

APR 1980
15N 1034
34

AN ASSESSMENT OF AID'S PROJECT EVALUATION SYSTEM

Office of Evaluation
Working Paper No. 34

Gerald M. Britan
Northwestern University
September 1980

SUMMARY

Concern has recently grown about how adequately AID's Project Evaluation System is meeting the agency's program and policy planning needs. This report is based on an analysis of AID's regular Program Evaluation Summaries (PESs) as well as an assessment of the larger evaluation system that these summaries reflect. Many current PESs were found to be relatively superficial project appraisals that failed to tap much of what should have been easily available project impact data. To a large extent, this problem was seen to result from AID's over-reliance on an "experimental" evaluation model that placed unrealistic demands on project managers, while ignoring their special experience and expertise. It is recommended: 1) that PES reports be oriented more clearly towards the kind of process and direct impact data with which field development officers are most familiar, 2) that a separate project monitoring system be established and that fewer regular evaluations be conducted for each project, 3) that AID's overall evaluation system be more clearly focused on direct project impact data, and 4) that a coordinated system of monitoring, regular evaluations, special evaluations, and studies be established with separate and clear responsibilities for project managers and centralized evaluation specialists.

INTRODUCTION

AID's regular project evaluation system is intended to provide data and analysis that can help improve both everyday project implementation and wider program and policy planning. At the core of this evaluation system is the Project Evaluation Summary (PES) which (1) records evaluation decisions for project participants, (2) notes that evaluation has occurred, (3) summarizes project "progress," and (4) clarifies any "lessons learned" for similar projects elsewhere.

AID's project evaluation system has been developed over more than a decade and has functioned smoothly for some time. Recently, however, there has been a growing concern about the evaluation system's adequacy. In part, this reflects questions about the appropriate role for regular evaluations within AID's increasingly complex structure of special evaluations, impact evaluations, ex post facto evaluation, evaluative studies and other efforts to measure and record AID's experience. In part, it reflects uncertainty about what kinds of data regular evaluations can and should provide for their varied audiences of project managers, regional bureau administrators, program coordinators, and policy analysts.

Just as project evaluation provides an important basis for project improvements, an assessment of AID's project evaluation system can suggest better ways of meeting the agency's data and analysis needs. The present report, and the research upon which it was based, represents the first step towards such an assessment. Although it focuses on the adequacy of AID's Project Evaluation Summaries, the report also addresses a number of important underlying issues for AID's overall evaluation system.

THE NATURE OF THIS REPORT

This report was developed through an iterative, interactive research process involving both the author and PPC/E/PES staff. First, a small sample of PESs was chosen for analysis, including both "average" submissions and those that PPC/E/PES staff judged to be of "high quality." The content of these PESs was then appraised and a series of intensive staff discussions formed the basis for the development of two documents: (1) a typology of PES evaluation criteria including issues and concerns that could be addressed within the existing PES structure (Appendix 1), and (2) an outline of questions and concerns that could be addressed within a restructured evaluation and reporting system (Appendix 2). These documents provided the basis for a second series of intensive discussions that focused on the nature of AID's wider evaluation strategy and the place of regular evaluations within it. The results of all of this analysis are presented in this final report.

OVERVIEW OF SAMPLE PESs

The following comments are based on an appraisal of twelve PESs selected from those submitted during the first three quarters of fiscal year 1980.

Since the goal was to develop a set of criteria that good evaluations could fulfill, the sample was biased to include a disproportionate number of PESs that PPC/E/PES staff considered especially well done. Even so, the quality of the PESs reviewed varied considerably: Some were clear and concise, others were disorganized and confusing. Some provided detailed assessments of project implementation and results, others provided little if any insight into either the process or magnitude of project impact. Few PESs were analytically sophisticated, Few included any detailed or rigorous appraisal of a project's broader implications for development practice. Many of these inadequacies seemed to result from PES reporting instructions which emphasized the collection of outcome indicators for inferential testing, but which ignored the insights, capabilities, and first hand experience of project managers. In general, the PESs placed too little emphasis on process-oriented data about project impacts--about how project outputs actually affected intended beneficiaries. Such information could have easily been obtained if PESs had sought the kinds of anecdotes, examples, case studies and specific results measures that could only be gained through direct experience in the field.

More specifically:

- (1) PES Summaries often proved extremely confusing to the uninitiated reader, since they usually turned directly to discussions of "project status" without considering project design or setting. Although such background data could sometimes be gleaned from later sections, those PESs that included an initial project abstract or overview were much more understandable. This deficiency is considered particularly important because staff turnover at both the field level and in Washington places a requirement as to at least succinct background information for evaluation reports.
- (2) Evaluation Methodology usually consisted of vague statements about how the Evaluation Review was conducted or at best, descriptions of the methods through which data were collected. Questions of analytical methodology (how conclusions were or could be validated) were rarely considered. Yet, this is precisely the crucial issue for an evaluation strategy that is, according to present guidance, strongly oriented towards summative, experimental inference.
- (3) External Factors were often ignored. PESs that considered "external factors" usually focused on problems of organizational coordination and project implementation. Little information was ever provided about a project's socio-economic setting or its basic theoretical assumptions.

When earlier project overviews were lacking, discussions of "external factors" could be extremely confusing.

- (4) Inputs, Outputs, Purposes, & Goals were often discussed repetitively. Problems with inputs, for example, might be repeatedly cited to explain failures to achieve expected results at each subsequent level of the logframe. The issues would often have been better expressed in a single narrative. The greatest emphasis, moreover, was placed on material inputs and outputs. Measures of purpose and goal achievement usually involved only vague judgments. There was little concern for substantiating findings or for assessing broader theoretical implications (i.e., the validity of hypotheses linking outputs and purposes, etc.).
- (5) Information about beneficiaries was not always included. Those PESs that did consider a project's impact on beneficiaries usually repeated pro forma statements from the initial Project Paper. Distribution and spread effects were rarely discussed. Again, the PESs exhibited little concern for applying rigorous analytical methodologies that could validate findings about a project's impact on beneficiaries.
- (6) Implementation problems were often considered in "Lessons Learned", but PESs rarely addressed larger issues of development policy, strategy, or theory.

In general,

- many reporters simply repeated earlier logframe or PP statements.
- The PESs were very "project oriented," that is, they focused on the monitoring of physical project inputs.
- The PESs exhibited little concern for the kind of rigorous analytical methodology that is essential for an evaluation strategy oriented towards inferential (experimental) testing of summative results indicators.
- The variable quality of the PESs suggests that some improvements in evaluation data could be obtained within the existing PES framework. At the same time, the findings also suggest that a more radical transformation of the regular evaluation system might be desirable.

TYOLOGY OF PES EVALUATION CRITERIA

The typology of PES Evaluation Criteria (Appendix 1) was developed through reviews and content analyses of background documents and sample PESs as well as intensive discussions with PPC/E/PES staff. The typology delineates the kinds of evaluation issues that could or should be addressed within the existing structure of PES reports. It is intended to provide a set of criteria for a program-focused analysis of the quality of existing PES reports, as well as a basis for new reporting instructions that are more clearly focused on broader program and policy concerns.

Although some of the questions included in the PES typology are open-ended, this is seen as precisely what is needed to tap the professional experience and judgements of development officers in the field. Certainly, any new PES instructions should be thoroughly field tested before wider adoption. Even so, it seems unlikely that altered instructions can solve all the PESs problems. What is probably needed is a broader restructuring of evaluation reporting to focus on those impact issues that can best be addressed during project implementation.

OUTLINE OF A RESTRUCTURED PROJECT EVALUATION SUMMARY

The complicated (and sometimes confusing) structure of the current PES reporting system is not well targeted towards larger program evaluation issues. Instead of emphasizing the process-oriented data with which Program officers and project managers are most familiar, current PES instructions focus on final results indicators that are rigidly tied to AID's use of the Logical Framework (logframe). This logframe model dramatically collapses the causal relations linking project "outputs" and "purposes" and relies on "black box" experimental inference to validate evaluation conclusions. Unfortunately, the comprehensive data and rigorously controlled comparisons that such experimental inference requires are rarely obtainable in the field.

To address this problem, an outline was developed for a new evaluation reporting system oriented towards more realistically answerable questions about project implementation and impact (see Appendix 2). This outline focuses on process-oriented data about how "inputs" produce "outputs," how "outputs" achieve purposes," and how "outputs" affect "beneficiaries." Particular attention is paid to the achievement of New Directions goals. The emphasis is on developing an evaluation and reporting system that can promptly provide the best available data on project implementation and direct project impacts. Ideally, these changes in PES reporting would be combined with a restructuring and simplification of the larger AID evaluation system. The nature of such changes are discussed below.

LARGER EVALUATION ISSUES

Project evaluation involves much more than collecting raw data about the achievement of project inputs, outputs, and purposes.

Evaluations must be designed so that conclusions about the nature and causes of project impacts can be validated. Evaluations must demonstrate that measured impacts are, in fact, due to project activities (internal validity) and that similar results could be expected in other project settings (external validity). Many important decisions about evaluation design must be made before project implementation even begins.

Despite the claims of some practitioners, evaluation research is not yet a truly rigorous science. Methodologists continue to argue about the merits of evaluation alternatives and their applicability in various project settings. Today, two different (but surprisingly complementary) evaluation models have become dominant.

Experimental evaluations seek to validate conclusions about project or program effectiveness through indirect logical and statistical inferences. Project "treatments" (in AID's terms, the delivery of outputs to intended beneficiaries) are viewed as a "black box." Very rarely controlled comparisons are made between groups of potential beneficiaries that differ only in the fact of their participation in the project. Any differences that emerge between these groups are then logically attributed to project effects.

In practice, it is often difficult to ensure that there are no uncontrolled differences between treatment and non-treatment groups. The ideal design for this kind of evaluation is a pure experiment in which participants and controls are randomly selected. Unfortunately, randomized experiments in directed social change are rarely possible or even desirable--project services may be legislated for entire populations; or the exclusion of interested participants (for use as "controls") may be seen as unethical. In these situations, evaluation methodologists have developed various "quasi-experimental" designs, which contain identifiable threats to internal validity.

Experimental evaluations face several other problems. Such evaluations aim primarily at definitive either/or judgments about whether project effects have occurred; they tell us little about how or why such results were achieved. Thus, experimental evaluations prove most useful when project hypotheses are clear, goals discrete, and treatments simple (for example, in assessing disease vaccination programs). Experimental evaluations also require sufficient "cases" (data points) to minimize chance variation, and are most relevant for assessing a project's impact on individual participants rather than on groups, villages, regions, or sectors. As a result, many of AID's development projects would be difficult to evaluate experimentally, even if this experimental methodology were otherwise desirable and affordable.

Process (or contextual) evaluations take a rather different approach. Instead of trying to assess project impact through directly controlled comparisons, contextual evaluations directly study the processes through which projects ultimately affect beneficiaries (see Britan 1978a, Britan 1978b, Patton 1980). Conclusions are not validated through logical and statistical inference, but rather by isolating smaller and smaller processual steps until causal connections become intuitively obvious.

This contextual approach is especially useful for assessing broad-aim programs that have complicated treatments aimed at multiple goals (Weiss & Rein 1974). Such an approach can also help the researcher interpret abstract indicators of quantitative results and can provide important insights about project implementation, project evolution, unanticipated project outcomes, and a project's fit with its broader sociocultural setting. The greatest weakness of a contextual approach is the difficulty of assessing the relevance of an evaluation of a single project for other project settings. However, this problem can be substantially reduced if appropriate comparisons are built into the initial research design for evaluations of similar projects or similar elements of projects.

It should be understood that the distinction between experimental and contextual evaluation models is very different from the distinction between qualitative and quantitative evaluation data. At least in theory, experimental comparisons can be based on qualitative distinctions. Conversely, contextual studies usually measure a wide range of indicators and results. The issue is not what kind of data is collected, but how those data are analyzed to validate conclusions.

THE AID EVALUATION SYSTEM

The evaluation "model" that AID applies is strongly biased towards indirect experimental inference. This inference is based upon a logical framework matrix that provides a simplified picture of causes and effects and that specifies results indicators for measurement.

Although the logframe can be an extremely useful tool in project design and in establishing the initial basis for project "evaluability", its rigid application in project evaluation often seems forced and artificial. Most project managers lack the time, experience, or training to conduct the kind of rigorous controlled comparisons needed to validate logframe hypotheses. Instead of producing

sounder evaluations, the logframe tends to focus research on those aspects of project implementation that are most easily measured. As a result, much too little emphasis is placed on obtaining the kinds of process-oriented impact data with which project managers are most likely to be familiar. Evaluations based upon the logframe's multi-dimensional matrix of inputs, outputs, purposes, goals, indicators, and assumptions often become mechanical exercises in misplaced precision.

Over time, AID's evaluation "system" has become increasingly complex. This "system" now includes "regular" evaluations, "special" evaluations, "interim" evaluations, "final" evaluations, "ex post facto" evaluations, "impact" evaluations, and evaluative "studies". It is not always precisely clear why these different evaluation categories exist, what they encompass, or how they differ. Certainly, there is a great deal of overlap. Yet, despite this complexity, AID's evaluation strategy also seems poorly related to different stages of project implementation. Monitoring and evaluation are often intertwined. An attempt to "rationalize" and simplify the evaluation system--to make it better tuned to program and policy planning needs--seems appropriate.

More specifically, the AID evaluation system could incorporate a much greater concern for project "impact" at all stages of project implementation (and "impact" evaluation, as a type, could probably be subsumed by other categories). This would involve re-orienting "interim" and "final" evaluations towards data on the process through which impacts are achieved and on the specific results of projects for participants. "Regular" evaluations could effectively tap the practical experiences of both project managers and project beneficiaries, while also providing preliminary impact data that could signal the need for broader (and experimentally more rigorous) "special" studies.

At the same time, the "regular" evaluation system could also be simplified. Basic project monitoring could be accomplished in a separate annual or semi-annual report. Regular evaluation would then only be conducted at the mid- and end- points of typical projects. Special evaluations and studies would be reserved for situations involving particular high priority issues, unanticipated problems, or broader cross-cutting assessments of program or sector impacts (see Appendix 3 for a preliminary typology of evaluation alternatives).

Even if the existing structure of the AID evaluation system is retained, the PES reporting system could still be improved. This should be based upon the realization that the logical framework is not so much a model for project evaluation as a model for designing more evaluable projects. The logframe's emphasis on results indicators implies a strong orientation towards experimental inference. Yet few projects in the field can support the kind of rigorously controlled comparisons that such inference requires. Treating complicated, multi-dimensional projects from a logframe perspective may artificially simplify complex cause and effect relationships, especially at the critical point where project outputs are translated into the achievement of project purposes.

More useful evaluations of ongoing or recently completed projects would concentrate less on abstract "indicators" and more on the "hypothesized" processes linking logframe categories. This would focus attention on the kinds of data--e.g., examples, case studies, participant experiences, and direct results measures--that project managers are most likely to discover in the field. In other words, less emphasis should be placed on final judgments about project success and more emphasis on understanding how and why desired impacts are occurring.

BRITAIN

RECOMMENDATIONS

- (1) Project Evaluation Summaries should be restructured (along the lines suggested in Appendix 2) to focus on process and specific results data that are most likely to be available and useful during (and immediately following) project implementation. This involves a much greater emphasis on the links between existing logframe categories (especially on how outputs achieve purposes) and on the experiences and qualitative judgments of development officers in the field. The specific format for such a revised PES should be thoroughly field tested before wider adoption.
- (2) If the existing PES structure remains unaltered, reporting instructions should be revised (along the lines suggested in Appendix 1) to obtain more relevant program and policy data. Again, these new instructions should be developed in consultation with project managers and should be thoroughly field-tested before wider dissemination.
- (3) In most cases, there seems to be little need for annual project evaluation. If a simpler annual or semi-annual monitoring system were established, regular evaluations could be limited to interim and final reports for most projects. (Present PES items that could be easily included in such a monitoring system are listed in Appendix 4). This separation of monitoring and evaluation functions is intended to reduce the evaluation workload for project managers while enhancing the utility of both monitoring and evaluation data.
- (4) The AID evaluation system should be simplified and more clearly oriented towards project impact data. Regular evaluations should remain the responsibility of field personnel, but should be re-oriented towards the process and direct impact data with which these personnel are most familiar. These evaluations should

provide the basis for Mission and centralized Bureau and PPC/E analyses which can signal the existence of problems and/or the desirability of more rigorous special evaluations or cross-cutting studies.

It should be recognized that most project managers lack the time, capabilities, or resources to design or implement rigorous, in-depth evaluations—experimental or otherwise. Where such evaluations are indicated, they should remain the responsibility of central evaluation and/or outside (contracted) specialists.

- (5) The rigor and scope of any evaluation should be flexibly tied to the size and importance of the project under consideration. A simple monitoring system can provide much useful information. Regular evaluations, conducted twice during the course of a typical project, should focus on easily available data and should draw on the special knowledge and skills of project managers. Special evaluations and cross-cutting studies should utilize the skills and resources of central evaluation staffs. Regular mechanisms should be developed to signal the need for more rigorous or in-depth evaluation approaches.

REFERENCES CITED

Britan, Gerald M.

1978a "The Place of Anthropology in Program Evaluation"
ANTHROPOLOGICAL QUARTERLY, 51:119-128.

Britan, Gerald M.

1978b "Experimental and Contextual Models of Program Evaluation"
EVALUATION AND PROGRAM PLANNING, 1:229-234.

Patton, Michael Quinn

1980 QUALITATIVE EVALUATION METHODS. Sage Publications:
Beverly Hills, California

Weiss, R.S. & M. Rein

1974 "The Evaluation of Broad-aim Programs: Difficulties in
Experimental Design and an Alternative" In C. Weiss (Ed.),
READINGS IN PROGRAM EVALUATION. Englewood Cliffs: Prentice
Hall.

APPENDIX 1

TYPOLOGY OF PES EVALUATION CRITERIA

(Information that could or should be appropriately reported within the existing PES structure)

**---maybe answered with reference to/or attachment of earlier data or reports
*---not relevant for all projects
---could be reported on a coded checklist

DOES THE PES:

(13 - Summary)

- 1) Indicate what kind of project is being considered? (pilot, one of a series, follow-up, innovative, etc.)
- ** 2) Provide an overview (abstract) of the project's design?
- ** 3) Summarize previous changes in project design?
- 4) Summarize the project's current status?
- 5) Summarize the likelihood and timing for achieving planned outputs, purposes, and goals?

(14 - Evaluation Methodology)

- 1) Indicate why the evaluation was conducted? (decision driven, spotcheck, special problem, normal monitoring)
- 2) Indicate the stage of implementation being evaluated? (interim, final, ex post facto, cross-cutting)
- 3) Indicate who was involved in the evaluation review and what groups were represented? (host country representatives, intermediary or implementing agency, external evaluator, etc.)
- ** 4) Indicate what data or documents were used (attach bibliography), how evaluation data were collected and by whom? (field methodology)

- ** 5) Indicate how evaluation data were analyzed?
(case study, comparison with baseline, comparison with primitive control, quasi-experimental statistical analysis, true experiment)
- 6) Indicate the reliability of the evaluation's conclusions?
(internal and external validity)

(15 - External Factors)

- * 1) Indicate changes in the project's political environment?
(local, national, or international)
- * 2) Indicate changes in the project's economic environment?
(local, national, or international)
(e.g. changing local opportunities, regional price changes, currency devaluation, international markets)
- * 3) Indicate changes in the project's physical environment?
(e.g. flooding, fire, disease, drought)
- * 4) Indicate changes in the project's organizational environment?
(e.g. loss of project champion, changes in local agency responsibility, bureaucratic politics, altered assumptions about local organizational capabilities, etc.)
- * 5) Indicate changes in the project's socio-cultural context?
(e.g. invalid baseline data or assumptions; invalid assumptions about traditional production; invalid assumptions about the characteristics, motivations, knowledge, or capabilities of participants; unanticipated socio-cultural change)

(16 - Inputs)

- 1) Summarize the extent to which planned inputs (commodities, funding, and technical assistance) are being provided?
- * 2) Explain the reasons for any shortfalls in the quality or quantity of inputs?

(17 - Outputs)

- 1) Summarize progress towards achieving planned outputs?

- * 2) Explain the reasons for any shortfalls?
(e.g. altered inputs, organizational/implementation problems, altered external conditions, invalid assumptions about socio-economic context, invalid theoretical hypotheses about the links between inputs and outputs)
- * 3) Suggest changes in planned outputs? (new expectations)
- ** 4) Consider the likely/observed impact of expected/observed outputs on the achievement of project purposes.

(18 - Purpose)

- ** 1) Summarize the project's planned purpose?
- 2) Summarize progress towards achieving EOPs?
- * 3) Explain the reasons for observed/expected shortfalls?
(e.g. altered outputs, altered external conditions, invalid assumptions about socio-economic context, invalid theoretical hypotheses about the links between outputs and purposes?)
- * 4) Suggest changes in planned EOPs/purposes? (new expectations)
- 5) Consider the likely/observed impact of expected/observed EOPs on the achievement of wider goals?

(19 - GOAL/SUBGOAL)

- ** 1) Define the goal/subgoal to be achieved?
- 2) Analyze the project's observed/expected impact on achieving this goal/subgoal?
- 3) Consider the relationship of this project to other projects contributing to goal/subgoal achievement and to the mission's CDSS? (Including projects sponsored by other donors).
- 4) Analyze or test the validity of theoretical hypotheses linking purposes to goals?

(20 - Beneficiaries)

- ** 1) Indicate what kinds of individuals have/will directly benefit from the projects and in what ways?
 - a) characteristics of the target population, size, % who will benefit.
 - b) differential impact/distributive effects on men and women, poor and not so poor, farmers and laborers, etc.
- ** 2) Indicate what kinds of organizations and formal groups have/will benefit from the project? (e.g. farmers co-ops)
- ** 3) Indicate the project's expected indirect effects on individuals or groups?
- ** 4) Indicate the likelihood/occurrence of spread effects, demonstration effects, or independent replication?
- 5) Summarize the participation of beneficiaries or potential beneficiaries in project design, implementation or evaluation?

(21 - Unplanned Effects)

- * 1) Summarize any unplanned (desirable or undesirable) distribution of benefits? (e.g. benefits for civil servants, larger farmers, private contractors, etc).
- ** 2) Summarize unplanned socio-cultural effects? (e.g. on land tenure, on the nature of production/work, on the structure of families, on sources of motivation or prestige, on traditional patterns of leadership or authority.)
- * 3) Summarize unplanned effects on the environment? (degradation, erosion, disease).

- * 4) Summarize unplanned effects on health and population?
(nutrition, infant mortality)
- * 5) Indicate if unplanned effects alter the project's
assumptions or its place/contribution in the CDSS
and in achieving larger development goals?

(22 - Lessons Learned)

- * 1) Analyze this project's implications for improving local
participation in project design, implementation and
evaluation?
- * 2) Analyze the project's implications for improving
development implementation by host country organizations?
- * 3) Analyze the project's implications for improving/facilitating
the use of intermediary organizations (PVO's, private
contractors)?
- 4) Analyze the project's implications for other development
projects in the country/region? (lessons learned about
this setting)
- 5) Analyze the project's implications for other projects
in this sector/program area (lessons learned about
broader program sector and policy concerns)?
- 6) Analyze the project's implications for targeting projects
towards "Basic Human Needs" and the "Poor Majority?"

APPENDIX 2

AN OUTLINE FOR RESTRUCTURED
PROJECT EVALUATION SUMMARIES

**—might be answered by
referencing or attaching
earlier data/reports
*—not relevant for all projects
—might be reported on
checklist

- I. Project Data
 - a. Project type, implementation stage, evaluation type, data sources, and analytical methodology.
 - ** b. Project abstract, status summary (including previous design changes)

- II. Baseline Data
 - ** a. Characteristics of Target Population
 - ** b. Characteristics of co-operating/participating organizations
 - ** c. Socio-economic-political context
 - ** d. Other external parameters (assumptions)

- III. Implementation Problems
 - a. Delivery of Inputs
 - b. Management of Inputs to produce Outputs
 - 1. Management by Implementing Organizations (s)
 - 2. Co-operating among other host country organizations, PVO's, private contractors, and/or AID Mission

* 3. Unanticipated burdens on AID Mission

* 4. Possible problem solutions

c. Outputs Produced

d. Validity of hypotheses and assumptions linking
Inputs to Outputs

IV. Effects on Target Population

a. Local participation in Design, Implementation, and/or
Evaluation.

b. Descriptive analysis of the Process through which
outputs are affecting beneficiaries (and
differentially affecting classes of beneficiaries
such as men/women, small farmers/large farmers,
poor/not so poor).

c. Direct results measures - the affects of project
outputs on beneficiaries (and classes of
beneficiaries) that have participated thus far.

d. Descriptive analysis of the process through
which project outputs are affecting local
organizations (bureaucratic agencies, regional
centers, co-operatives, self-help groups, etc.)
and the ability of these organizations to
sustain development initiatives.

e. Other expected or observed impacts

f. Aggregate results indicators, as available.
(e.g. changes in income, farm production, health
indicators, etc.)

g. Validity of hypotheses and assumptions linking outputs
to purposes and goals. (How do project outputs
contribute to the achievement of higher level project
purposes.)

V. Lessons Learned

- a. Implications of this project for larger program and policy.
- b. Implications of this project for better targeting development efforts towards the "basic human needs" of the "poor majority".
- c. Implications of this project for improving "local participation" or enhancing the capabilities of local organizations for sustaining development initiatives.
- d. The validity of the underlying development hypotheses upon which this project was based.

5
1

APPENDIX 3

EVALUATION ALTERNATIVES

Decisions about data collection and evaluation design should reflect the nature of the project being assessed, the stage of project implementation and the reasons a project is being evaluated. The following table provides some preliminary guidelines:

1) Pilot Projects

Experimental evaluation to test any development hypotheses that will be the basis for project follow-ups.

Process evaluation of implementation as formative feedback for changes in project design, but also including baseline measures and available data on direct results.

2) Routine Projects/Follow-on Projects

Process-oriented case studies, including available baseline data, direct results data, and available data on aggregate results.

3) Large Projects

Comprehensive baseline data, process-oriented case studies, including direct results data and available data on aggregate results.

Experience evaluations may be appropriate if there are clear, widely applicable hypotheses for testing.

Expost facto collection of comprehensive results data and quasi-experimental comparisons may be indicated depending on regular evaluation findings.

4) Small Projects

Process-oriented case studies including available baseline, direct impact, and available aggregate results data. If regular evaluations indicate broader significance or special problems, expost facto quasi-experimental comparisons may be indicated.

- 5) Early in project implementation emphasis should be on monitoring the implementation process.

By a project's midpoint, process-oriented direct impact data (outputs to purpose) and some aggregate results indicators should be available.

By a project's end more aggregate results data should be available, but emphases must still be placed on direct impact measurers.

Ex post facto evaluations will usually concentrate on aggregate results measures and cross cutting quasi-experimental comparisons, but process data on continuing delivery of project services may still be important.

Pure experiments and some quasi-experimental comparisons must be designed from the start. This should be included in important pilot projects and for large, highly visible efforts.

APPENDIX 4

INFORMATION TO BE INCLUDED IN A PROJECT
MONITORING SYSTEM

I. From Appendix 1 - Typology of PES Evaluation Criteria

Items: 13-1, 13-2, 13-3, 13-4, 13-5,
14-2,
15-1, 15-2, 15-3, 15-4, 15-5,
16-1, 16-2,
17-1, 17-2, 17-3, 17-4.

II. From Appendix 2 - An Outline for Restructured PES Evaluation
Summaries

Items: Ia, Ib,
IIIa, IIIb1, IIIb2, IIIb3, IIIb4, IIIc
IVa

Of course, other fiscal and managerial information would also be included in the monitoring system, as needed.