

#25

100-25247

THE IMPLICATIONS OF THE PVO IMPACT STUDY  
FOR AID'S EVALUATION ACTIVITIES

By Dr. Elliott R. Morss\*

June 1979

Office of Evaluation Working Paper Number 25

\* Work for this paper was sponsored by the Office of Evaluation, PPC, USAID. The conclusions reached are those of the author, who is Director of Research, Development Alternatives, Inc., 1823 Jefferson Place, N.W., Washington, D.C. The author wishes to acknowledge the confidence in him evidenced by PPC in asking him to assess a study produced by his own firm.

The Implications of the PVO Impact Study  
for AID's Evaluation Activities

by Dr. Elliott R. Morss

Introduction

Development Alternatives, Inc. (DAI) recently completed a study of the impact of private voluntary organizations in two developing countries, Kenya and Niger.<sup>1</sup> The purpose of this study was :

- . To document evidence of development impact in PVO-assisted projects and to identify the determinants of impact; and
- . To test for and develop a cost-effective methodology for documenting and explaining such impact within a comparative framework.

This paper will assess the study, with particular attention given to its implications for AID's evaluation activities.

AID's Evaluation Needs

The assessment to be made in this paper will be based largely on my perception of AID's evaluation needs and current evaluation activities. They may differ significantly from those of others, and will be discussed under four major headings: policy, program, project and Agency effectiveness.

a. Policy

Policy evaluation is a review of Agency development strategies "at the highest level of meaningful generality". While there is

---

<sup>1</sup> Footnotes to this report appear on page 16.

considerable room for disagreement over what constitutes the highest level of meaningful generality, it should include, among other Congressional mandates, the following concerns:

- What progress has the Agency made in realizing "New Directions" objectives?
- To what extent are the methods being used to attain "New Directions" objectives better or worse than use of more traditional methods?
- Is there a meaningful role for the private sector in developing nations to play in AID as distinct from OPIC activities?
- In implementing the so-called "basic needs" strategy has adequate attention been given to the time-phasing of social service and income generating activities?
- Inasmuch as it takes a decade to run through a major policy cycle, is adequate attention being given to future problems of development, or more generally, to future realities?

It should be emphasized that this list is intended as exemplary of policy considerations and not necessarily the most important set of issues. It is worth noting that much of the data to answer the questions posed should come from project-level evaluations that permit both cross-project and cross-program assessments.

#### b. Program

The Agency uses a series of programs to realize its legislated objectives, and clearly these programs should be evaluated periodically. There is probably no one "program" definition that is best; periodic evaluations should be made of whatever the Agency defines to be programs inasmuch as they will have, at least to a certain extent, lives of their own. This means

functional programs such as potable water and rural roads should be evaluated; it also means that periodic evaluations should be made of non-functional programs such as PVOs and even Regional Bureaus. Once again, much of the information for program evaluations should come from comparative assessments of project impact. The comparisons are useful in that they provide a basis for analyzing what works, what does not and why.

### c. Project

Theoretically, at least, the same evaluative cycle should take place at policy, program and project levels: each should be designed, initiated, evaluated with the evaluation results used to make improvements. At the project level, a different evaluation strategy is called for than for either policy or program purposes. Considerable effort should go into formative assessments where, unlike the case of summative evaluation, the emphasis is on helping project management improve upon what they are doing.

The argument can be made that impact assessments are not appropriate for project-level purposes. This can be discussed under the headings of cost, timing, state of the art, and integrity. First, an impact analysis is going to be expensive relative to AID's customary evaluation work. But even at \$100,000 per project, assuming there are 1,000 "significant" Agency projects worth evaluating annually, this amounts to only about \$10 million, which is simply not a lot of money.

The timing issue is more troubling. The critical question, which will be discussed in greater detail later, is how much time must elapse after project start-up before meaningful impact can be seen or foreseen. Again, in anticipation of the discussion that

will follow, benefits that are likely to be self-sustaining can rarely be documented until a number of years have elapsed, say three or four, and by then foreign assistance is winding down. There is a danger that impact assessments will be attempted too soon.

Ultimately the two preceding points relate directly to the state of the art in impact work, which is in the early stages. High caliber people are needed to come up with new approaches; time is needed to experiment with alternatives. High quality people and time to experiment are costly. We should be able to develop techniques that allow us to predict impact as a means to get around the timing problem.

The objectivity of evaluative work should always be questioned when those responsible for projects also sponsor the evaluations. This is largely the case in AID, where the Regional Bureaus both approve projects and sponsor their evaluations. This will not be a problem so long as Bureau leaders believe the evaluative cycle will lead to future project improvements.

The arguments against moving ahead with impact assessments for project-level purposes are not convincing. It is time that project managers were forced to think in goal/purpose terms rather than just input/output terms. The costs are not excessive and, while many thorny issues remain on "how to do it," there is great potential for advancing the state of the art if a sustained and systematic effort is applied to it.

#### d. Agency effectiveness

Assessments of Agency effectiveness require two types of information. Firstly, information on policy, program and project impact is required. The second type of information has to do with the costs of Agency activities. As to the latter, budget information is immediately relevant and management studies are often employed to find ways to reduce costs.

#### AID's Current Evaluation Activities

Set against these evaluation needs, what can be said about AID's current evaluation activities? A summary assessment would be that they are woefully inadequate, but improving. Little, if anything, is being done at the policy level. Some program assessments have recently been initiated. The majority of AID's evaluation work is focused at the project level, but because methodological standards are not effectively enforced and because virtually all evaluations stop far short of measuring project impact, it is impossible to draw any conclusions concerning long-term project effects. In such circumstances, cross-project assessments are next to impossible using the existing project evaluation data base. Finally, there is no system whereby the results of evaluations are fed back into the project design process.<sup>2</sup>

There are reasons to expect improvement. There is a general awareness of evaluation inadequacies at all levels within the Agency, and Congress has called for the Agency to improve upon its evaluation work (the so-called Glenn Amendment). It is against this backdrop that the PVO study will be considered.

### What the Study Attempts to Do

Given the state of affairs described above, the PVO study was an extremely ambitious undertaking. It was an attempt at cross-function comparative impact assessment with analysis for programmatic purposes. Some elaboration on this statement is in order. The study was not limited to a particular PVO activity, but in no way was it a comprehensive experimental prototype. It looked at a limited number of PVOs in two African countries. Rather, an attempt was made to develop measures to assess a variety of PVO activities intended to increase the income of small farmers. The comparative analysis component was intended to develop recommendations for the Agency's PVO program.

Questions can and will be raised as to whether the attempt was overly ambitious. It can be rather persuasively argued that the study should have been limited to doing a more thorough job on a more restricted agenda. On the other hand, it can be argued that a complete run-through is useful every so often to give a comprehensive picture of the problems and potentials of what might be done. Regardless of the relative merits of each position, the PVO study attempted the latter and some of its component parts will now be examined.

### Impact Categories (Constituent Variables!?!): Which ones are right?

It is relatively easy to judge a project against a set of stated objectives. An attempt to assess development impact moves one into uncharted waters. By defining development impact as significant intended and/or unintended project effects, we should

be able to identify effect categories. This is necessarily a somewhat abstract exercise, however it is a critical stage in the process of insuring that major effects are neither overlooked nor double-counted. In the PVO study these categories were: direct benefits, benefit continuation and benefit growth. While the first two categories are well-defined, a broader and more detailed definition of benefit growth is warranted. The study defines benefit growth as project outputs that lead to a diversification of development benefits not specified in the original project. Benefit growth, however, may encompass a host of effects that should be broken out analytically into at least the following sub-categories:

- A catalyst among project participants: project activities serve as a spur for project participants to take on additional development activities;
- Spread: persons not part of the project copy project and project-related activities;
- Backward linkages:<sup>3</sup> project activities provide incentives for new activities early in the "production" cycle, e.g., project activities create a market for the production of labor-intensive farm tools;
- Forward linkages: project activities provide incentives for new activities late in the production cycle, e.g., the project increases production thereby justifying the construction of a processing plant; and
- Replicability: somewhat related, the project generates a level of benefits at a low enough cost to warrant wide-spread replication elsewhere.

At the measurement level, caution should be taken to avoid the double-counting of growth and other project benefits. However, the effects outlined above are at least conceptually different enough to merit separate consideration.

In summary, it is important to remember that impact analysis as a methodology is in its infancy. The question of what constitutes the "right" categories is far from resolved. And yet, some consistency is required if comparative work on project impact is to be carried out. Moreover, the objective of virtually all project impact studies to date has focused on observable effects that can be linked back to the project. Yet, there is a formidable body of literature that argues that if project effects cannot be traced into a far broader socio-economic setting, one cannot be sure that the net effects will not be far greater or less than the gross effects.<sup>4</sup> Conceptually, this argument is unassailable; project effects can have wide-ranging impacts. Operationally, there are budget and time limitations to what can be done, and, at this point, it is pushing the state-of-the-art to get project impact data, much less link that into a broader context.

#### Category Indicators

Having settled on categories, indicators for each category must be developed. This is where creativity is required. Usually one must find a middle ground between the ideal indicator and available data, and sometimes the best proxies appear to have very little direct relation to what is being measured. In the laudable effort to be as rigorous as possible, there is the danger that indicators most easily quantified will be given a disproportionate share in the analysis. The ordinary problems of getting the right indicators are compounded in this study by the need to cover a wide range of development activities.

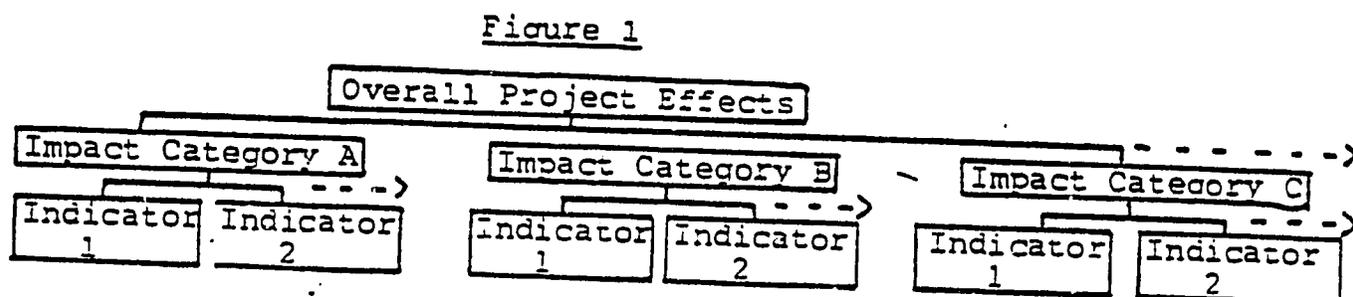
In such circumstances, the ideal procedure is probably to develop an "effects tree" for each activity separately, decide at what level on the tree you want to measure effects, and develop indicators accordingly. In the case of the PVO study, such a procedure was not possible inasmuch as the project activities to be examined were not known in a number of cases until the field visits. This necessarily required the development of some admittedly seat-of-the pants indicators.<sup>5</sup>

Numerous technical questions can and should be raised concerning the indicators employed. Without dwelling on particular technical issues, it should be emphasized that the selection of indicators is critical to the success or failure of evaluation work, and usually far too little time is devoted to thinking about and experimenting with different indicators.

#### Weighting and Timing

In analyzing project impact, two important factors need more attention: weighting and timing.

The use of impact categories and category indicators leaves one with the situation depicted in Figure 1.



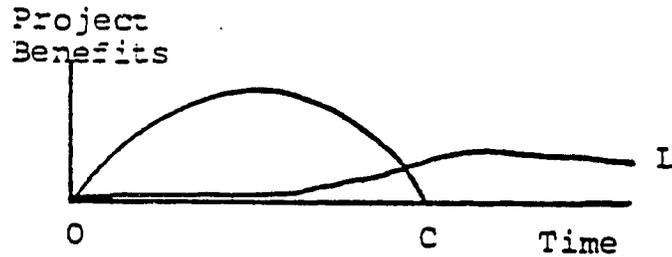
It should be immediately obvious from the figure that some weighting scheme is needed to get from indicators to categories to

overall effects. The weighting scheme employed in the PVO study was to assign equal weights first to indicators and then to categories.<sup>6</sup> While this approach has an easily understandable appeal it would be a remarkable coincidence if all categories and indicators therein should be weighted equally. For example, so-called direct benefits will often be little more than a reflection of the foreign assistance resource transfer and have no real bearing on development impact. This line of argument would suggest giving a lesser weight to the direct benefits category than to others.

It is extremely difficult to talk with much certainty about appropriate weights. It is, on the other hand, quite possible to talk about possible ranges for weight values and run simulations using different weighting combinations to see how important the weighting assumptions are to the conclusions. Such simulations should be part of any future comparative impact studies.

Projects examined in the PVO study had been operating for quite different lengths of time. In such circumstances, one should wonder whether a comparative assessment is not quite unfair to certain projects. Perhaps the issue can best be seen graphically. In Figure 2, project benefits are portrayed vertically, while time is measured on the horizontal axis.

Figure 2



It is reasonable to hypothesize that for most projects there are two "benefit profiles" that occur. The first, measured in Figure 2 by OC results directly from foreign assistance (money, personnel or equipment) being provided to the project area. These transfer benefits will cease when foreign assistance ceases (at time C in the figure). Lasting benefits are likely to follow a different time profile. As line OL in Figure 2 suggests, they are not likely to even start to manifest themselves until considerable time after the foreign assistance has commenced and probably not reach "full bloom" until several years after foreign assistance has ended.<sup>7</sup>

Clearly, different "benefit profiles" or "gestation periods" exist than the one portrayed in Figure 2, but the important point is that whatever the profiles are, when a project is evaluated may make an important difference on how the project appears to be doing. It is possible for experienced development professionals to anticipate with a reasonable degree of accuracy what the profiles will look like and base their evaluations on them (provided they donot need hard documentation for their conclusions). It would, however, also help if AID would commission some studies on what these profiles look like for different project types.<sup>8</sup> Until such studies are undertaken, the time criticism will be a legitimate one.

### Distributional Issues

The "New Directions Mandate" calls for AID-assisted projects to help the poorest segment of society. Suppose that it can be documented that an AID-assisted project caused a \$100 increase in the incomes of project participants, but only half the project participants qualify as poor. How should benefits be measured? Can it be assumed that benefits from PVO projects accrue to the poor, thereby eliminating the problem? This appears to be the assumption underlying the PVO study. Perhaps it was realistic for the projects examined. Presumably, the proper way to deal with this problem when it emerges is to give benefits that accrue to wealthier members of society less weight in calculating benefits than those accruing to the poorest segments of society.

### Analysis: Qualitative or Quantitative

Development dynamics are complex. In attempting to measure the impact of a project it is necessary to separate project effects from other operative forces. This is a complex undertaking, and ideally, one would employ multivariate statistical measures to separate out the project effects. In development work we fall far short of the ideal because of data shortcomings and quantitative models that are not sophisticated enough to distinguish between project and other effects.

The PVO study falls between the extremes of what has been done quantitatively and qualitatively. It does not resort to regression analysis or other statistical methods,<sup>9</sup> but it does scale projects quantitatively and speculates on the causes of relative success and failure.

The effort to be judiciously rigorous is to be applauded. The attempt at rigor forces a structured approach to data collection in the field, and such an approach is more likely than not the best way to go about accumulating data for the purpose at hand. Another positive reason for rigorous analysis is that it is then clear what the "analysis" is doing and hence it is easy to be critical.

#### Perceptions Data

The PVO study might be criticized for not giving much importance to the perceptions of the people the projects were intending to benefit. Concretely, for relatively small additional sums of money, local firms could have been hired to survey the perceptions of relevant populations. DAI has never been much of a supporter of such surveys: it is difficult to get true responses to such surveys; it is costly and time-consuming to construct anything that comes close to a representative sample of relevant populations; and respondents are likely to be far more impressed by large resource transfers than by projects that rely heavily on self-help. However, these shortcomings (and others) are well-known, and it would probably be useful if some experimentation with formalized survey work was undertaken.

#### Accurate and Adequate Supporting Data

What criteria should be employed to determine the accuracy and adequacy of data? This is a troubling and important question in the context of summative evaluations. Frequently, evaluators do not have or take the time to get a true picture of what the project is achieving. By the same token, the written criticisms of evaluators will frequently touch off a frenzy of defensive paperwork by those being criticized. Who is to be believed?

Firstly, it is important to distinguish between accurate and adequate. Adequacy has more to do with the volume of data and the way in which it is presented. We all know what accuracy means. While it is possible from Washington to reach judgements about adequacy, the question of accuracy is another matter. It is rarely a black or white issue, but rather one that calls for judgement. It is sobering for development experts to reflect upon the fact that there are few real project successes as measured against true development objectives. For the most part we are simply involved in a resource transfer process that has little sustaining effect. In such circumstances, the evaluator should not be too harsh in his assessments. By the same token, there are projects that should have been terminated and some that never should have been initiated.

The judgement on adequacy and accuracy should not rest with either the evaluators or those being evaluated. Regarding the PVC study, the Agency has the study and supporting documentation. It is now in the process of receiving PVO reactions and must be the final arbiter.

If the PVO reactions are convincingly negative, three issues need to be considered. Firstly, is the reaction largely due to the summative nature of the evaluation? Clearly, formative evaluations, i.e., evaluations intended to assist the PVOs in making their projects better would be more to their liking. Such evaluations are useful and should be done. However, the PVO study was not intended for that purpose; rather it was a summative evaluation intended to inform AID if PVO activities were having a positive impact. Secondly, is the reaction due largely to the fact that the

contractor simply did not have enough time/resources to do an accurate/adequate job? There is a discussion of costs and resources in the study and this should be helpful in reflecting on what is needed to do the job better. Finally, is the reaction contractor-specific? It is quite possible that the persons and/or methods employed are simply not satisfactory from the PVO viewpoint. On all these possibilities, the Agency must reach its own conclusions.

### Conclusions

This review of the PVO study has attempted to raise a number of points concerning impact assessments that warrant more discussion than they have heretofore received. Impact studies, as this study demonstrates, are expensive and complex. As a methodology, they are in their infancy. However, if AID is to take its new "Evaluation Mandate" seriously, there would appear to be no alternative to developing this approach to project assessment for program, policy and Agency effectiveness.

Footnotes

1. "The Development Impact of Private Voluntary Organizations: Kenya and Niger", prepared for the Office of Private and Voluntary Cooperation, Agency for International Development by Development Alternatives, Inc., February 1979.
2. For a further discussion of staff incentives and accountability for project effectiveness, see: Michelwait, Sweet, Morss, The New Directions Mandate: Studies in Project Design, Approval and Implementation (Westview Press, forthcoming).
3. I am grateful to Nick Owens, of the New Transcentury Foundation, for discussions on backward and forward linkages.
4. See Bela Belassa, "The 'Effects Method' of Project Evaluations", Oxford Bulletin of Economics and Statistics 38, (November 1976) : 219-231.
5. This approach should be contrasted to the procedure being followed in the sector studies being carried out by the Office of Evaluation in PPC.
6. More accurately, a one/zero weighting scheme was employed. All indicators/categories in the study were given one weight, while all indicators/categories not employed were given zero weight.
7. Of course, Figure 2 presents an optimistic situation. In less successful projects, OL never rises above the horizontal axis.
8. Several years ago Development Alternatives, Inc. proposed that such a study be undertaken by the Agency. For details on the methodology to be employed, see "A Proposal to Study the Time Profile of Benefits to Small Farmers" submitted to the Office of Technical Assistance, February 1975.
9. See, Elliott R. Morss, et al., Strategies for Small Farmer Development, 2 vols.--(Colorado: Westview Press, 1976).
10. "Information for Decision-making in Rural Development", prepared for the Office of Rural Development, Agency for International Development by Development Alternatives, Inc., May 1978.