

## **Monitoring and Evaluation for Cost-Effectiveness in Development Management**

Paul Clements<sup>1</sup>

### **1. Development Assistance Requires a High Analytic Standard**

In the Malawi Infrastructure Project, the World Bank planned to rehabilitate 1500 rural boreholes at a cost of \$4.4 million, with an estimated economic rate of return of 20%. At the project's Midterm Review, two years later, the rate of return was reduced to 14%, but the reasons for the reduction were not clear. The plan had anticipated that 85% of project benefits would come from the value of time the villagers saved that they would have spent collecting water, and 10% from the incremental water consumed. The Midterm Review estimated 31% of benefits from time savings, however, and 56% from incremental water consumed.<sup>2</sup> No reason was given for reducing the estimate for time savings or for increasing the value for water consumption.

The World Bank's Fourth Population Project in Kenya aimed to decrease Kenya's total fertility rate to six births per woman by improving family planning services.

---

<sup>1</sup> Corresponding author: Paul Clements, Department of Political Science, Western Michigan University, Kalamazoo, MI 49006, e-mail: [clements@wmich.edu](mailto:clements@wmich.edu).

<sup>2</sup> Carlos Alvarez, Ebenezer Aikins-Afful, Peter Pohland and Ashok Chakravarti, 1992, "Malawi Infrastructure Project: Mid-Term Review Report, September 1992," Lilongwe, Malawi: The World Bank, Appendix B. The economic analysis from the project plan comes from the Malawi Infrastructure Project's Staff Appraisal Report. I was given it upon agreeing not to reference it.

The project was approved in 1990, and Kenya's total fertility rate fell from 6.4 in 1989 to 5.4 in 1993. The project's Implementation Summary Reports consistently indicated that "All development objectives are expected to be substantially achieved," and a 1995 supervision report asserted that "The project development objectives have been fully met."<sup>3</sup> Project activities, however, mainly supporting the National Council for Population and Development, were largely unsuccessful, and in 1994 a large part of the project budget was reallocated to the fight against AIDS. There were many other development agencies with family planning projects in Kenya, some with much stronger performance. Documents for the Fourth Population Project do not explain how its development objectives were related to the activities it funded.

The World Bank's Water Supply and Sanitation Rehabilitation Project in Uganda aimed to rehabilitate the water and sewerage system in Kampala, the capital city, and in six other major towns. Its plan calculated a 20% economic rate of return based on incremental annual water sales of \$5.5 million from 1988 to 2014. The completion report estimated actual returns at 18% because water production in 1991 was 10-20% below expectations.<sup>4</sup> The project had indeed achieved its construction goals, but its efforts to strengthen the National Water and Sewerage Corporation (NWSC) had been undermined by the government's failure to raise water rates amidst hyperinflation and late payments on its water bills. The NWSC would have been unable to maintain the system without ongoing support, and

---

<sup>3</sup> World Bank, 1995, "Form 590," (unpublished project implementation summary for Third and Fourth Kenya Population Projects), Washington, DC: The World Bank.

<sup>4</sup> World Bank, 1991, "Project Completion Report: Uganda Water Supply and Sanitation Rehabilitation Project (credit 1510-UG)," Washington, DC: The World Bank, p. 30.

indeed by 1993, even with a major new project supporting the water company, it was once more operating in the red.<sup>5</sup>

These examples come from a blind selection of four World Bank projects that I studied for my doctoral dissertation.<sup>6</sup> What is remarkable about these inconsistencies—an economic analysis in a midterm review that does not follow from the one in the project plan, development objectives that do not reflect project activities, an economic rate of return that anticipates 23 additional years of water sales based only on the current state of the infrastructure—is that even though at least the second two are at face value analytically incorrect, they are presented as routine reporting information, with no attempt to hide them such as in obfuscating language. Indeed they reflect common analytic practice in the international development community, and this common practice reflects a structural problem of accountability.

I would like to argue that the tasks undertaken by the large multilateral and bilateral donor agencies require a particularly high analytic standard, but several incentives that influence development practice—political incentives for donor and recipient governments, organizational incentives for development agencies, and personal incentives for managers—have led to positive bias and analytic compromise. These incentives are “structural” in that they result from the pattern

---

<sup>5</sup> Paul Clements, 1996, *Development as if Impact Mattered: A Comparative Organizational Analysis of USAID, the World Bank and CARE based on case studies of projects in Africa*, doctoral dissertation for the Woodrow Wilson School of Public and International Affairs, Princeton University, p. 325.

<sup>6</sup> Along with four projects of the US Agency for International Development and four from CARE International, all located in Uganda, Kenya and Malawi. The projects were selected based on descriptions of less than a page with no information on results.

of the flow of resources inherent in development assistance. The problem therefore requires a structural solution, and this paper proposes a possible solution involving a dramatic improvement in the quality and consistency of project evaluations. We can be confident that such an improvement is possible, first, because the evaluation problem facing development agencies has determinate features with specific analytic implications, and second, because a similar structural problem has already been addressed in the management of public corporations.

Sooner or later, development assistance comes down to designing investments and managing projects. Unlike private sector investments, development projects aim not to make a profit, but to improve conditions for a beneficiary population—to reduce poverty, or to contribute to economic growth. There is no automatic feedback such as in sales figures, and no profit incentive to keep managers on task. Typically one needs to strengthen existing institutions or to build new ones, and/or to encourage beneficiaries to adopt new behaviors and to take on new challenges. Yet in the project environment there is likely to be weaker infrastructure, a less well-educated population, and more risk and uncertainty than in the environments facing most for-profit enterprises. Furthermore, in places that need development assistance one cannot assume that institutional partners will be competent and mission-oriented. These conditions in combination place particular demands on development managers. Project managers need to maintain a unified conception of the project, its unfolding activities, and its relations with its various stakeholders, a conception grounded in a view of its likely impacts. Donor agency officials need a conception of the relative merits of many actual and potential projects, and an analysis that turns problems on the horizon for developing countries into programmatic opportunities.

The central challenge in the management of development assistance is to maintain this kind of consciousness—this analytic perspective—among the corps of professional staff. Some might like to think that development can be achieved by getting governments to liberalize markets or by getting local participation in project management, and these may well be important tactics. Intuition suggests and experience teaches, however, that there can be no formula for successful development. Each investment presents a unique design and management challenge. There are two problems in maintaining the will and the capacity to address this challenge: an incentive problem and one we can call intellectual or cognitive. The key to solving both problems, or so I will argue, is strong evaluation.

## **2. But Accountability in Development Assistance is Weak**

### *2.1 Donor agencies are responsible for the success of their projects*

According to the World Bank's procurement guidelines, "The responsibility for the execution of the project, and therefore for the award and administration of contracts under the project, rests with the Borrower."<sup>7</sup> One might think that a development loan to a government is like a business loan to an entrepreneur. The donor agency makes the loan, but it is entirely the responsibility of the borrower government to spend the money. Whether a government manages its projects well or poorly, one might imagine, is primarily its own affair, with the donor providing technical assistance upon request. We know, of course, that this image is incorrect—donor agencies typically have the predominant influence over project design, and substantial influence over project administration—but it is useful to

---

<sup>7</sup> The World Bank, 1985, "Guidelines: Procurement under IBRD Loans and IDA Credits," Washington, DC: The World Bank, pp. 5-6.

recall why this is so. One reason is parallel to a private bank's prudential interest in the management of its loans. As the World Bank's Articles of Agreement state,

The Bank shall make arrangements to ensure that the proceeds of any loan are used only for the purposes for which the loan was granted, with due attention to considerations of economy and efficiency and without regard to political or other non-economic influences or considerations.<sup>8</sup>

The Bank wants to be repaid, and it also has an interest in promoting economic growth and enhancing well-being in borrower countries, so it may take pains to see that its loans are well spent. Many loans are to governments with limited bureaucratic capacity in countries with inconsistent management standards, so the Bank must retain enough control to ensure that the projects it supports are properly administered.

By this logic we would expect relationships with bureaucratically stronger governments to be closer to the private sector model, and indeed some governments with coherent industrial strategies (consider South Korea in the 1970s) have succeeded in using Bank loans very much for their own purposes.<sup>9</sup> Many development loans, however, are for projects at the edge of the borrower's

---

<sup>8</sup> International Bank for Reconstruction and Development, 1991, "International Bank for Reconstruction and Development: Articles of Agreement (As amended effective February 16, 1989)," Washington, DC: The World Bank, p. 7.

<sup>9</sup> See e.g. Mahn-Je Kim, 1997, "The Republic of Korea's Successful Economic Development and the World Bank," in Devesh Kapur, John P. Lewis, and Richard Webb, ed., *The World Bank: Its First Half Century*, Volume Two, Washington, DC: Brookings Institution Press, pp. 17-48.

frontier of technological competence, and the Bank (like other donor agencies) is a repository of expertise in the sectors it supports. The Bank also has demanding requirements for project proposals, and many governments have been unable independently to prepare proposals that the Bank could accept, particularly in earlier years when patterns of Bank-borrower relations were established. Therefore the Bank has generally taken primary responsibility for designing the projects it funds,<sup>10</sup> and the responsibility that comes with authorship cannot be lightly abandoned during implementation.

A second reason that donor agencies take an interest in how their funds are spent is that donor funds come from (or are guaranteed by) governments, and, the Bank's Articles of Agreement notwithstanding, governments do not release funds without taking an interest in their disposition. On one hand this political logic reinforces the prudential logic discussed above. Donor governments want their funds to contribute to the borrower's development, so they insist that donor agencies take responsibility for project results. Foreign aid is also, on the other hand, enmeshed in donor governments' general promotion of their foreign policy agendas.<sup>11</sup> It matters that the United States is not indifferent as to whether and when the World Bank will make loans to Cuba, and World Bank loans to Côte d'Ivoire have been subject to particular influence from France, the country's former colonial master.<sup>12</sup> Bilateral aid is even more closely linked to donor government interests than aid

---

<sup>10</sup> Warren C. Baum and Stokes M. Tolbert, 1985, *Investing in Development: Lessons of World Bank Experience*, New York: Oxford University Press for the World Bank, p. 353.

<sup>11</sup> Paul Clements, 1999, "Informational Standards in Development Agency Management," *World Development* 27:8, 1359-1381, p. 1360.

<sup>12</sup> Jacques Pégatiéan and Bakary Ouayogode, 1997, "The World Bank and Côte D'Ivoire," in Kapur, Lewis and Webb, ed., pp. 109-160.

through multilateral institutions. Not only from the donor side but from that of recipient governments too, the parameters of development spending cannot be understood merely in terms of the requirements for maximizing development impacts.

As intermediaries between donor and recipient governments, donor agencies are required to take more responsibility than private banks for managing the loans they make. The analogy with the private sector breaks down even further, however, when we consider the incentives governing a donor agency's management of its portfolio. The main cause for different incentive structures between donor agencies and private banks, of course, arises from differential exposure to financial risk. With private loans, the borrower and often the lender suffer a financial loss if the investment fails. With most development projects, by contrast, neither the donor nor the implementing agency faces a financial risk if impacts are disappointing. For projects funded by loans it is the borrower government, typically the treasury, that is responsible for payments. But the treasury seldom has control over individual development projects.

## *2.2 The usual watchdogs are not there to hold donor agencies accountable*

The structural conditions of development assistance, therefore, create an accountability problem. Donor agencies have control over development monies but they face no financial liability for poor results (and no financial gain when impacts are strong). In this context their orientation to their task will depend largely on the demands and constraints routinely placed on them by other agents in their organizational environment, on the individual and corporate interests of their leaders and employees, and on the mechanisms of accountability that are institutionally (“artificially”) established.

In regard to external agents, Wenar notes that there has been a “historical deficiency in external accountability” for donor agencies.

Aid organizations have evolved to a great extent unchecked by the four major checking mechanisms on bureaucratic organizations. These four mechanisms are democratic politics, regulatory oversight, press scrutiny, and academic review.<sup>13</sup>

The electorates in donor countries want to believe that aid is helping poor people, but democratic politics also leads to pressures on donor agencies to support the agendas of well-organized interest groups.<sup>14</sup> Some promote humanitarian and progressive agendas, but others have aims that create tensions with development goals. Generally, since the intended beneficiaries of aid cannot vote in donor country elections, the reliability of democratic politics as a source of accountability is limited. There has been significant regulatory oversight aiming to ensure that aid funds are not fraudulently spent, but external oversight of project effectiveness faces major practical hurdles. Aid projects are so widely dispersed, and the period between when monies are spent and when their results transpire is typically so substantial that effective oversight would require major bureaucratic capacity. Responsibility for project evaluation, however, has normally rested with the donor agencies themselves. This clearly leads to conflicts of interest, and it is the aim of this paper to suggest how these conflicts could be, if not removed, at least substantially ameliorated. Donor agencies have not, in any case, been subject to

---

<sup>13</sup> Leif Wenar, 2003, “What we owe to distant others,” *Politics, Philosophy & Economics*, 2:3, 283-304, p. 296.

<sup>14</sup> For example, American farmers have influenced U.S. food aid programs, which are overseen by the U.S. Department of Agriculture.

significant external accountability by way of regulatory oversight. Press scrutiny and particularly academic review, in contrast, have been significant sources of accountability, and academic studies have contributed to many foreign aid reforms. Given the strength of the political and bureaucratic interests that drive the programming of aid, however, and the above-noted dispersal of aid projects, scholars and journalists can only be expected to hold aid agencies accountable in a limited and inconsistent manner. Also, they are largely dependent, for information on aid operations on the donor agencies themselves.

Few who have spent much time with development agency personnel can doubt their generally admirable commitment to development goals, and the reforms this paper will propose depend heavily on the personnel's sustained interest in professionalism and effectiveness. Their behavior is also influenced, however, by their individual and corporate interests, and these interests take shape in the specific task environments that they face in their home offices and in the field. There are two aspects of the way their interests come to be constructed that are particularly relevant to the problem of accountability. First and most obviously, while institutional norms require donor agencies to maintain the appearance of a coherent system of responsibility for results, their institutional relationships require them to maintain the appearance that their operations are generally successful. They must evaluate, but it serves their individual and corporate interests if evaluation results are generally positive (or at least not often terribly negative). Since donor agencies have generally controlled their own evaluation systems, they have had the opportunity to design these systems in such a way that they would tend to reflect positively on the agencies themselves. Second, due in part to the long time span between the commitment of funds and the evaluation of results, internal personnel evaluations have tended to focus on variables only loosely

correlated with good results, and sometimes on variables that conflict with good practice.

### *2.3 Lacking secure accountability for results, other less relevant criteria inform resource allocation decisions*

We will later consider some of the approaches donor agencies have taken to evaluation below. For the purposes of understanding the accountability problem in development assistance, it is enough for now to note that donor agencies have controlled their own evaluation systems. In the context of the general deficiency in external accountability, the priorities that have been enforced within donor agencies take on particular significance. Perhaps the most longstanding and sustained critique of donor agencies' internal operations involves the imperative to "move money."

The classic account of the "money-moving syndrome" is Tendler's *Inside Foreign Aid*.<sup>15</sup> Focusing on the U.S. Agency for International Development (USAID) and the World Bank, Tendler identifies a "pressure to commit resources that is exerted on a donor organization from within and without," and finds that "standards of individual employee performance ... place high priority on the ability to move money."<sup>16</sup> In the context of her organizational analysis, she gives several examples of aid officials knowingly supporting weak projects in order to reach spending targets.<sup>17</sup> Tendler also finds, reinforcing the present argument about evaluation,

---

<sup>15</sup> Judith Tendler, 1975, *Inside Foreign Aid*, Baltimore: Johns Hopkins University Press.

<sup>16</sup> *Ibid.*, p. 88.

<sup>17</sup> *Ibid.*, p. 88-96.

that in a political environment often hostile to foreign assistance, aid officials learned to self-censor reports that could provide ammunition for critics.

For writing what he considered a straightforward description of a problem or a balanced evaluation of a project, an AID technician might be remonstrated with, “What would Congress or the GAO [General Accounting Office] say if they got hold of that!?” ... Words were toned down, thoughts were twisted, and arguments were left out, all in order to alleviate the uncomfortable feeling of responsibility for possible betrayal. ... Such a situation must have resulted in a certain atrophy of the capacity for written communication – and, inevitably, for all communication through language.<sup>18</sup>

The World Bank typically required economic analysis of proposed projects, but Tendler found that many ostensibly economic projects were selected by non-economic criteria.<sup>19</sup> Much of the economic analysis that was carried out amounted to a “*post hoc* rationalization of decisions already taken.”<sup>20</sup>

While Tendler offers several political and organizational reasons to explain the money-moving imperative, I would like to emphasize what is absent from the organizational culture she describes. We do not find a sustained effort to consider how development funds can be employed to maximize their contribution to development. In such an environment we might expect well-intentioned professionals, once they win some organizational power, to act like policy

---

<sup>18</sup> Ibid., p. 51.

<sup>19</sup> Ibid., p. 93.

<sup>20</sup> Ibid., p. 95.

entrepreneurs, promoting their individual conception of a good development agenda in large measure despite the prevailing incentives. We might expect segments of a donor agency that have strong external allies to develop coherent agendas that they can implement themselves, as I believe reproductive health professionals at USAID have done. What we cannot expect, however, is that organizational decisions will routinely be taken on the basis of expected impacts.

The World Bank's "project approval culture" was recognized in its internal 1992 study, "Effective Implementation: Key to Development Impact" (popularly called the Wapenhans Report). The report cites a "pervasive preoccupation with new lending,"<sup>21</sup> in part because "signals from senior management are consistently seen by staff to focus on lending targets rather than results on the ground,"<sup>22</sup> noting also that "[t]he methodology for project performance rating is deficient; it lacks objective criteria and transparency."<sup>23</sup> Although the report describes the Bank's evaluation system as "independent and robust," it finds that "[l]ittle is done to ascertain the actual flow of benefits or to evaluate the sustainability of projects during their operational phase."<sup>24</sup>

Since the appearance of the Wapenhans Report, the Bank has moved increasingly to spending modalities that further dilute accountability for results. The two kinds of programs that have become most central to Bank strategies particularly in lower income countries are adjustment loans of various kinds (structural, sectoral) and

---

<sup>21</sup> Portfolio Management Task Force, 1992, "Effective Implementation: Key to Development Impact," Washington, DC: The World Bank, p. iii.

<sup>22</sup> Ibid., p. 23.

<sup>23</sup> Ibid., p. iv.

<sup>24</sup> Ibid.

Poverty Reduction Strategy Papers (PRSPs).<sup>25</sup> Adjustment loans require borrowers to adopt free market reforms in order to better align economic incentives with development goals. They tend to operate on a wider scale than traditional projects, with more diffuse impacts. There is often a feeling that they are imposed, as the government receives the loan for policy changes it presumably otherwise would not have made, and they are often implemented only partially and inconsistently. These factors make them harder to evaluate. Poverty Reduction Strategy Papers typically push a larger part of the responsibility for evaluation onto the borrower government, and it seems that their more participatory approach to policy formation and implementation is intended to substitute, to some extent, for rigorous agency evaluation. They ask the government, as part of the process of generating a poverty reduction strategy, to identify a set of indicators for measuring the strategy's impacts. If the World Bank has had such a hard time ascertaining the level and sustainability of impacts from its own portfolio, however, it is questionable whether governments of low-income countries will be able to do much better.

### **3. Independent and Consistent Evaluation Can Improve Accountability and Learning in Development Assistance**

#### *3.1 The basic idea of the proposed evaluation approach*

The problems discussed above present formidable obstacles to maintaining accountability in foreign assistance on the basis of program and project results. We should recall, however, what is at stake. In the absence of meaningful accountability there is little to counter-balance the pressures for aid resources to

---

<sup>25</sup> David Craig and Doug Porter, 2003, "Poverty Reduction Strategy Papers: A New Convergence," *World Development* 31:1, 53-69.

support political interests of donor and recipient governments, organizational interests of donor and implementing agencies, and personal interests of management stakeholders. The inconsistency and mixed reliability of evaluations have also undermined learning from experience, so the aid community has been slower than it would otherwise have been to identify successful strategies and to modify or abandon weak ones.<sup>26</sup>

One way to address the historical deficit of external accountability, for example to push the focus of management attention forward from moving money to achieving results, and to improve the incentive and the capacity to manage for impacts, is to

---

<sup>26</sup> Indeed despite development agencies consistently reporting positive results from their overall operations, there have been persistent doubts about the basic effectiveness of development assistance at improving economic and/or social conditions in recipient countries. In their comprehensive 1994 review of foreign aid on the basis of donor agency documents, *Does Aid Work? Report to an Intergovernmental Task Force*, second edition, Oxford, UK: Oxford University Press, Robert Cassen and associates find that most project achieve most of their objectives and/or achieve respectable economic rates of return. A series of cross-country econometric studies, however, have failed to find evidence of positive impacts from foreign aid. These include Paul Mosley, John Hudson, and Sara Horrell, 1987, "Aid, the public sector and the market in less developed countries," *The Economic Journal*, 97:387, 616-641; P. Boone, 1996, "Politics and the effectiveness of foreign aid," *European Economic Review*, 40:2, 289-329; and Craig Burnside and David Dollar, 2000, "Aid, policies and growth," *The American Economic Review*, 90:4, 847-868. These results are reviewed and contested, however, in a recent paper by Michael Clemens, Steven Radelet and Rikhil Bhavnani, 2004, "Counting chickens when they hatch: The short-term effect of aid on growth," Center for Global Development Working Paper 44, <http://www.cgdev.org/Publications/?PubID=130>. Clemens, Radelet and Bhavnani find positive country-level economic impacts from aid based on cross-country econometric studies focusing on the approximately 53% of aid that one would expect to yield short term economic impacts.

institute independent and consistent evaluations of the impacts and cost-effectiveness of donor-funded projects.<sup>27</sup> This would mark a significant departure from existing practice, so I first explain the concept, then suggest how it could be implemented, and finally consider how it compares to established evaluation approaches.

For both accountability and learning, the appropriate frame of reference is not the individual project but the donor agency's overall portfolio, and, for learning, the world-wide distribution of similar and near-similar projects. The donor agency's question is how to allocate its resources so as to maximize the impacts of its overall portfolio. The project planner or manager's question is, in light of relevant features of the beneficiary population and the project environment, (and, for managers, in light of how the project is unfolding,) how to configure the project design so as to maximize impacts. In both cases the relevant conception of impacts is one that supports comparisons among projects.

Both accountability and learning, for donor agencies, start from impacts and then work backwards in causality. They start, for example, with strong or weak results, and while accountability uses the discovery of causes to allocate responsibility, learning uses it to construct lessons based on schemes of similarities (so the lessons can be applied to other contexts). In this way rewards and sanctions can be allocated based on contributions to impacts and managers can gain a feel for what is likely to work in a new situation.

---

<sup>27</sup> Impacts are defined as changes in conditions of the beneficiary population due to the project, i.e. compared to the situation one would expect in the project's absence (compared to the counterfactual).

Now this logic may sound quite general. It applies with particular force to large donor agencies because their other sources of accountability are so sparse, the tasks they undertake are so costly and complex, and the contexts in which they work are so often difficult and demanding. Accountability and learning bear a heavier burden than in other contexts. This is why projects should be evaluated not by the extent to which they achieve their individual, idiosyncratic objectives, but in terms of impacts expressed in consistent and comparable units. An evaluation's units of analysis establish a perspective or orientation in terms of which the project and its activities come to be understood. In order to establish a consistent orientation across a donor agency's portfolio, therefore, (or, for a given type of project, across countries and donor agencies,) evaluations should be conducted in consistent units. Accountability requires consistent units precisely because the appropriate frame of reference for accountability is the donor agency's overall portfolio.

### *3.2 Comparing projects in terms of cost-effectiveness*

I would like to suggest that the unit that provides the appropriate frame of reference for donor agency evaluations is cost-effectiveness. Cost-effectiveness aims to achieve the greatest development impacts from the available resources. We can compare evaluation in terms of cost-effectiveness with two other approaches that donor agencies have often used. Bilateral donor agencies have typically evaluated their projects in terms of how far they have achieved their stated objectives,<sup>28</sup> and the multilateral development banks (such as the World Bank) have historically evaluated most of their projects in terms of their economic rates of return.

---

<sup>28</sup> This is the 'logical framework' approach.

When projects are evaluated in terms of their objectives, comparisons among projects are likely to be misleading. Some projects have ambitious objectives while others are more modest, so a project that achieves a minority of its objectives can clearly be superior to another that achieves most of its own aims. Also, the criterion “to achieve objectives” bears no clear relation to costs. If this criterion is taken as the basis for accountability, it establishes an incentive to set easy targets and/or to over-budget.

A project’s economic rate of return (ERR) expresses the relation between the sum of the economic value of its benefits and its costs.<sup>29</sup> It can also be described as the return on the investment, and the World Bank has typically expected an ERR of 10% or higher from its projects in the economic sectors. The difference between cost-effectiveness as I am defining it and an ERR is that an ERR measures benefits in terms of their economic values (ideally at competitive market prices) while cost-effectiveness measures benefits in terms of the donor’s willingness to pay for them. The economic analysis of projects typically does not include improvements in health or education, and benefits to the poor generally count the same as benefits to households that are already well off.<sup>30</sup> For an evaluation system based on cost-effectiveness, a donor would need to establish a table of values specifying how much it is willing to pay for each of the various benefits that it expects from its projects, including those that may be expressed in qualitative as well as in

---

<sup>29</sup> Specifically, the ERR is the discount rate at which the discounted sum of benefits minus costs is equal to zero.

<sup>30</sup> J. Price Gittinger, 1982, *Economic Analysis of Agricultural Projects*, second edition, Baltimore, MD: Johns Hopkins University Press.

quantitative terms. In this way a basis would be established for comparing, for example, primary health care and agricultural extension projects.<sup>31</sup>

### *3.3 The proposed evaluation approach in practice*

In practice, the proposed evaluation approach would work like this. At the completion of a project, an evaluator estimates the sum of impacts up to the present point in time and the magnitude of impacts that can be expected in the future that can be attributed to the project. The project's total impacts are compared to its costs based on the donor's table of values, and on this basis the evaluator estimates the project's cost-effectiveness. To estimate impacts the evaluator lists the relevant impacts from a project of the present type and design,<sup>32</sup> and carries out the appropriate quantitative and/or qualitative analysis of the project's activities and their results. The evaluator assigns each form of impact a numeric value and/or a qualitative rating based on his/her judgment of the project's likely effects in the respective areas over the lifetime of the project's influence. The impacts are summed together with appropriate weights from the table of values and compared to costs, and on this basis the evaluator estimates the project's likely cost-effectiveness, for example, on a scale from one to six, with one representing failure and six indicating excellence (see Table 1). The evaluator also notes her degree of confidence in the cost-effectiveness score, and, if her confidence is moderate or low, indicates the range of cost-effectiveness scores in which she is confident that the true value of likely impacts lies. In this case she also specifies the additional

---

<sup>31</sup> This approach to establishing the value of project impacts is described in Paul Clements, 1995, "A Poverty Oriented Cost-Benefit Approach to the Analysis of Development Projects," *World Development*, 23:4, 577-592.

<sup>32</sup> These may be found in the project plan.

information that could plausibly be collected that would allow a more precise estimate. The estimate of the project’s cost-effectiveness anchors the evaluator’s analysis of the project’s design and implementation. All four components—the analyses of impacts, cost-effectiveness, design and implementation—serve as a basic unit to support accountability and learning within the project and the donor agency and across the development community.

Table 1: Scale of Cost-Effectiveness

Economic Rate of Return	Degree of Cost-Effectiveness	Interpretation
30% and above	6	Excellent
20% - 29.9%	5	Very good
10% - 19.9%	4	Good
5% - 9.9%	3	Acceptable
0% - 4.9%	2	Disappointing
Below 0%	1	Failure

### 3.4 An evaluation association to address bias and inconsistency

While evaluations in terms of cost-effectiveness may support learning and accountability, there are (at least) three problems with the proposed evaluation approach. First it does not address the bias arising from donor agency control of the evaluation process. Second the estimates of impacts that it requires, including impacts in the future, present methodological challenges. There are no widely accepted methodologies for some of the required impact estimates (e.g. for reproductive health and AIDS education projects). Third, even where accepted methodologies are available (such as economic cost-benefit analysis, for economic projects), the results are often highly sensitive to minor changes in assumptions. When evaluations are contracted out on a project by project basis, different

assumptions are likely to be applied to different evaluations, undermining the validity of comparing and aggregating their results.

There are strong parallels between the conditions for the problem of bias and inconsistency in the evaluation of foreign aid and conditions facing public corporations in the management of their internal finances. Stockholders want corporation managers to employ the corporation's resources in such a way as to maximize profits, but managers face incentives to use the resources for their private purposes. There are elaborate rules governing how managers may appropriately use a corporation's resources, and it is the task of accountants and auditors to ensure that these rules are followed. As with evaluators in foreign aid, however, accountants and auditors are employed by the very managers whom they are expected to hold accountable. In order to protect their independence from management, and to ensure that they have mastered the relevant techniques, accountants and auditors have established professional associations. These associations establish qualifications that their members must achieve and rules that they must follow in order to retain professional membership. It is these rules that are the source of accountants' and auditors' independence from corporate management. Although independence is not maintained perfectly, the consequences of major lapses can be quite severe, as evidenced by the collapse of the international accounting firm Arthur Anderson after the accounts it managed for the Enron Corporation were found to be unreliable.

Amartya Sen lists transparency guarantees as one of five sets of instrumental freedoms that contribute to people's overall freedom to live the way they would like to live. Society depends for its operations on some basic presumption of trust, which depends in turn on guarantees of disclosure and lucidity, especially in relations involving large and complex organizations. Sen points out that where

these guarantees are weak, as they appear to be in foreign aid, society is vulnerable to corruption, financial irresponsibility, and underhand dealings.<sup>33</sup> A professional association of development project evaluators could play a role guaranteeing disclosure and lucidity in the management of international development assistance similar to that of associations of accountants in the management of public corporations. Such an association could also address the problems of estimating project impacts and of comparing impacts in common units.

In order to address the problems of bias and inconsistency, such an association would need the same structural features as associations of accountants—qualifications for membership, a set of rules and standards governing how evaluations are to be carried out, and procedures for expelling members who fail to uphold the standards.

In order to ensure that impacts are estimated and then compared in common units, the association would establish a constitutional principle asserting that each end-of-project evaluation conducted by its members would estimate the project's impacts and cost-effectiveness to the best of the evaluator's ability. One task in establishing the association would be to work out impact assessment approaches for different kinds of projects. The technical difficulties in estimating project impacts are objective problems, so it is possible to identify principles and practices for addressing them. An evaluation association would provide a forum for identifying better evaluation approaches and for ensuring consistency in their application. Over time, as its members gained experience, these approaches would be refined.

---

<sup>33</sup> Amartya Sen, 1999, *Development As Freedom*, New York: Knopf Publishers, pp. 38-40.

There are dozens of donor agencies and many thousands of implementing agencies in the development assistance community, and each agency has its own management culture and approaches. The universe for the evaluation association's operations would be the development assistance community overall, and it would support learning about better practices on the basis of the type of project throughout this community. Each evaluation completed by a member of the association would be indexed and saved in an online repository, which would be accessible to the entire development community. Since each evaluation estimates the project's cost-effectiveness, it would be a simple operation for someone planning, say, an urban water project, to review the approaches of the five to ten most cost-effective water projects in similar environments.

#### **4. Monitoring and Evaluation (M&E) for Cost-Effectiveness Compared to Other M&E Approaches**

##### *4.1 Monitoring and evaluation for empowerment*

To explain M&E for cost-effectiveness it is useful to compare it with other evaluation approaches. The strongest challenge to standard approaches to aid evaluation in the last two decades has involved the elaboration and application of participatory approaches.<sup>34</sup> These have aimed to involve beneficiary populations in project management, to assist them in taking responsibility for improving their own conditions and to incorporate them in more democratic processes of development decision making. Authors such as Korten and Chambers,<sup>35</sup> whom

---

<sup>34</sup> See e.g. B. E. Cracknell, 2000, *Evaluating Development Aid: Issues, Problems and Solutions*, Thousand Oaks, CA: Sage Publications.

<sup>35</sup> David C. Korten, 1980, "Community Organization and Rural Development: A Learning Process Approach," *Public Administration Review*, 40:5, 480-511; Robert Chambers, 1994, "The

Bond and Hulme describe as “purists,”<sup>36</sup> have sought to reorient the development enterprise to support the goal of empowerment. They have promoted an approach I call “M&E for empowerment” because it emphasizes learning at the local level, seeking to empower project beneficiaries by involving them in the evaluation process. While M&E for cost-effectiveness appreciates that empowerment is an important development goal, it identifies the locus for the primary learning that evaluation should support among those who are responsible for resource allocation decisions. Donor agency officials are the primary audience for aid evaluation because they exercise primary control over these resources. It turns out, however, that the form of evaluation that can best inform these officials will also best inform officials of developing country governments, project managers, and the overall development community, as well as, with some additional synthesis, the legislatures that appropriate aid budgets.

Evaluation and empowerment goals overlap in their management implications, and empowerment was certainly neglected by the development community prior to the mid-1970s. In many instances participatory strategies are more cost-effective than projects based on so-called blueprint approaches, so M&E for cost-effectiveness would promote participation in these cases. M&E for cost-effectiveness does not assume, however, that participatory approaches are right for all projects. The empowerment of project beneficiaries is interesting from an analytic viewpoint,

---

Origins and Practice of Participatory Rural Appraisal,” *World Development*, 22:7, 953-969; Robert Chambers, 1994, “Participatory Rural Appraisal (PRA): Analysis of Experience,” *World Development* 22:9, 1253-1268; Robert Chambers, 1994, “Participatory Rural Appraisal (PRA): Challenges, Potentials and Paradigm,” *World Development*, 22:10, 1437-1454.

<sup>36</sup> R. Bond and D Hulme, 1999, “Process Approach to Development: Theory and Sri Lankan Practice,” *World Development*, 27:8, 1339-1358, p. 1340.

because it can be seen both as a means to improving project designs and as an end in itself. For this reason M&E for cost-effectiveness views empowerment in a dual light. As a means, M&E for cost-effectiveness considers empowerment like any other possible means to be considered in program design. As an end, M&E for cost-effectiveness considers successful empowerment to be a benefit which must be valued and counted along with other benefits in the assessment of a project's cost-effectiveness. Under M&E for cost-effectiveness both more and less participatory projects are considered within the same evaluation framework.

#### *4.2 Monitoring and evaluation for truth*

It is possible that a great practical barrier to useful evaluation arises from some of those most knowledgeable of and committed to evaluation as a science. It has been common practice to begin discussions of aid evaluation methodology with the experimental method of the natural sciences,<sup>37</sup> and to present the various evaluation methods as, in effect, more or less imperfect approximations to randomized and controlled double-blind experiments. This approach often uses household surveys that measure conditions that a project seeks to influence, so that through appropriate comparisons changes attributable to the intervention can be identified in a statistically rigorous manner. I call it "M&E for truth" because it emphasizes making statistically defensible measurements of project impacts. This approach is right to insist that projects should be assessed primarily on the basis of their impacts, and that impacts should be understood as changes in the conditions

---

<sup>37</sup> E.g. Dennis J. Casley and Denis A. Lury, 1982, *Monitoring and Evaluation of Agriculture and Rural Development Projects*, Baltimore, MD: The Johns Hopkins University Press; Judy L. Baker, 2000, *Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners*, Washington, DC: The World Bank.

of the population compared to what would be expected in the project's absence (in evaluation jargon, as compared to the counterfactual). It is arguable, however, that in its orientation to statistical rigor it has established a "gold standard" that many evaluators are all too quick to disavow. Only a very small proportion of project evaluations present statistically rigorous impact estimates, and evaluations that do not often use the demanding requirements of statistical rigor as an excuse not to address the question of impacts at all. Also, evaluations that adhere rigorously to the maxims of statistical rigor seldom estimate the future impacts that can be attributed to a project.

Monitoring and evaluation for cost-effectiveness is methodologically eclectic in its effort to reach reliable judgments of cost-effectiveness. It is grounded not in the first instance in the scientific method, but in the causal models of change inherent in project designs. Each project design presents a hypothesis as to the changes in beneficiary conditions that can be expected from the actions the project undertakes. It is the evaluator's task to analyze how this hypothesis has unfolded, and on this basis to estimate the quantity of benefits that beneficiaries are likely to realize. A given project is taken as an instance of a project of its type, so impact estimates for other similar projects serve as a first approximation for the benefits that may be anticipated from the present project. Evaluators locate the present project along the continuum established by other similar projects based on how its design hypothesis unfolded as compared to theirs. Clearly, baseline surveys often provide critical information for estimating impacts, and statistical methodology of course provides central criteria for their analysis. As suggested above, M&E for cost-effectiveness employs participatory methodologies in many instances to elicit beneficiaries' judgments of the significance of project outputs. The evaluator's final estimate of a

project's impacts and cost-effectiveness, however, is based on triangulation taking into account all the forms of information we have so far considered.

### **5. A More Analytic Development Assistance Community**

Although I have described the proposed approach as monitoring and evaluation for cost-effectiveness, the discussion up to this point has focused on evaluation only. For the proposed evaluation approach to address the accountability problem in foreign aid, however, it is essential that planners and managers should know in advance that upon its completion there will be an independent evaluation of their project's impacts and cost-effectiveness. The development assistance community as we know it has evolved under conditions of inconsistent and often limited and biased evaluation, but one could anticipate, if the proposed evaluation approach were implemented, that its effects would gradually suffuse through all stages of project planning and implementation. Project planners would soon learn to include an estimate of cost-effectiveness in their project designs, and to establish monitoring systems that would track the relevant impacts (or their main contributing factors) through the life of the project. It would soon be taken for granted that when targets or systems for estimating impacts are altered during project implementation, the reasons for these changes should be clearly documented. The development assistance community would soon learn what outcomes need to be tracked for different kinds of projects to inform subsequent impact estimates.

While many individuals and groups in the contemporary development community are engaged in promoting development agendas of their own conception, the proposed reforms would enhance the experience of development work as a cooperative venture with shared goals. Development professionals would become

more confident that others would endorse their sound justifications for their management strategies, and management strategies would be more rigorously grounded in expected impacts. Members of the development community generally would become more conscious of the pathways by which their actions contribute to improvements in beneficiary conditions, and their shared concern for efficiency would be enhanced. The development community would be quicker to identify and to adopt more successful strategies. Although I believe that outright corruption on any significant scale is uncommon in the development community,<sup>38</sup> the higher analytic standard that the proposed reforms would bring about would reduce corruption even further.

The general public has tended to be fairly skeptical of foreign aid, and the management standards described in this article provide good reasons for skepticism. The proposed reforms would make it straightforward to aggregate project impacts, for example, by country or by agency. The tax-paying public would receive better information on the consequences of foreign aid, and they would have better grounds for confidence in its integrity. In due course this could be expected to increase the generosity of citizens in the wealthier countries towards people in need.

---

<sup>38</sup> I found large scale corruption in only one of the 12 projects in the sample for my dissertation.