the more important studies were professionally edited. At the beginning, in the confusion and the chaos of the Bank's pre-1987 reorganization, all kinds of good people, really very good editors, had been helping us for free. As the dust settled, we had to find the resources to continue the effort. Ram Chopra and I were finally able to persuade Rachel Weaving to join my office. She managed the recruitment and supervision of editors as well as editing a few things herself. She also took charge of developing a new concept, the *OED Précis*, where in one to four pages, we would highlight the findings of a study or an audit and draw lessons from experience. These were an immediate success.

Feedback, or the reinvestment into new operations of experience gained in past efforts, is an integral part of operational quality and controls and is the responsibility of operational staff and managers. The DGO reports annually to the Board and management on the effectiveness of the feedback process. The DGO and evaluation staff play a limited though important role in facilitating the process.

We promoted a few ideas. I tried, for example, getting the Regions to indicate what they had done to reinvest the lessons from earlier studies of value to their particular Regions, and to have their signed statements incorporated in our annual report. Some of these statements were quite good. Others were not. Where the wooden tongue of the bureaucrat was speaking loudly, it was very difficult to get uniform quality out of this effort, but we tried.

The other thing that we tried to do was to support best practice. The regional vice president for East Asia once told me, "I have an idea. What if I asked, I instructed, staff to have a section in the appraisal report that states how experience relevant to this project has been used and applied, how it has been reinvested in the design of this project?" I said, "Great." He said, "Would you help me?" I said, "Yes, we will help to the

extent we can,” although I was not welcomed in OED when I announced that we have this extra piece of work, which is to read appraisal reports. The LAC Region also took up this system.

So, in the decentralized system of the Bank, there are obviously myriad ways to proceed in dissemination and feedback.

Evaluation capacity development

As my predecessors Mervyn Weiner and Shiv Kapur also recognized, helping countries to develop their own capacity for evaluation is just as important a task for the DGO as promoting self-evaluation in the Bank. Both tasks are key to the effective process of learning and enhancing operational quality.

Why evaluation capacity development should be a priority was obvious to me. In an ideal world, borrowers would evaluate their development policies and investments, and donors would evaluate their contributions, as we were evaluating ours. Of course, trading experiences among participants would add value and impact to the work of each of them, and as more participants joined in, the value to each of them would grow—just as in a telecommunications system, the addition of one subscriber makes the system more useful for all the others. Participants should not be limited to the borrowers and ODA participants; there is considerable benefit in expanding the evaluation network to private sector actors, or agencies such as highway authorities in developed countries that have valuable experience to research and share.

Coming back to the real world, I learned quickly that I had to focus selectively on countries that showed real commitment and willingness to allocate resources to evaluation, and that would build and animate the decentralized systems necessary to do effective evaluation and to sustain their efforts.

I saw my role as to initiate the dialogue, wherever possible both with countries and other development institutions, and to follow up wherever prospects—that is, the commitment of the other party—made it worthwhile. To ensure follow up, I assigned the function initially to an adviser in my office, Robert van der Lugt, and when Robert left, I appointed as its manager Pablo Guerrero, whom I had met when he was heading the internal evaluation unit of the InterAmerican Development Bank. Together, we proceeded to involve the staff of OED whenever their work was relevant and the effort needed was affordable.

I would like to bring out two stories in this connection. I had been told that in Brazil, a central evaluation unit had been created a few years earlier but had ceased to exist. I was asked to meet with the Brazilian delegation to the Bank’s annual meeting and was informed that evaluation had been made mandatory in Brazil’s last constitutional reform and that our help would be welcome. I suggested that the Brazilians appoint a senior official and send him to spend time with us to develop the project with whatever help we could provide. So they did. And we—Pablo Guerrero in particular, and others in OED—developed a fruitful cooperation over several years. As part of his central duties, Dr. Lapa undertook the difficult task of ensuring meaningful comments on our evaluation

work in Brazil. When we undertook the complex evaluation of the environmental impact of Bank projects in Brazil, the government agreed to contribute to the effort.

An example from Colombia also illustrates the role that a good evaluation study, and collaboration with evaluation staff, can play in helping to launch evaluation in a country. At my first Bank annual meeting I met a minister from Colombia, who asked that we do an evaluation of the Bank's involvement in the power sector in Colombia. I agreed without hesitation, because I was already aware that there was a problem. Rene Ribí, civil engineer and a former power division chief, and Trevor Byer, who had led the Bank's Colombia energy assessment mission and was a star energy planner, undertook this work, helped by a financial analyst, and did, I think, a superb job. Then the Colombian Minister of Finance, Mr. Holmes, showed up in my office, on the occasion of a later annual meeting, and I said, "Mr. Minister, you are aware, I suppose, that one of your predecessors asked for an evaluation of the power sector in Colombia?" He said, "Yes, certainly, that's why I'm here." He leaned forward and said, "The study is very critical of the Bank." Then he smiled and said, "It is also very critical of Colombia." And so it was. He proceeded to tell me he was arranging a seminar in Colombia where all policymakers and other interested parties would be invited to discuss the needed reform of the power sector, and he asked for the participation of the authors of the OED study to help lead the discussions.

Then I said to the minister, "You have just seen the power of evaluation to turn things around. Would you consider creating an evaluation capacity of your own?" He said, "Not a bad idea, but how do I go about this?" I said, "You choose a high official whom you trust and you ask that person to think through the idea and propose a program to you. We would be delighted to welcome this person, and share our knowledge and the network of our friends to help in this endeavor."

And so he appointed Eduardo Wiesner, the retiring executive director for Colombia, who came to spend time with us, meeting with others in the evaluation community—including Eleanor Chelimsky, who was heading the evaluation arm of the U.S. General Accounting Office—and developing in the process a passion for evaluation. The minister approved his proposals and the project took off rapidly. A year or so later, Eleanor Chelimsky, Pablo Guerrero, and I were invited to a very interesting seminar on evaluation in Colombia where we were presented with four evaluation studies done by consultants and where, in effect, every official in the country, including the attorney general, wanted a piece of the action. (In Colombia as in Brazil, evaluation had been made mandatory in the last constitutional reform.)

Conclusion

In retrospect, referring back to Mr. Clausen's question, we did achieve impact in the Bank. I will cite only one example in this regard: the Wapenhans Review of the Bank's portfolio performance. This review would not have been launched were it not for OED's reporting of deteriorating portfolio performance in a series of Annual Reviews and studies. And it would not have been as credible and effective except for the material and analyses provided in a number of OED studies—on project supervision inter alia—or an

important review of economic analysis that I had asked management to undertake, and which Joanne Salop carried out so effectively with the help of our own Pablo Guerrero and a few others.

As my term ended, I felt I was leaving a stronger base for the future, and a lot more wanting to be done by my successors.

At this point, I think I would like some questions.

PARTICIPANT: I am Patrick Grasso, and I am the knowledge manager here. Mr. Rovani, one your points that caught my attention was about using advanced technologies, and particularly, I gather, for capturing our reports electronically. I think it's worth mentioning that the effort you made there is still paying off quite a number of years later, and is very useful. In fact, I don't think we could do our knowledge management job without it. It's been a very effective tool for us to have this kind of technology. What gave you the thought to begin to move in that direction?

MR. ROVANI: Well, two things. First, I was hit in the face when I came to OED by something that was called the Concordance. The Concordance was a huge book, manually, clerically, assembled by someone looking for key words. It was awful, and I stopped that work. Also I was struck by the fact that every year, when the Annual Review had to be prepared, the compilation of data had to start from scratch. Then, at this famous retreat I mentioned, there was interest in introducing personal computers.

So everybody got equipped, and we made a deal with the information technology department of the Bank to have one of their staff in residence to upgrade our data management. The support we got from them, particularly from Peter Gutterman, is what helped us develop a state-of-the-art data management system, and also to conceive the idea of the searchable evaluation text base and make it a reality. I was not an actor in this. I just was interested; I opened the door, and supported this thing. I went with Pablo to see what the Canadians had been doing, and they had a very good system. I think that had cost them several million dollars, and we couldn't spend \$100,000. So we waited a little, and when off-the-shelf technology had progressed a little further and costs had come down, OED and Peter were able to develop a "user-friendly" text retrieval system tailored to our needs.

PARTICIPANT: A little-known fact is that the rural development report was only the second Bank report published on the desktop. The precursor was the Uruguay Round report.

MR. ROVANI: I did not know this. It is another piece of evidence of the talent and innovative spirit in OED. Let me use this opportunity to make another point. Nowhere in the Bank is there under one roof as much diverse expertise, experience, and concern for quality as you have in this outfit. I think you are blessed with the fact that you cross the hallway and you meet somebody with a different background, different ideas. This is one of the great things about OED, and one of the things that made OED attractive to prospective staff members.

PARTICIPANT: You stressed the importance of impact evaluation for getting a better understanding of the determinants of sustainability. But in practical terms, the meaning of impact evaluation is rather elusive. One MDB defines it as four years after project completion. Here in OED, it's sort of a mix. And as you move the time horizon out, the evaluation becomes much more challenging, more like a research project, with difficulties of attribution, and extra expense. I'd be interested in your thoughts on this issue.

MR. ROVANI: I could not give you the expert answer you would like. A number of staff, I am sure, could answer this better than I can. But for me, two things matter: to be able to report meaningfully on the economic worth of projects well into the operational life of the project, and to do this cost-effectively—because ultimately this is what makes evaluation credible. Of course, the task of evaluating the impact of projects varies greatly with the type of project. Irrigation projects are probably the most difficult ones because of their long life and the many facets of their impact. In my time as DGO, we could not afford, without substantial borrower involvement, the costs of the long-range research involved. I thought we could do better by building into projects, from the beginning, the basic data needed for eventually evaluating the operational work. This said, we did produce meaningful evaluations of the impact of projects, with some surprising and worthwhile results. I can recall an impact evaluation of an irrigation project in Morocco, from which material was developed for EDI courses, and comparative evaluations of area cotton projects in several countries in West Africa.

PARTICIPANT: How do you see evaluation in the immediate operating context of today? The Bank is facing new kinds of criticism, new pressures, and new shaping forces for its future. Evaluation is being questioned and certainly being looked at more critically. So the question of accountability is key. Do you see this as a shaping force for the future of OED and its role in the Bank?

MR. ROVANI: Quite clearly, the Bank's evaluation function cannot escape the wholesale criticism being laid at the Bank's doorstep. The function is indeed part of the Bank; I know of no other course. Accountability will always be a key feature of the evaluation function. The issue is what you evaluate and how well you do it. I would say, stay the course, do the best job you can, and strive to continue to be independent.

I will touch on your second question in a different way. When I was briefing the executive directors in 1986, the United States executive director told me, "Mr. Rovani, evaluation is very important. We wish you to get to the top." And I said, "Mr. Keating, I promise I will climb. I don't know how far the top is." And my answer to you today is that we did a lot of climbing but, on my watch, evaluation did not get to the top.

PARTICIPANT: I just want to make a little postscript, Yves, if you'll excuse me, to say that for those who have known him as long as I have, Yves hasn't changed. The very first time I arrived in OED, in 1987, I think it was, he was holding a meeting, a staff meeting, and I got there late, and I inched open the door, and Yves was already in full flow, and I got in, and there was a little chair against the wall. I thought I'm going to make this, you see, and I was just sitting down in my little chair, having not made a sound, when he interrupted what he was saying, looked straight at me, and said "Nice of you to come."

PARTICIPANT: The latest review by the Bank's Quality Assurance Group shows that among the Bank projects entering the portfolio last year, 93 percent were satisfactory. If you put out a report showing that 93 percent of your projects are satisfactory, do you have a credibility problem? A lot of the measurement techniques that we use are looking at Bank *processes* and how effectively they are performed. And they miss in some key respects the ultimate *impact* the Bank is having on the beneficiary. Some elements of measurement are beginning to capture impact, and I think they should be part of the further discussion that really needs to take place.

PARTICIPANT: I wanted to make one comment that sums up my personal experience in OED. It is about what evaluation does to the evaluator. It's a tremendous discipline for training your mind for doing other work, where you have to look at a situation and discern the essence of what is determining that situation, in order to prescribe and make suggestions for policy and other things. I retired from OED seven years ago now, and since then I have been busy, but I must say that I always can tap into the reservoir that I accumulated by my experience in OED, because it's a tremendous training for the evaluator. Quite apart from that, you benefit from the quality people you work with in OED and the interchange that happens there.

MR. PICCIOTTO: Yves said two things today that particularly resonated with me. One is that his Bank career was always at the center (Projects, Central Projects Staff, Energy, Evaluation). V. S. Naipaul has a book called *Finding the Center*, and I think that is a job for evaluation, as Yves highlighted. Yves also talked about getting to the top. I think today we are still trying to get to the top. I don't think we'll ever make it, but I think that OED also still has both continuity and change, as Yves mentioned. So it has made us very proud to have you here today, Yves.

WB50377

O:\PUBs\History\Rovani 091802 rw draft.DOC
September 19, 2002 4:43 PM