EVALUATING THE IMPACT OF AID PROJECTS

A Final Report

VOLUME II: STUDY FINDINGS

Prepared for:

Agency for International Development
PPC/E/PTS
Washington, DC

Prepared by:

PCI (Practical Concepts Incorporated)
1730 Rhode Island Avenue NW
Washington, DC 20036

Study Team:

Molly Hageboeck (Director)
Lawrence D. Posner, Ph.D.
Brian Goodhart
Samuel R. Daines, Ph.D.
Michael Dalmat, Dr.PH

April 4, 1980

BEST AVAILABLE COPY

TABLE OF CONTENTS

PAGE

VOLUME II

CHAPTER ONE: OBJECTIVES AND APPROACH

A. STUDY OBJECTIVES I-1

1. General I-1
2. Specific Study Outputs I-2

B. STUDY APPROACH I-3

CHAPTER ONE: ANNEX A: AID SCOPE OF WORK

CHAPTER TWO: THE NEED FOR INFORMATION

A. THE INFORMATION AID NEEDS FROM MATURE & TERMINATED PROJECTS II-1
B. THE AVAILABILITY OF IMPACT INFORMATION IN AID II-6

CHAPTER TWO:

CHAPTER THREE: THE SCOPE OF IMPACT EVALUATIONS

CHAPTER THREE: ANNEX A: CANDIDATE EVALUATION SCOPE ELEMENTS & THEIR IMPLICATIONS

CHAPTER FOUR: THE TIMING, DURATION, & INTENSITY OF IMPACT EVALUATIONS

A. EVALUATION TIMING IV-1
B. THE DURATION & INTENSITY OF PROJECT IMPACT EVALUATIONS IV-6

CHAPTER FOUR: ANNEX A: FACTORS THAT AFFECT THE TIMING OF PROJECT IMPACT EVALUATION

CHAPTER FOUR: ANNEX B: THE EFFECTS OF SCOPE & METHODOLOGY CHOICES ON DURATION & INTENSITY

CHAPTER FIVE: IMPACT EVALUATION APPROACHES & METHODS

CHAPTER FIVE: ANNEX A: IMPACT EVALUATION APPROACHES & METHODS

CHAPTER FIVE: EXHIBIT A.1: MEASUREMENT OF MANDATE IMPACT INDICATORS

CHAPTER FIVE: EXHIBIT A.2: GETTING SPECIFIC ABOUT PARTICIPATION: ANALYSIS FOR PROJECT DESIGN, IMPLEMENTATION & EVALUATION

CHAPTER FIVE: ANNEX B: THE SKILLS REQUIRED TO IMPLEMENT PROJECT IMPACT EVALUATIONS

BEST AVAILABLE COPY

PAGE

CHAPTER SIX: IMPACT EVALUATION COVERAGE

CHAPTER SIX: ANNEX A: THE ADVANTAGES & DISADVANTAGES OF COVERAGE & SELECTION OPTIONS

CHAPTER SEVEN: ROLES & RESPONSIBILITIES FOR IMPACT EVALUATION

A. CHOICE OF TOPICS & PROJECTS VII-2
B. METHODS VII-4
C. IMPLEMENTATION VII-7
D. UTILIZATION OF EVALUATION FINDINGS VII-10

CHAPTER SEVEN: ANNEX A: THE OPTIONS FOR LOCATING A SERIES OF EVALUATIONS RESPONSIBILITIES

CHAPTER SEVEN: ANNEX B: LESSONS FROM THE EXPERIENCE OF OTHER ORGANIZATIONS

CHAPTER EIGHT: THE PLACE OF IMPACT EVALUATION IN AID'S EVALUATION SYSTEM

A. THE RELATIONSHIP OF IMPACT EVALUATION TO THE PES VIII-1
B. APPROACHES TO FUNDING IMPACT EVALUATIONS VIII-3
C. STAFFING IMPACT EVALUATIONS VIII-4

CHAPTER NINE: THE COST OF IMPACT EVALUATION

A. THE COST OF PROJECT IMPACT EVALUATIONS IX-2

1. Analysis of Existing Data IX-3
2. Case Study Data Collection & Analysis IX-3
3. Sample Survey Data Collection/Analysis Using an Existing Sample Frame IX-3
4. Sample Surveys where the Sample Frame must be Developed IX-3

B. COSTS ASSOCIATED WITH MANAGEMENT OF THE IMPACT EVALUATION PROGRAM IX-6
C. LABOR ESTIMATES FOR ADDITIONAL PROJECT DESIGN PHASE WORK THAT WILL PROVIDE THE AGENCY WITH AN IMPROVED BASIS FOR EVALUATION IX-11

1. Improved Project Logic IX-11

BEST AVAILABLE COPY

TABLE OF FIGURES

	PAGE
VOLUME II	
CHAPTER ONE: OBJECTIVES AND APPROACH	-
CHAPTER TWO: THE NEED FOR INFORMATION	-
CHAPTER THREE: THE SCOPE OF IMPACT EVALUATIONS	
ANNEX A: TABLE III.A.1: A LOGICAL FRAMEWORK OF THE OBJECTIVES AND INDICATORS FOUND IN THE INTRODUCTORY SECTIONS OF THE FOREIGN ASSISTANCE ACT	III.A.-4
ANNEX A: TABLE III.A.2: ANALYSIS OF THE FREQUENCY OF USE OF EACH FAA MANDATED OBJECTIVE/INDICATOR IN 1,940 AID PROJECT PURPOSE AND GOAL STATEMENTS	III.A.-5
ANNEX A: TABLE III.A.3: APPARENT DIFFERENCES IN THE MEANING OF THE TERM MANDATE IMPACT IN CONGRESSIONAL & AID DOCUMENTS	III.A.-15
ANNEX A: FIGURE III.A.1: NARRATIVE SUMMARY OF THE EXTENDED LOGIC OF THE PROJECT	III.A.-21
CHAPTER FOUR: THE TIMING, DURATION, & INTENSITY OF IMPACT EVALUATIONS	
FIGURE IV-1: HYPOTHESIZED OVERLAP BETWEEN INTERMEDIATE & SUSTAINED IMPACT OF A PROJECT	IV.-4
EXHIBIT A.1: A ROUGH ANALYSIS OF IMPACT EVALUATION TIMING FROM THE PERSPECTIVE OF PROJECT PLANS & SCHEDULES	IV.A.1.-1 to IV.A.1.-15
ANNEX B: TABLE IV.B.1-1: ESTIMATED ALLOCATION OF COSTS IN A SAMPLE SURVEY FROM AID'S MANAGER'S GUIDE TO DATA COLLECTION, BY MOLLY HAGEBOECK, et.al., PCI	IV.B.-3
CHAPTER FIVE: IMPACT EVALUATION APPROACHES & METHODS	
ANNEX A: FIGURE V.A.1	V.A.-6
EXHIBIT A.1: Table 1	V.A.1.-3
EXHIBIT A.1: Table 2	V.A.1.-4
EXHIBIT A.1: Table 3	V.A.1.-7
EXHIBIT A.2: Figure 1	V.A.2.-4

BEST AVAILABLE COPY

TABLE OF FIGURES

	PAGE
EXHIBIT A.2: Figure 2	V.A.2.-7
EXHIBIT A.2: Figure 3	V.A.2.-32
CHAPTER SIX: IMPACT EVALUATION COVERAGE	-
CHAPTER SEVEN: ROLES & RESPONSIBILITIES FOR IMPACT EVALUATION	
ANNEX B: FIGURE VII.B.-1: WORLD BANK PROJECT EVALUATION PROCESS	VII.B.-3
CHAPTER EIGHT: THE PLACE OF IMPACT EVALUATION IN AID'S EVALUATION SYSTEM	-
CHAPTER NINE: THE COST OF IMPACT EVALUATION	
TABLE IX-1: DIRECT LABOR & OTHER DIRECT COST COMPONENTS OF EACH TYPE OF STUDY (ESTIMATE APPLIES ONLY TO "BASIC" SCOPE)	IX-4
TABLE IX-2	IX-5
TABLE IX-3: DIRECT LABOR & OTHER DIRECT COSTS FOR "SCOPE SUPPLEMENTS" IN EVALUATIONS	IX-7
TABLE IX-4: IMPACT EVALUATION PROGRAM MANAGEMENT COSTS -- YEAR 1 (LABOR ONLY)	IX-8
TABLE IX-5: IMPACT EVALUATION PROGRAM MANAGEMENT COSTS -- YEAR 2 (LABOR ONLY)	IX-9
TABLE IX-10: IMPACT EVALUATION COST ESTIMATE SUMMARY	IX-10

BEST AVAILABLE COPY

Objectives & Approach

CHAPTER ONE

In April, 1979 AID's Office of Evaluation, Bureau for Program and Policy Coordination (PPC/E) asked PCI (Practical Concepts Incorporated) to carry out a "Feasibility Study of an Ex-Post Project Evaluation System." The present volume together with the Executive Summary constitute PCI's final report on this study. The study was carried out under Contract Number AID/ort/C-1377, Work Order #25.

A. STUDY OBJECTIVES

1. General

The objective of this study was to assess the feasibility of improving evaluative information about mature and terminated projects sponsored by the Agency for International Development. AID's scope of work anticipated that the primary mechanism for securing this information would be an "ex-post" evaluation system that supplemented AID's existing management-oriented Project Evaluation System.

AID's scope of work and it's initial contractor briefings indicated that the Agency continues to need the type of management and replanning information generated by its existing Project Evaluation System. However, it also needs better information on the impact and sustainability of projects it undertakes in the developing countries. This additional information is needed for policy development, program and project design, financial planning, and AID's dialogue with the U.S. Congress and other external bodies.

PREVIOUS PAGE BLANK

The scope of work for this engagement called for an analysis of the constraints, benefits, costs and other requirements of an ex-post project evaluation system. It also required a description of such system characteristics as scope, methods, coverage, management roles and responsibilities, financial arrangements, staffing requirements, and the duration, intensity and timing of ex-post project evaluations.

Based on a preliminary review of AID documents and discussions with AID/W personnel, a slight shift of emphasis was made in the study's terms of reference. Specifically, it was determined that PCI should begin with a review of the Agency's needs rather than with the *a priori* conclusion that an ex-post project evaluation system was required. If PCI found that an ex-post project evaluation system was an optimal approach for meeting the Agency's needs, its final report would so indicate. However, if the study identified alternatives that would better serve AID's needs, these alternatives were to be presented and recommended.

2. Specific Study Outputs

Two types of outputs were required of this study:

a. Specific action-oriented recommendations for securing adequate evaluative information on AID's mature and terminated projects;

b. Recommendations for integrating the proposed approach for evaluative mature and terminated projects into AID's on-going systems for project design and evaluation.

Actionable approaches for securing evaluative information on mature and terminated projects are identified and described in this report. In preparing these recommendations, PCI began with the premise that the recommendations should, to the extent possible, build on the evaluation foundation AID has already developed. PCI's effort to integrate proposed new procedures and activities into existing systems and approaches is reflected throughout the report, as well as summarized in Chapter Eight.

B. STUDY APPROACH

The feasibility study AID envisioned was to be carried out in Washington, D.C. Documents available in AID and from other organizations, as well as interviews, were to be the primary sources of data for the analysis PCI would perform. A three-month time frame was set for the study.

The process used by the PCI team in conducting the feasibility study was an iterative one. The questions posed by PPC/E in the scope of work are interrelated; some questions logically must be answered, at least tentatively, before others can be addressed. Thus, the PCI team divided the investigation into three parts:

Step One:

The need for evaluation findings (that had led AID to consider the development of an ex-post project evaluation system). This step in the study focused on Question 1 in the scope of work. (The AID scope of work for this study is presented in Annex A following this chapter.)

Step Two:

The characteristics of evaluations that would satisfy such needs as were documented in Step One. The elements of the scope of work addressed in this step included:

- Evaluation scope (Question 4);
- Evaluation methodologies (Question 2);
- The timing of evaluation (Question 6);
- The level of effort:
  - a. Duration and intensity of the evaluations (Question 6);
  - b. The cost of evaluations (Question 3);
- The skills required to carry out the evaluations (Question 10).

Step Three:

The characteristics of a mechanism for carrying out evaluations with the characteristics identified in Step Two. The elements of the scope of work addressed in this step included:

- Coverage (the number of projects for which this type of evaluation is to be conducted) (Question 4);
- Roles and responsibilities for conducting these evaluations (Question 8);
- Financing the evaluations (Question 9);
- Relationship of the plans for these evaluations to other Agency procedures/systems and effects on this project design process (Question 5);
- Work force requirements/training implications (Question 10);
- Procedures for installing the new evaluation mechanism/approaches (Question 7).

The following chapters address PCI's findings, conclusions and recommendations concerning each of the major characteristics of impact evaluations in AID.

The data used by PCI in the course of this feasibility study came from the following sources:

- AID documents and memoranda that deal with AID's evaluation system, past and current project and program evaluations, evaluation issues, and evaluation experiences;
- Pertinent documents generated by other organizations that fund economic and social development projects domestically and in the less developed countries;
- AID's automated databases;
- Congressional hearings and the Foreign Assistance Act;
- PCI files on evaluation (developed over the course of many engagements with AID and other clients);
- Formal and informal interviews with AID personnel, conducted in the course of this and other engagements;
- Interviews with the evaluation staff at the World Bank and discussions with evaluation personnel from DHEW and the Evaluation Research Society.

AID SCOPE OF WORK

The feasibility study will examine the intrinsic features, benefits, costs and other aspects of creating and operating such a sub-system. It will address the relationships between such a sub-system and the larger evaluation system of which it is a part. Thus analysis of any of the questions to be addressed below may also require an examination of such intra-system questions as complementarity, competition for scarce resources, comparative benefit/cost characteristics, relative priorities, etc.

Specifically the questions to be addressed are as follows:

1. The need for and uses of ex-post project evaluation findings in various Agency functions such as policy and program design, budgeting, Congressional Presentation, and other management functions.

2. What evaluation methodologies and techniques would be appropriate to assess achievement of planned project targets, unplanned results, and impact on development in general and on target groups specifically?

3. The benefits and costs.

4. The scope, coverage, and other quantitative dimensions of the sub-system.

5. What changes, if any, would be needed in the design of projects to facilitate ex-post evaluation?

6. Timing, duration, intensity of effort and other operational aspects.

7. The possible need for phased installation over a period of time, e.g., three years.

8. Organizational roles and responsibilities including the role of the host country, of Mission and AID/W staff, etc.

9. What changes, if any, would be needed in Agency regulations and procedures for financing and conducting ex-post evaluations after the termination of AID contributions and services, e.g., what changes would be required in project budgets, project agreements, contracts and PASA/RSSA, etc.?

10. What are the work force and skill requirements in quantitative terms? What is the most appropriate use of direct-hire, contract, PASA/RSSA resources of combinations thereof. What kinds and amounts of training will be required? What is the workload, i.e., total direct-hire time required?

The Need for Information

CHAPTER TWO

A. THE INFORMATION AID NEEDS FROM MATURE & TERMINATED PROJECTS

The important additional information AID requires about its projects was found to be information on project impact. AID needs:

- Evidence of changes and accomplishments in the priority areas designated under the Foreign Assistance Act;
- Evidence of the distribution of project benefits to the target groups AID is mandated to serve;
- Evidence of the connection between development assistance and impact.

The need for information on project impact is most strongly felt at the top levels of the Agency, and in the offices that bear a responsibility for evaluation. AID's current need for information was found to be an outgrowth of the Agency's perceived inability to respond to Congressional and other external inquiries concerning the impact of its projects.

Congressional demand for evaluative information is generally acknowledged, within AID, to be the primary stimulus for a number of activities that are underway in the Agency, including both this feasibility study and AID's decision in 1979 to move ahead with 20-30 ex-post project evaluations this year. The AID Administrator and the Agency's senior evaluation personnel believe that the Agency must improve its ability to answer questions posed about the impact of its past projects as well as on the likelihood that the projects AID proposes for the future will yield the benefits they promise.

BEST AVAILABLE COPY

While PCI found that Congressional pressure for information was the "demand" to which new initiatives in AID are responding, we found no data to suggest that Congress expected AID to perform a miracle (e.g., come up with all the answers to all potential questions in less than a year). That is, the Congressional demand for impact information was not translated, by Congress, into a program that calls for a specific number of evaluations, of a specific type, for a Congressionally selected set of projects, completed within some fixed time period. Rather, Congress expects AID to take the steps that are necessary to secure information on the impact of its development assistance efforts.

A sense of urgency and priority attached to improving AID's supply of information about project impact is not evenly distributed through the Agency. Nor is it clear that all of the offices that do consider this a high priority also consider impact information to be a central focus.

Interviews conducted in Washington with AID field personnel and data from past studies by PCI and other contractors suggests that project impact evaluation is not normally a high priority in the Missions. The exception to this rule occurs when Missions are considering investing in "repeater projects". The reasons why impact evaluations are of low priority at the Mission level have been reported before:

- The lack of an immediate need for the information that an impact evaluation must provide;
- Lack of time and/or resources to mount and manage such evaluations;
- A tendency in AID's personnel/reward system to acknowledge individual achievement when it is associated with developing new projects rather than for serious efforts to facilitate implementation or rigorously examine the progress and impact of projects;
- Periodic turnovers in Mission staff which have the effect of diminishing individual commitment to project results;

BEST AVAILABLE COPY

- The subjective perception on the part of junior staff that (a) higher levels of the organization would rather not receive evaluation reports that find projects to have been ineffective and (b) individuals who are associated with evaluations that are controversial may find themselves being viewed as irritants in the system.

There are also differences in the way the need for information on project impact is viewed within AID/Washington. The sense that evaluations of mature and terminated projects are a time-urgent priority was found to be highest in PPC and in the Office of the Administrator. In these offices, discussions about evaluation have examined the possibility of evaluating all of AID's projects, i.e., developing a system that is applied to the full portfolio in the future, combined with a fairly comprehensive look at the past. Outside of PPC, the general sense of this situation is that AID needs to examine carefully the types of projects it has already carried out if it intends to propose new projects of that same type.

While the stimulus for improving AID's understanding of the impact of its projects is recognized by most AID personnel as coming primarily from the outside, PCI found that a good number of AID/W personnel view the Congressional interest in impact data as catalytic. These staff members sense that AID will be the main beneficiary of whatever impact evaluations are undertaken. PCI found that the perception of AID as a beneficiary of impact evaluations was held by those who feel such evaluations will be a burden as well as by those who welcome a movement in this direction.

AID personnel feel that impact information will be useful in AID for project design, and to some degree for policy formulation. On the other hand, they do not have a clear sense of how outside organizations such as Congress will use this information. There is some sense that Congress may use impact evaluation findings to "cut off" AID investments in types of projects that do not show a direct relationship to mandated benefits. Conversely, some AID staff members feel that impact data will lead Congress to authorize more investment in types of projects that prove their development value through

BEST AVAILABLE COPY

evaluations. This dual sense of Congress, as the primary audience for AID's evaluations of mature and terminated projects was a source of ambivalence in some discussions.

Host government utilization of impact evaluation findings came up a number of times in the interviews PCI conducted during this study. While AID personnel generally felt that such information would be useful to the host countries, the specific uses that they would be likely to make of such information was generally left unstated.

In the course of this study, AID has used the term "ex-post" evaluation, fairly consistently, to refer to the type of action it felt it should take to meet its current needs for information. Throughout the contract period, PCI has questioned the use of this term. We have been bothered, we believe, not by a semantic issue, but rather by the question of focus implied by a choice of terms. Early on PCI reported to AID that all the evidence we had gathered pointed to the need for information on project impact. PCI also indicated to AID that "impact evaluation" and "ex-post" evaluation were not necessarily the same thing. Impact evaluations are, by definition, focused on the substance of program or project results. The term "ex-post" refers to when an evaluation is conducted, not directly to its substance. For the following three reasons, PCI recommends that AID use the term "impact" to refer to the evaluation it plans as a means of gathering evidence on project performance against Purpose and Goal level objectives.

- AID's "pilot" ex-post evaluation studies which examined a series of rural health projects did not concentrate on impact information. If these studies become a reference point for the Agency, some of the energy needed to gather solid evidence of impact and benefit distribution may be siphoned off into the collection of information about Input Management and/or Output Level performance that can be more easily secured;

- PCI's examination of the question of when impact can be measured has suggested that in many projects it should be possible to observe impact shortly after the project Outputs are provided.

BEST AVAILABLE COPY

AID documents and some of our discussions with AID indicated that periods of three-five years after termination were being considered appropriate time frames for ex-post assessments. There appears to be no need to wait this long in most cases. Further, in some projects, it may become more difficult to establish a connection between impact and AID's project assistance as additional years pass;

- Our prior work with AID has suggested that when objectives and terms are not precise AID personnel can, over time, lose sight of why an activity is being undertaken. Ten years ago PCI developed the Logical Framework for AID, in part to solve this problem at the project level. If there is as much turnover in AID's evaluation staff in the next ten years as there has been in the past ten, there may be a tendency for "ex-post" evaluation to start collecting data that can easily be collected "ex-post" rather than to face the difficult problem of measuring impact.

BEST AVAILABLE COPY

B. THE AVAILABILITY OF IMPACT INFORMATION IN AID

AID is well aware of the fact that it does not have information on the impact of its projects. On the other hand, AID finds it difficult to explain to external groups why this is so, since the Agency has an ongoing Project Evaluation System and occasionally undertakes "special evaluations" as well.

AID's ongoing Project Evaluation System is designed and operated to produce information at the local (project) level to support management improvement for impact (Purpose and Goal) evaluation as well as management (Input to Output) evaluation. In practice it has been used primarily for management evaluation. The Mission-useful orientation of this system is such that little data on Outputs, Purpose or Goal reach AID/W. Data collected in the project is used in a Mission level evaluation. AID/W receives a report on this process, and the actions that are being taken based on the evaluation. The report provides minimal data on performance and impact. Thus, even to the degree that evaluations carried out under the Project Evaluation System have secured impact data, AID/W may not be aware of or have access to this information.

In addition to its regular life-of-project evaluations at the project/Mission level, the Agency has, over the years, conducted a number of "special evaluations". These evaluations have originated at the Mission level, in the regional bureaus, and in the central bureaus of AID. In general they have been carried out by individuals who are not connected with the project(s). In the mid-70s, AID also experimented with a pilot set of ex-post evaluations. The approaches and methods used in these supplementary evaluations have varied, as have the reasons for undertaking them.

During the course of this study, PCI examined several reviews of the products of the current project evaluation system, including one undertaken by Mr. Robert Hubbell for AID, and one undertaken by Dr. Samuel Daines of PCI in connection with a different engagement. PCI used these reviews to update a fairly complete review it undertook in 1974 in conjunction with the study of

the "Horizontal Expansion" of the Project Evaluation System. Several important points can be taken from this review:*

- AID attempts to evaluate the impact of its projects will more complex for projects where the objectives have not been fully and logically articulated, i.e., most pre-1971 projects and many others as well;

- Contrary to the thrust of current discussions in AID, past evaluation efforts by the Agency offer little evidence that securing impact information is difficult. Rather, the evidence these studies seem to offer is evidence that AID and its outside evaluators have made few attempts to evaluate project impact. Most evaluations concentrate on whether Inputs were delivered and Outputs provided, as well as on the quality of the management/implementation process found in the project;

- AID's past evaluations vary dramatically in terms of the quality of the methods used to gather data and in terms of the data used to formulate conclusions and recommendations. Past variability in these areas suggests that if AID were to mount a large program of evaluation for mature and terminated projects without quality control procedures and the ability to enforce them, AID would find that its evaluations again yielded some "garbage" as well as some good information;

- Anecdotal information about AID's evaluations suggests that evaluations which are "flawed" (i.e., used weak methods or data) may be just as influential in the Agency as those that rely on some methods and data as long as the recommendations turn out to be meaningful, or just "seem right", to AID. To the degree these anecdotes describe reality, they suggest an inability or lack of interest in AID in discriminating between professional and unprofessional work. The internal capacity to administer a "quality control" process for evaluation may be lacking.

PCI's review of AID's needs for information, the current availability of impact information and the basic characteristics of AID's Project Evaluation System have led us to conclude that the best approach for AID involves modification of the current project evaluation system rather than the development of a totally new system. The modification which is needed will retain the basic

---

* A fuller description of this update is provided in Annex A to this chapter.

project design concepts used by the PES as a basis for evaluation, but it will divide the responsibilities for securing evaluation data.

Primary responsibility for securing information on performance at the Input and Output levels, on the validity of assumptions at these two levels (i.e., those which affect Output and Purpose level performance), as well as process information on project management; participation of the target group in the design, implementation, evaluation and benefits of the project; the performance of implementing agents; host cooperation, etc., should all be retained as part of the existing Mission-oriented, project level system. The option to call for special evaluations at this level should also be retained. In addition, this level should continue to have a responsibility for identifying whatever Purpose of Goal level achievement the project/Mission observes during the life of the project. To this system should be added an explicit requirement to signal AID/W at the point in time when all Outputs are in place, and more rigorous work on impact evaluation can commence.

The responsibility for securing information on the Purpose, Goal and other higher level impacts of AID projects is the subject of Chapter VI of this volume.

BEST AVAILABLE COPY

THE AVAILABILITY OF INFORMATION FROM THE AID PES

PCI's review of the products of the PES and related special evaluations, during the course of this study, confirmed what AID already knew: There is very little evidence of project impact, even for projects that have undergone one or more regular, PES, evaluations. Further, PCI found that while many project evaluation reports discussed or speculated about impact in general terms, only a handful had actually attempted to measure project effects at the Purpose and Goal levels.

This lack of impact information has been taken by AID's senior staff as evidence of the need for a new system or program that will produce impact information. In examining AID's files, and in interviews with the Agency's evaluation staff, PCI's study team found that there were a number of possible explanations offered to account for the lack of impact data in evaluations conducted under the existing system. The following list, which was compiled from a variety of sources (not all of which were in agreement with each other), illustrates the range of difficulties AID perceives. Factors to which the lack of impact information was attributed included:

- A lack of effective demand for this information;
- Inadequacies in AID's current project evaluation system, or at least inadequate use of the existing system's potential;
- A presumed need to wait until projects are over to measure their impact;
- Inadequate time and resources and a lack of personnel who have the background required to plan and implement impact evaluations;
- Difficulties associated with the collection of field data on the impacts of interest to AID and Congress;
- Methodological difficulties associated with virtually all attempts to attribute observable effects to AID's development assistance.

The foregoing list is in one sense an index of Agency frustration. AID has been evaluating its projects for nearly ten years in a systematic way, and yet it cannot answer straightforward questions about its program. In this climate, the solution of "starting all over with a new system" has certain attractions. On the other hand, an AID decision to invent a new system simply to satisfy its felt need to take action might not be optimal. Hence, a more careful review of the current project evaluation system was indicated. PCI divided that examination into three parts: The adequacy of the AID Project Evaluation System, discrepancies between system potential and current practices, and exogenous factors affecting the system and its use.

1. The AID Project Evaluation System

In order to assess the adequacy of AID's existing Project Evaluation System for meeting AID's needs for impact information, PCI first examined the conceptual and procedural aspects of the system (without reference to the way the system is actually applied in AID). The conclusion PCI reached based on this analysis was that while the systems concepts (e.g., the idea of a hierarchy of objectives, the utility of objectively verifiable indicators, etc.) were adequate to provide a framework for evaluating projects in terms of their impact, both the system's focus (i.e., a Mission/management orientation) and current practice were not.

The key system concepts which were found to be supportive of AID's current information needs are summarized below:

- Project objectives must be specified if a basis for evaluation is to exist;
- Project objectives can be expressed as a logical hierarchy;
- The logical hierarchy of a project's objectives may be clear only when a number of levels of that hierarchy are expressed;*

* Limiting the expression of the hierarchy to four levels in a Logical Framework matrix is a procedural convenience rather than a conceptual requirement.

BEST AVAILABLE COPY

- Independent measures of success are required for each level of such a logical hierarchy;
- Objectives do not stand alone as a description of the project and its hypotheses. The assumptions made at each level must be taken into account in a full statement of each of the development hypotheses put forward by a project.

All of these statements, which apply to AID's current Project Evaluation System, are consistent with the theory and practice of impact evaluation as it is expressed in documents such as Evaluation: A Systematic Approach (Sage, 1979).

What the foregoing review of the AID Project Evaluation System's concepts and procedures suggests to PCI is that the current system may provide a partial basis for meeting AID's needs for impact evaluation. Since the conceptual aspects of the system are appropriate to the task, it would appear to be logical to incorporate them in AID's approach for securing impact data. The Mission-orientation of the present system, which does not call for detailed reporting to AID/W on evaluation findings, may cause AID/W some difficulty now that a valid need for this material has arisen. On the other hand, AID need not and probably should not eliminate the existing mission/management oriented procedures that are used by the PES. There is no evidence that suggests they are inappropriate for the function they serve which is helping projects and Missions generate and use evaluative information as a management tool during the life of a project.

2. AID's Evaluation System: Current Practice

While only a partial view of the effectiveness of AID's Project Evaluation System can be gleaned from Washington, several elements of current practice are discernable. As in the prior section, an examination of current practice can be divided into a review of the use of concepts and the use of procedures. Having determined that the procedures of the current system are not potentially supportive of AID's current need for impact information in

Washington, an extended discussion of procedural compliance would serve little purpose here. On the other hand, current practice with the use of the AID evaluation system concepts may provide some insight into AID's ability to build its impact evaluation approach on things the Missions are already doing.

On the positive side, PCI has observed and the Agency has documented the adoption of the evaluation system's ways of describing the levels of objectives, assumptions, measures, targets and the like as the *lingua franca* for Agency discussions of project designs and project evaluations. The reaction to this fact in AID is mixed. Some find the concepts of the system invaluable for sorting out what the projects are attempting to achieve; others are vehement in their dislike of the system's concepts, often for the very reason others appreciate them. The system's concepts about the logic of a project are fairly rigid, and proper application requires a level of discipline that is not generally viewed as being comfortable.

The system's concepts encourage AID personnel to begin the project design process by articulating a problem and the higher level objectives they want to achieve. From that point on, AID's Handbook Three envisions a project design as being logically derived. Once the objectives are set, all subsequent design elements should be selected based on an examination of what options will best lead to the achievement of the stated objectives. The "top-down" thinking advocated by Handbook Three runs counter to the way many of AID's personnel planned projects before the system was installed, i.e., by specifying what resources could be made available and then trying to determine how those resources could be combined to achieve some worthy objective.

During the period of system operation there have been a number of internal reviews of current practice. In addition, several contractor studies have reviewed current practice in the course of carrying out an assigned task. While these various summaries normally attempt to look at current practice in a positive light, they also tend to point out a number of striking weaknesses, such as the ones identified in the following statements:

BEST AVAILABLE COPY

- A good number of AID projects are still designed from a "resource-first" perspective. The system concepts are applied only after the design is complete. The Logical Framework is sometimes the last element of a Project Paper to be prepared;

- A review of Project Papers suggests that many AID personnel have not fully assimilated the system concepts. The Project Papers often violate system guidelines by defining more than one project Purpose, failing to identify plausible and objective indicators of achievement at each level of the project's logical hierarchy, failure to target those indicators that are offered, failure to fully specify alternatives to the proposed design, failure to fully specify the project's assumptions, and so forth;

- A review of the process reports submitted to AID/W after project evaluations are undertaken in the field suggests that while there has been some progress in terms of gathering and using evidence on performance at the Output level, project personnel do not take seriously the system requirement to measure Purpose level achievement and/or monitor the indicators of Goal level achievement;

- A review of the contracted evaluations that have been carried out for AID projects at the request of Missions, regional bureaus and central bureaus suggests that the concern with impact is minimal. For any of a number of reasons (e.g., lack of Mission interest, poor scopes of work, lack of skills on the part of the evaluation contractors, etc.) the majority of AID's major evaluation studies, to the degree they address impact at all, tend to "talk about" impact (i.e., offer observations, judgments and assertions) rather than "measure" impact;

- Commentary on all of the above from Missions and in AID/W has suggested over the years that AID/W lacks the capacity and/or will to insist on quality in the application of its system concepts and procedures. Poor designs (from an evaluation standpoint) are approved, e.g., approved projects can readily be found in which the project logic is best described as being "upside down and backwards". Projects are approved that pay only lip service to AID's requirement for an evaluation plan. Scopes of work for contracted evaluations that are poorly formulated and will not likely result in evidence concerning project performance and project impact pass through AID/W without a full quality review, even when they fail to present evidence to support the conclusions and recommendations they offer.

The picture of AID's Project Evaluation System painted by these summaries of current practice conjures up an image of a runaway horse. AID has attempted a great deal. It has been ambitious and aggressive in making evaluation an element of the project system. There have been positive results from this

BEST AVAILABLE COPY

approach, and it is to AID's credit that its evaluation system is far more advanced than the evaluation systems that are to be found elsewhere in the federal community, or for that matter in other development assistance organizations around the world. On the other hand, AID's ability to control quality at key points in the evaluation system has not kept pace with AID efforts to extend its evaluation approach to the full range of Agency projects.

3. Exogenous Factors that Affect System Performance

PCI found that prior reviews of AID's Project Evaluation System had pointed to four exogenous factors that appear to affect the AID Project Evaluation System. They are identified briefly below because they may also affect AID's attempts to institute an approach for securing the impact information it needs. The four factors involved are motivation, skills, time and resources.

Past reviews of AIDrs Project Evaluation System have pointed out that the reward system in AID is closely tied to those functions that result in the expenditure of resources. Personnel perceive themselves to be judged in terms of the number of projects that are designed and approved under their supervision. The same mechanism operates to encourage the initiation of evaluations and the type of service projects that DS and other central bureaus can develop.

Because of the distortion that is apparently introduced by this perception of the Agency reward system, its personnel view follow-through (i.e., high quality project implementation, critical review of evaluation products, careful monitoring of service arrangements and the like) as having relatively low priority and value. Accordingly, they focus their energies on the "new starts" that they believe higher level Agency management intends to encourage. As it proceeds with the development of impact evaluation procedures, AID may find that:

- Project level personnel do not share AID/W's interest in impact information on things that are over rather than information that has immediate bearing on "new starts";

BEST AVAILABLE COPY

- Those who administer the impact evaluation program may themselves be affected by the "new starts" psychology, and may give inadequate attention to the quality of impact evaluations once they are begun.

Good personnel tend to rise in the AID hierarchy. While this statement is obvious and consistent with everything we know about organizational behavior it may have implications for AID's attempt to introduce impact evaluation. The personnel who move up in the Agency start out by being good at their various technical specialties; they end up being managers who rarely have the opportunity to exercise their technical skills.

The net result is that direct hire personnel who have the technical foundation required to serve on evaluation teams are often too busy or, due to an overexposure to managerial tasks, have lost some of their confidence about the strength of their expertise. This phenomenon, which tends to make direct hire personnel unavailable for evaluation duties, also tends to discourage highly qualified specialists (including evaluation specialists) from joining the AID staff. Individuals in this category, who want to retain their strong technical credentials and well developed skills, tend to avoid situations in which they will be absorbed in managerial tasks.

AID has attempted to address the issue of evaluation skills at all levels of the Agency through its training programs, the short course on project design and evaluation, and its longer Development Studies Program. While these programs appear to be well received and effective in the short run, it is often the case that by the time an AID officer needs a particular skill, he or she is eight to ten months away from the training course in that skill and much has been forgotten. Thus, while AID's effort to train its own personnel has been fairly continuous, it has never succeeded in building the strong cadre of evaluators AID needs to independently operate a high quality program of evaluation.

Outside personnel who offer themselves as "experts" in evaluation and other technical areas that relate to specific evaluations are not always as skilled as AID would hope. If the evidence available in past evaluation reports is

BEST AVAILABLE COPY

taken as an indicators of the real "expertise" in evaluation AID secures by going outside the Agency, then AID may need to proceed cautiously. It is doubtful that AID will want to present to Congress evaluations that make recommendations based on assertion rather than fact. Yet this is what many of the evaluations AID has on file offer. It will take a concerted effort on AID's part to make both direct hire and outside personnel understand and respond to the fact that AID needs to be able to present evidence rather than opinions to Congress, and that impact evaluations that fail to produce evidence will not be acceptable to the Agency.

There are many demands on the time of AID staff members. These time demands, combined with the relatively low priority accorded to evaluation by most Agency staff, work against the aspirations of the Agency's evaluation staff. The time constraints that AID officers face on a daily basis have had some unfortunate repercussions for its evaluation efforts.

There is a tendency to want evaluations to be carried out quickly. Evaluation designs are sometimes developed that provide little in the way of usable data. These evaluations take two forms: The quick visit of an expert who uses his/her time to examine the situation and form a judgment which is presented to the project team, and what is being called the "rapid survey" approach to data collection. The former tends to be subjective and the latter to be based on a small number of observations (for which no measure is normally available of the degree to which the sample represents the intended study population).

A lack of financial resources has often been cited as the reason why a full-scale evaluation was not attempted for an AID project. In other cases, the cost of the evaluation was considered to be high given the cost of the project. In still other projects, the failure to define an evaluation plan (and set aside funds) at the time of project design has hampered evaluation efforts. While all of these reports of financial constraints may be valid, PCI is not convinced that a lack of resources is the central factor that explains the paucity of truly high quality evaluation in the Agency. Motivation and skills, combined with the time to carry out evaluations, are probably more binding

BEST AVAILABLE COPY

constraints on the current evaluation system. These three factors, rather than financial constraints, are also likely to be the main impediments AID will face in attempting to secure high quality information on project impact.

3

BEST AVAILABLE COPY

The Scope of Impact Evaluations

CHAPTER THREE

The term *scope* is used here to refer to the range of questions that are to be answered by an evaluation. The duration, intensity, cost, methods and staff required for an evaluation are all affected by what is identified in an evaluation scope.

PCI's discussions with AID personnel over the summer suggested that the question of the scope of the evaluations of mature and terminated projects was far from closed. PCI found that while Congress had implied, and AID had specified (in a memorandum from Shakow to the AAs on June 25, 1979) a limited scope for these evaluations, AID personnel tended to offer a much broader description of an evaluation scope. The following items make up the comprehensive list of "scope" elements suggest to PCI. Only Items one-three, and the portion of Item 4 that deals with negative impact were explicitly identified by the Shakow memorandum that discussed the information AID needs for its future Congressional presentations.

1. Evidence of impact (improvements and accomplishments) per Section 101 of the Foreign Assistance Act;

2. Evidence of target group impact (i.e., benefit distribution);

3. Evidence of the connection between these impacts and AID project action, specifically evidence that the effects are *stimulated by* or *associated with* or *caused by* AID's development assistance. (Note that these alternatives were explicit in the AID memorandum. Their methodological implications are quite different);

4. Evidence that the lessons learned through impact evaluation are applied by AID and the host countries;

5. Evaluative information concerning a number of types of effects projects may yield including:

BEST AVAILABLE COPY

- Second generation, as well as first generation effects;
- Unplanned, as well as planned effects;
- Secondary, as well as primary effects;
- Negative, as well as positive effects;
- Spread effects.

6. Evaluative information on the sustainability of project impact over the long-term (alternatively called "benefit continuation");
7. Information about problems in the project design or in the way the project was managed;
8. Methodological observations that can be employed by later evaluation teams;
9. Information pertaining to the external validity of the project approach, i.e., the assumptions/conditions that are required;
10. Information/hunches/partial data on subjects that go beyond the immediate evaluation framework; e.g., hypotheses about approaches that are not being used which might stimulate development in the project situation;
11. A formal articulation of the findings, conclusions and recommendations of the study;
12. A formal statement of the methodology used by the evaluators.

PCI's review of the list of candidate items led us to conclude that some of the potential elements of the evaluation scope might be better handled in other ways.

Item 3 was judged by PCI to be a management responsibility that goes well beyond that normally assigned to an evaluation team. The utilization of lessons learned from evaluation will normally be known by the time a final report is prepared. Thus, AID will find it difficult to include this item in the scope of studies it requests contractors, PASAs and RSSAs to perform. The item could be better handled if it were made the responsibility of the office(s) that oversee AID's program of evaluations of mature and terminated projects.

BEST AVAILABLE COPY

Item 6, benefit continuation, differs from basic studies of impact on the issue of training. In Chapter 3, PCI discusses the issue of timing and relates it to this potential scope item. Our conclusion is that AID's best approach involves low-cost follow-up studies of impact evaluations carried out close to the time at which the full intensity of impact can first be measured for a project.

Item 7, PCI concluded, is still a primary responsibility of AID's regular project evaluations, and it should remain so. The Agency's major effort to secure this type of data should be made with the current PES. Impact studies will invariably pick up some information on these issues, but they should not be a primary focus of the evaluations of mature and terminated projects, except in so far as these evaluations provide data on the internal and external validity of the project hypotheses.

The remainder of the candidate items for inclusion in the scope of project impact evaluations were judged to be proper scope items. In the long run, AID should expect its project impact evaluations to address all of the questiosn raised under these legitimate scope items. However, this full set of impact evaluation scope items may not be appropriate for the first project impact evaluations AID carries out.

The problem AID faces in making a decision about whether to add items to what appears to be a "minial" scope for its evaluations of mature and terminated projects has more to do with technical and management capacity than with cost. It will be difficult enough, at first, to simply do a good job of securing the minimum information needed about project impact. On the other hand, AID may see these evaluations as its one good opportunity to try to answer questions that go beyond the minimum items in an evaluation scope. Further, a number of the items on the list could be added to a minimal scope at very little cost.

Unfortunately, some of the supplemental items, such as unplanned effects and some forms of spread effect may prove difficult to measure. And every addition

BEST AVAILABLE COPY

to an impact evaluation scope can, potentially, reduce the degree of focus in the study and diffuse the energy that will be needed to gather impact data. In its first project impact evaluations AID should concentrate on the first three scope items, which constitute the minimum but critical set of questions AID must answer for Congress.* AID should itself be able to judge when it begins to have some success in securing this basic information. At that time, the Agency should begin to add other scope elements as rapidly as it is able, while still maintaining an ability to secure high quality information on each scope item it includes.

Annex A to this chapter provides a detailed description of each of the main items, and many of the subtopics, on the list of candidate scope elements presented above. It also discusses the implications of adding various items to a "minimal" scope for the evaluation of mature and terminated projects.

---

* It will also be important to keep track of methodologies and methodological lessons during this period. But it may be more appropriate to do this in a centralized book of "lab notes" than in the evaluation reports in the beginning.

BEST AVAILABLE COPY

CANDIDATE EVALUATION SCOPE ELEMENTS & THEIR IMPLICATIONS

Seven major candidate scope items are examined in this Annex:

1. What impact AID projects have yielded;
2. Whether such impact as is found was associated with or caused by AID development assistance;
3. Whether impact will be sustained over the long run;
4. What lessons a project can offer to guide future Agency efforts;
5. What recommendations can be made based on the evaluation of mature projects: What range of actions can be realistically considered, and to whom these recommendations should be directed;
6. What actions AID and host governments have taken based on the lessons learned from successful and unsuccessful projects;
7. What quality of information is to be gathered and used to address the six questions listed above.

To place the items on this list in current perspective, it is worth noting that only Items 1, 2 and 6 were explicitly mentioned in AID's June 25 memo from the Acting Director of PPC to the Assistant Administrators. These are the three items that are currently seen as being central to the Agency's dialogue with Congress. In the following paragraphs, each question on the list given above is examined separately to determine whether it should be incorporated in a general statement of the scope of project impact evaluations, or handled by the Agency in some other way.

1. Determining What Impact AID Projects Have Yielded

Different people and offices in AID mean different things when they use the term "impact". This is a potential problem. Any attempt to standardize the

BEST AVAILABLE COPY

way in which the term is used is likely to meet some resistance. On the other hand, if AID fails to develop an unambiguous operational definition of "impact", it's possible that no two impact evaluations will address the same questions.

The term was found to have two basic definitions in the AID context. It connotes both:

- The achievement, through projects, of the specific objectives set forth in the Foreign Assistance Act;
- Achievement of the development objectives identified by projects; specifically, the achievement of project Purpose and Goal.

In the course of this engagement, PCI undertook a systematic comparison of the objectives and indicators of impact identified in the FAA and the Purpose and Goal statements of 1,940 AID projects in the DIS database. This analysis indicated that at present there is an 85% overlap between these two definitions. Of the 1,940 projects in the database (all of which were active on or after October 1974), a total of 1,659 stated objectives at the Purpose and/or Goal level that are equivalent to one or more of the 27 objectives and indicators PCI found in the current version of the FAA.

The magnitude of this overlap suggests clearly that an approach to securing impact information that focuses on project Purpose and Goal achievement will potentially capture 100% of the explicitly stated project expectations concerning impact in mandate terms. An approach that focuses on impact as defined in the FAA would ignore only a small portion of AID's portfolio.

PCI's comparison of the FAA to AID project Purpose and Goal statements further suggested that AID's approach to securing information on project impact might well be a fairly standardized one. Of the 1,659 projects that claim impact in mandated terms, over 80% state that the impact the project will yield corresponds to one of seven of the 27 objectives and indicators PCI found in the FAA. The strong tendency for AID project impacts to cluster

BEST AVAILABLE COPY

around a few objectives/indicators suggests that the majority of AID's projects could be evaluated using a handful of impact measurement approaches.

Table II.A.1 is a Logical Framework of the objectives and indicators PCI found in the first portion of the FAA. The material in the Act was arranged into a hierarchy of objectives, and the major indicators identified by Congress were placed parallel to the objectives they seemed to address. On the Logical Framework there are 27 numbers, placed in circles. Each of these numbered statements was considered by PCI to be a separate objective or indicator, even though the language of the Act often combined items in single sentences.

Table II.A.2, its companion, summarizes the apparent overlap between these Congressionally mandated objectives and the portion of AID's total projects found in the DIS database. This database, which includes an abstract of project Purpose and Goal, contains 1,940 projects, all of which were alive on or after October, 1974. The database thus includes virtually all of the AID projects for which Logical Frameworks were developed, i.e., the database includes the vast majority of projects planned during 1971-72 when the PES was installed. It also includes a number of projects with earlier start dates. (PCI did not go beyond this database in its analysis of the overlap since our prior work with DS/DIU had indicated that projects not on the DIS tend to have poorly formulated/articulated objectives, which would make them difficult to analyze and which may cause AID difficulty if it attempts to assess their impact.)

In Table II.A.2, the first column identifies each of the 27 numbered objectives and indicators from Table II.A.1. It also identifies two other categories PCI used in its analysis: (a) Statements that were impact objectives, but not mandated impact objectives and (b), statements that were logically below the mandate impact level but appeared to us to lead toward mandate impact. Column 2 of the table shows the absolute number of times each type of statement was found in the 1,940 projects examined. Columns 3 and 4 provide subtotals of the number of times each of the 27 statements from the Logical

BEST AVAILABLE COPY

TABLE III.A.1:

A LOGICAL FRAMEWORK OF THE OBJECTIVES AND INDICATORS FOUND IN THE INTRODUCTORY SECTIONS OF THE FOREIGN ASSISTANCE ACT

	NARRATIVE SUMMARY	OBJECTIVELY VERIFIABLE INDICATORS	ASSUMPTIONS
HIERARCHY OF AID'S MANDATED OBJECTIVES	SUPER-GOAL 2 (Section 101, quoted) (1) "A principle objective of the foreign policy of the United States is the encouragement and sustained support of the people of developing countries in their efforts to acquire the knowledge and resources essential to development and to build the economic, political and social institutions which will improve the quality of their lives".		
	SUPER-GOAL 1 (Section 101, quoted) "United States development cooperation policy should emphasize four principal goals: (2) (1) The alleviation of the worst physical manifestations of poverty among the world's poor majority (3) (2) The promotion of conditions enabling developing countries to achieve self-sustaining growth with equitable distribution of benefits (4) (3) The encouragement of development processes in which individual civil and economic rights are respected and enhanced, and (5) (4) The integration of the developing countries into an open and equitable international economic system".	(6) (B) Reduction of infant mortality (C) Control (of) population growth (7)	
	GOAL (Section 102, paraphrased) (8) The poor majority of the people in developing countries *increase their income and* (their access to public services which will enable them to) *satisfy their basic needs and lead lives of decency, dignity and hope.* (9)	(D) Promotion of greater equality of income distribution including such measures as more progressive taxation and more equitable returns to small farmers. (10)	
	PURPOSE (Section 102, paraphrased) (11) The poor majority of the people in developing countries *participate in a process of equitable growth* (through productive work) *and influence decisions* (12) *that shape their lives.*	(A) Increase(s) in agricultural productivity per unit of land through small-farm, labor-intensive agriculture. (13)	
	SUB-PURPOSE STATEMENTS (Section 102, paraphrased) The poor majority of the people in developing countries: (14) (1) *Are engaged in productive work* (15) (2) *Use their expanded access to their economy* (through public sector services and institutions at the local level) (16) (3) *Increase their participation in making decisions that affect their lives* (17) (4) *Improve their lives through their own efforts*	(E) Reduction(s) of rates of unemployment and underemployment (18) (F) Increase(s) in literacy (19)	
AID'S MANAGEABLE INTEREST	OUTPUTS (Section 102 paraphrased) AID funded projects: (1) Lead to *increased labor-intensive production* (20) and the *use of appropriate technology* (21) (2) *Expand productive investment* (22) and *services* (23) out from major cities to small towns and rural areas (3) Expand the access of the poor majority to their economy through *services* (24) and *institutions at the local level* (25) (26) (4) Provide opportunities for the poor to improve their lives through their own efforts (27) (5) Involve the poor in increasing their participation in making decisions that affect their lives		"Congress finds that the efforts of developing countries to build and maintain the social and economic institutions necessary to achieve self-sustaining growth and to provide opportunities to improve the quality of life for their people *depends primarily upon successful marshalling their own economic and human resources*".
	INPUTS AID's programs and projects		"The magnitude of (the) efforts required exceeds the resources of developing countries".

* Section 102 states: "The President shall assess the commitment and progress of countries in moving toward the objectives and purposes...by utilizing criteria, including but not limited to...the items shown in this column of the Logical Framework".

NOTE: Circled numbers identify objectives and indicators PCI reviewed against projects.

BEST AVAILABLE COPY

TABLE III.A.2: Analysis of the Frequency of Use of Each FAA Mandated Objective/Indicator in 1,940 AID Project Purpose and Goal Statements.

OBJECTIVE OR INDICATOR	TOTAL NUMBER OF TIMES LOCATED	NUMBER OF TIMES USED IN PURPOSE	NUMBER OF TIMES USED IN GOAL	TOTAL # OF USES AS % OF ALL USES
1. Any use of Super-Goal 2 words	10	0	10	.4
2. Super-Goal 1, (1)	151	10	141	6.0
3. Super-Goal 1, (2)	58	0	58	2.3
4. Super-Goal 1, (3)	1	0	1	.0
5. Super-Goal 1, (4)	5	1	4	.2
6. Super-Goal mortality indicator	7	0	7	.3
7. Super-Goal population indicator	49	2	47	1.9
8. Goal income objective	238	9	229	9.5
9. Goal basic needs objective	127	2	125	5.0
10. Goal income distribution indicator	121	6	115	4.8
11. Purpose equitable growth objective	8	1	7	.3
12. Purpose Participation objective	1	0	1	.0
13. Purpose production indicator	372	77	295	14.8
14. Sub-Purpose objective (1)	57	6	51	2.3
15. Sub-Purpose objective (2)	6	3	3	.2
16. Sub-Purpose objective (3)	19	3	16	.8
17. Sub-Purpose objective (4)	21	12	9	.8
18. Sub-Purpose employment indicator	40	4	36	1.6
19. Sub-Purpose literacy indicator	11	4	7	.4
20. Output (1) labor intensive	5	3	2	.2
21. Output (1) appropriate technology	21	19	2	.8
22. Output (2) expand investment	18	7	11	.7
23. Output (2) expand services	11	5	6	.4
24. Output (3) access to services	376	218	158	15.0
25. Output (3) access to institutions	453	378	75	18.0
26. Output (4)	26	21	5	1.0
27. Output (5)	20	9	11	.8
Objectives were parallel to mandate objectives, but they were not the same.	260			10.3
Objectives were below the level of the mandate objectives, but appeared to lead toward them.	21			.8
TOTAL OBJECTIVES & INDICATORS LOCATED IN 1,940 AID PROJECTS IN THE DIS DATABASE: 2,513				

BEST AVAILABLE COPY

Framework were seen at the Purpose or Goal level. The final column in the table indicates the number of times each type of statement from the Logical Framework was observed as a percent of the total number of observations of statements from the Logical Framework.

In addition to the two basic definitions of impact cited above, AID is also interested in both the various types of effects (impacts) projects can produce (e.g., planned and unplanned effects) and in the specific sector or macro-goal areas in which these impacts are felt (e.g., changes in health status, income, etc.). In the following section, PCI examines first, the types of project effects that concern AID, and secondly, the specific sectoral and macro-level changes in the lives of poor people that AID expects will result from its projects.

a. The Types of Effects AID May Want to Assess in Impact Evaluations

AID discussions of the types of effects projects can yield tend to revolve around a set of descriptive terms that can be applied to a variety of examples. Over time, some of these descriptors have come to be used as if they had a clear and unequivocal meaning in the absence of a specific example. For other terms, definitions are hard to come by, but examples abound.

The situation is one in which an attempt to develop a taxonomy of project effects would not only be time consuming; it might also tend to make the subject more complex. An alternative and relatively easy way of dealing with the various types of effects that concern AID involves recognizing that any example can be described along a number of dimensions. Put another way, it should be possible to pick one adjective from each of the following sets to describe any example of a project effect that might come to mind:

- First generation/second generation;
- Planned/unplanned;
- Primary/secondary;

BEST AVAILABLE COPY

- Positive/negative;
- Spread or not.

The sum of the descriptors selected to characterize any given example places the project effect in question in a fairly well defined category. The number of possible combinations, or categories, is too large to allow us to discuss every possibility here.

A review of the documentation for a typical AID project will normally yield a composite statement of those impacts that are first generation, planned, primary and positively valued. This composite is an articulated subset of all of the impacts the project might yield. In PCI's judgment, it is this subset that AID must, at minimum, examine in its project impact evaluations.

This minimum position is not necessarily the best choice for AID. Evaluations that examined only the subset of first generation, planned, primary and positively valued effects would leave many questions unanswered. If AID decides to exceed this minimal position, it still may not want to examine all of the possible effects projects may have. That is, AID might want to add to the minimum set an examination of spread effects and negative effects, but leave aside second generation effects, unplanned effects and secondary effects. Other combinations are obviously possible. In selecting the set of effects to be examined in an "optimal" scope for project impact evaluations, AID will need to consider the additional level of effort involved in identifying and measuring each type of effect it adds to the minimal set.

The following paragraphs provide brief definitions of each of the paired types of effects listed above. In addition, the implications of including types of effects that go beyond a minimal scope are discussed.

BEST AVAILABLE COPY

(1) First Generation/Second Generation Effects

First generation effects are effects caused by the project action. Second generation effects are effects for which the project's first generation effects are the cause. (As suggested above, these effects can be planned or unplanned, primary or secondary, positive or negative, and in some instances the term *spread* will also apply, e.g., the replication of a project effort is both a second generation effect and a spread effect.)

If AID elects to include the assessment of second generation effects in "optimal" scope for impact evaluations, it should expect that the time frame for evaluation may need to be extended. The basic characteristic of second generation effects, from the point of view of evaluation, is that they cannot be observed until after the primary effects have occurred. PCI's review of the question of evaluation timing suggests that, in general, first generation effects will be observable in the period from 0 to 18 months after project Outputs have been provided. Some second generation effects may be observable very early; others could lag and not be observable until well after the last of a project's first generation effects are felt, e.g., some time in the second or third year after the project's Outputs are provided.

A second difficulty that may be associated with second generation effects and their measurement may occur if the effects were not anticipated. If second generation effects from a project are unplanned, or if they could have been anticipated but were not, evaluators will have to undertake a two-step process in order to assess them: (a) Identify a project's second generation effects and (b) measure them. The process of identifying effects that were not articulated by the project adds a layer of complexity to the evaluation task, as is pointed out below.

(2) Planned & Unplanned Effects

Planned effects are the articulated (documented) changes expected as a result of project action. However, the line between planned changes and unplanned

BEST AVAILABLE COPY

ones is not always a tidy one. In order to capture all effects that do not meet the definition of a planned effect, our use of the term "unplanned effect" must incorporate both those effects that could not have been anticipated and effects that could have been anticipated, but were not articulated (documented).

These two types of unplanned effects differ from an evaluator's point of view. At the beginning of an evaluation period, an evaluator may be able to carry out an analytic effort to identify the unarticulated effects that can be anticipated, and thus build ways to measure them into the evaluation design. For the set of effects that cannot be anticipated, this process will not work.

Project effects that could not have been anticipated must be handled in a different way. In effect, these impacts must be "discovered" in the course of an evaluation. The evaluator must take three steps in a process that leads to the assessment of unanticipated effects. First, the evaluator must notice aspects of the project situation that are not on the list of specific changes she/he is attempting to measure. Second, the evaluator must discern whether these aspects of the situation have changed during the project period. Finally, the evaluator must define a plausible connection between project action and the unanticipated changes that have been observed. Each step in this discovery process poses its own difficulties.

While inclusion of unplanned effects in an "optimal" scope for AID's project evaluations might not increase the total time required, it would affect the level of effect required. In order to identify and measure the unplanned effects of its projects, AID would probably find that (1) more sophisticated fieldworkers would be required than if only planned change is measured and (2) more staff, overall, are required to complete the tasks within the evaluation scope.

BEST AVAILABLE COPY

(3) Primary & Secondary Effects

Primary effects and planned effects often mean the same thing in a project description. The distinction made here between primary and secondary effects relates to who/what is affected by project action. Primary effects, for the purposes of evaluation, would include all changes in the peoples, areas, institutions, etc., that were explicit targets of project action. Secondary effects would, in a parallel manner, subsume all changes that affect the target groups/areas in dimensions that differ from what was planned (e.g., negative side effects) and/or changes that affect populations/areas other than the target populations/areas (e.g., spontaneous adoption of a project approach by a different group of people). Spread effects, which are of special interest to AID (and receive separate treatment below) are often secondary effects.

If AID includes secondary effects as part of the optimal scope for its project impact evaluations, the effects on its evaluation designs could be multiple. Secondary effects may occur well after the primary effects are observable, thus extending the time frame for evaluation. Some secondary effects may fall in the unplanned category, thus requiring that a process be instituted to identify them before measurement can begin. And finally, where secondary effects are felt in populations/areas other than those targeted by the project, the number of sites visited during an evaluation may increase, implying a greater level of effort overall.

(4) Positive & Negative Effects

While the other descriptors of a project effect can be defined abstractly, great care needs to be taken in addressing the positive-negative dimension. Positive and negative are terms that convey our value judgments. They take on meaning only in a context, and not everyone will value a given effect in the same way. Thus, it is fully possible that AID will find situations in which the Mission and the host government, or the central ministries and the villages, do not agree on how an effect is to be valued. From an evaluation

BEST AVAILABLE COPY

perspective, the prudent approach would be to document effects and then indicate how those effects are valued by key actors in the project situation. The labels *positive and negative* might safely be applied where all key parties agreed on the value judgment for a given effect. Where no concensus is found, AID's evaluators would probably do the Agency a service by pointing this out, rather than trying to apply a label to which some parties do not subscribe.

A cursory review of AID's project documents suggests that the effects identified by projects tend to be effects that AID values positively. If the Agency elects to include within an optimal scope for its project impact evaluations, two issues will need to be addressed: (1) The degree to which the effects AID views positively are also positively valued by other actors in the project situation and (2) a way to identify and measure negative effects (or costs) will need to be developed. In order to examine the first issue, AID will need to secure value judgments from those involved in the project or in the area in which a project is mounted. In some cases, these actors will be familiar to AID, e.g., the ministry with which AID is working, individuals involved in project implementation at the site, etc. Other actors may not be so familiar to AID, e.g., local political representatives, people working in sectors that are indirectly affected by project action, etc. If AID is to understand how each of these actors values a given project effect, it will need to both identify the actors (taking care not to exclude groups that are expected to differ from AID in their views) and solve the problem of how to elicit judgments from these groups. One obvious problem will be a bias problem: People may tend to tell AID what it wants to hear rather than face the discomfort of indicating that they negatively value something AID is working hard to produce. The second issue involves many of the problems associated with the first: Additional complexity is added when AID must "discover" the effects it wants valued, as was the case for unplanned effects discussed above.

(5) Spread Effects

Spread effect is a term AID uses to refer to several patterns of effect that may also be primary or secondary, planned or unplanned, etc., in character. The

BEST AVAILABLE COPY

term is used to refer to two types of multiplier effects: Planned replication and the spontaneous adoption of one or more aspects of a project. It also characterizes some of the "ripple" effects of projects, which also tend to be secondary in nature.

When a spread effect is identified, it usually involves either the adoption of a process used in a project, or the adoption of both a process and some or all of the specific content of a project. For example, spread effects from a cooperative marketing project could involve (a) the use of the cooperative mode to serve other purposes such as savings and loan, or (b) the planned or spontaneous development of marketing cooperatives outside the target area/group. While spread effects that involve the adoption of both a process and the specific content from a project are likely to be recognized quite easily, instances in which only a process is reapplied may be somewhat more difficult to identify. Further, AID should recognize that the "state-of-the-art" in process assessment/measurement is not highly advanced. This is particularly true for cases in which the process of interest involves "participation" (of the sort AID is mandated to encourage through its projects).

If AID elects to include "spread effects" in its optimal scope for project impact evaluations, the effect will be similar to that expected if secondary effects are added. The number of personnel and their skill levels may need to increase, and the evaluation will likely cover a more extensive area/population than would be the case if "spread effects" were not included.

In summary, it appears that project impact evaluations which include the assessment of second generation effects, unplanned effects, secondary effects, spread effects and those effects that are valued negatively may take longer to execute and require the attention of more personnel, a higher percentage of which would need to have previous experience/training in identifying and measuring effects that go beyond those formally planned and articulated for a project.

Nevertheless, for AID, the definition of an optimal scope for project impact evaluations may well involve requiring that all of AID's questions are answered

BEST AVAILABLE COPY

and, hence, that all effects not listed in the minimal scope be listed in the optimal scope. If AID does develop a standardized optimal scope that includes all project effects, it may want to develop a mechanism that can be used to delete items on a case-by-case basis. Thus, for example, if AID were evaluating a series of projects of the same general type, it might wish to use the minimal scope for all but one, and with a single project in the series apply the optimal scope.

b. The Substance of Mandate Impact & Other Project Impacts

As suggested by the foregoing discussions, mandate impact is a subset, albeit an important subset, of all impact yielded from AID projects. Every project identifies its impact expectations in its Purpose and Goal statements. By definition, Purpose begins the chain of important changes AID expects to render in a project situation through its development assistance (i.e., the provision of Inputs which are transformed by AID's implementing agents into Outputs during the life of the project). While some projects identify their project Purpose in terms that are consistent with the AID mandate, others do not. Many identify impacts that converge with the mandate in their Goal statements, or imply that they are Super-Goals. In the following paragraphs, the content of AID project impact statements is examined. The changes sought, in mandate and other terms, will be the substance that impact evaluations attempt to assess.

(1) Mandate Impacts

The question of what specifically is to be assessed to determine whether AID's projects have mandate impacts cannot be answered with a simple list. PCI's review of Congressional and AID statements at first suggested that there was a lack of congruence in what each of these parties saw as "mandate impact". The two documents that best illustrate that lack of convergence are the 1978 version of the FAA and AID's February 1978 report to Congress on "Proposed

BEST AVAILABLE COPY

Critiera and Factors for Assessing Country Performance". Table II.A.3 identifies some of the areas where AID and Congressional statements appear to take different directions.

A more detailed review of the Congressional and AID documents, from an historical perspective, suggested two reasons for this apparent lack of congruence:

> In the 1975 version of the FAA, Section 102(d) required that the President establish appropriate criteria to assess the commitment and progress of countries toward the objectives of the FAA. One of the uses identified for these criteria was as a guide for allocating AID's resources. The office charged with formulating an appropriate set of criteria was PPC/PDA. The criteria which were developed and submitted in 1978 focus on the allocation use of such measures. It is not clear to PCI that AID fully assessed the criteria it submitted from the perspective of their second potential use as mandate impact indicators from an evaluation perspective.
>
> In AID's response to the 1975 version of the FAA (which is found in its report "Implementation of 'New Directions' in Development Assistance"), the Agency questioned Congress concerning its references to "direct" assistance to the poor (and hence direct measures of such assistance). AID pointed out to Congress that the Agency actually worked through public and private sector institutions and with policy changes, not in the villages. AID asked Congress for a "definitive statement" on this topic.
>
> Meanwhile, AID began work on the development of a set of criteria for assessing performance and commitment. In February 1977 it provided Congress with a progress report and by the same time the following year it had submitted its criteria to Congress. The criteria development exercise appears to have been based, in part, on the assumption that the Congressional "definitive statement" on "direct assistance" would lead to a revision of the language of the FAA. In 1978, Congress revised the FAA. *The revision maintained the language of "direct" assistance and restated the original list of direct measures as minimal criteria for assessing performance and commitment,* thus continuing the original construction of the law on this question.

The fact that there is a lack of congruence between Congressional and AID statements of what constitutes mandate impact (and what are appropriate indicators of performance) suggests that AID's Office of Evaluation has several options concerning the substance addressed by its project impact evaluations:

- AID can evaluate its projects in the mandate terms used in the 1978 version of the FAA, without additions;

BEST AVAILABLE COPY

TABLE III.A.3:

APPARENT DIFFERENCES IN THE MEANING OF THE TERM MANDATE IMPACT IN CONGRESSIONAL & AID DOCUMENTS

- *Differences in emphasis concerning the point at which AID's impact will be felt, and hence most appropriately measured.* Congress has focused on "direct" impact on the poor majority of the people in developing countries. The criteria it identified for assessing progress toward the objectives it identified were all stated in terms of measures of changes in the lives of people. AID has responded that its does not work directly with people, but rather with public and private intermediary institutions and on changing LDC policies. The criteria for assessing progress it submitted to Congress in 1978 focus extensively on government spending and government policies. The AID criteria also include the direct measures identified by Congress, but AID's list often gives these measures a relatively low priority, e.g., below "top three" indicators, which AID suggests are the key indicators for each of its seven criteria. If AID focuses on the "top three" indicators for its criteria in impact evaluations it will, in some instances, skip measuring criteria identified by Congress.

- *Differences in the number and types of criteria/indicators.* Not only does the point of impact differ from Congressional to AID documents, the absolute number of criteria/indicators differs significantly. Congress listed a minimum of six items that were to be measured to show progress toward mandate objectives. AID's 1978 submission identifies seven criteria and a total of 37 indicators. This difference, when considered as a total measurement burden, is substantial. One additional difference, that may partially equalize the implied measurement burden, is that the two lists suggest the use of different data. The Congressional list suggests that primary data on its indicators are needed, and that data is to be collected/analyzed only for those who fall inside the definition "poor majority". The AID list, on the other hand, suggests that a good deal of secondary (host government) data will be re-examined to generate the ratios, etc., that AID has identified.

- *Differences in the definitions of certain critical terms, such as participation.* Congress uses the term "participation" frequently. It speaks of participation in direct terms: It envisions poor people making the key decisions that affect their lives, and influencing local institutions, and of "democratic participation in private and local government activities". AID's use of the term is somewhat different. AID discusses participation in decisions, and suggests that "decisions concerning activities to be carried out are made, preferably by those benefited (that is, the poor), or if not, at least with effective consultation and substantial acceptance by those benefited". The difference is one of tone, or style, but it is an important one from an evaluation perspective. The ways in which AID would measure to ascertain whether a democratic process was in place would be different than the measurement approaches used to determine whether the poor had been consulted concerning a decision.

- *Differences in the means used to achieve stated impact objectives.* Once again, this is a difference of tone/style, but it too has implications for evaluation. The Congressional documents, if they do not weight public and private sector approaches evenly, appear at points to favor private sector and self-help modes. AID's list of criteria are heavily weighted toward public sector approaches, though this list does include some indicators of private investment and self-help.

- *Differences in the "vision" of the end state of the development assistance effort.* To grossly oversimplify the picture painted by the total statement of objectives/criteria in Congress and in AID, it appears that Congress envisions the development process as resulting in a highly participatory, self-motivating, economic and social system where people invest in themselves and work together to solve problems. The AID portrait is one of healthy public sector organizations that bring goods and services to the poor people and succeed in raising their standard of living. While such differences in "vision" may not in fact exist, the statements made and criteria suggested by the two parties suggest the possibility.

BEST AVAILABLE COPY

- It can evaluate them by following the list of criteria it forwarded to Congress in 1978;
- It can "start over" and examine Congressional intent and come up with a set of performance criteria that seem appropriate from an evaluation (as opposed to an allocation) perspective.

The choice is a difficult one from several perspectives, but it needs to be made. AID itself may consider the first option too simplistic, or inadequate to measure what AID projects actually accomplish. If it chooses the second option, the evaluation staff may find itself gathering and analyzing data that will yield little by way of guidance for the designers of future projects. The third option is also a difficult choice, since it involves both "homework" and the bureaucratic snarls involved in either challenging the first set of criteria or creating space for the existence of two sets of criteria--one on commitment and the other on progress, or some other mix, e.g., one for allocation and the other for evaluation.

While PCI cannot make this choice for AID, it does appear to us that since Congress is the primary audience for impact evaluation identified by this feasibility study, AID should consider seriously the implications of evaluating its projects in terms that differ significantly from those set forth in the law. Deviations of this sort will not necessarily make AID projects "look good" nor will they help AID learn what is really required to implement the "New Directions" on the ground.

To facilitate AID's consideration of its three options, PCI took an initial "homework" step and prepared a skeleton Logical Framework of Congressional intent as documented in the FAA, Table II.A.1. We then compared this Logical Framework's statements to existing project objectives, rather than to AID discussion papers on mandate impact, and came to the realization that AID's projects are in better agreement with the FAA's version of the mandate than are AID/W's discussion papers and criteria.

BEST AVAILABLE COPY

(2) Effects that are Explicitly Identified as Intermediate Steps in a Chain that Leads to Mandate Impact

As suggested earlier, in some projects either Purpose or both Purpose and Goal will identify impacts that have bearing on AID's mandate, but do not reach to the level of mandate impact. In the course of its project impact evaluations, AID may wish to assess performance against this type of objective. Included among the intermediate effects AID may wish to assess are efforts to build and strengthen the various agencies that it expects will eventually provide services to the poor, service and/or research efforts carried out by AID's contral bureaus (PPC and DS), management improvement and other training efforts for both AID and LDC audiences, education programs that reach beyond literacy to provide an upgraded human resource capacity in the LDCs, etc.

While these intermediate objectives will tend to cluster into families, PCI does not expect that AID will be able to "standardize" these impact statements in a way that will allow their inclusion in a general statement of what is to be measured in project impact evaluations. Rather, we expect that the terms of reference for the impact evaluations that cover projects which have intermediate objectives will require the inclusion of an instruction to this effect:

> The evaluator is expected to specify each of the discrete levels of impact between project Outputs and mandate impact and, following AID's review of the approaches proposed for assessing impact at each of these levels, the evaluator is expected to apply these measurement approaches to determine the impact of AID's project at each intermediate level, as well as at the project's mandate impact level.

Behind this type of instruction AID will probably need to supply potential evaluators with information on common families of intermediate impact and the Agency's experience in assessing these types of impact.

BEST AVAILABLE COPY

(3) Effects that are Not Explicitly Identified as Being Linked to Mandate Impacts

While some projects focus on the creation of intermediate effects that are explicitly defined as being linked to mandate objectives, AID has other projects where no such explicit connection is identified. The fact that these projects do not claim a connection to mandate impact cannot, however, be taken as an indicator that no such connection exists.

To the degree that AID decides to include objectives of this sort within the framework of its response to the Congressional mandate, and evaluate the impact of these projects, it will be important to examine the degree to which these objectives group into clusters. If AID finds that these objectives do cluster, it should expect that the evaluation methodologies it needs to assess these projects also cluster.

c. Evaluating Impact in Mandate Terms

As noted in Section Three, not all projects define their expectations concerning impact in mandate terms. PCI classified AID projects into three clusters with regard to mandate impact:

- Projects that claim "mandate impacts" as a direct (Purpose or Goal) result in project action;
- Projects that claim "mandate impacts" as a distant (Super-Goal) result of project action;
- Projects that make no claim to "mandate impacts".

PCI's examination of the AID databases suggests that there are projects in each of these categories. The feasibility of conducting impact evaluations, in mandate terms, differs for each cluster.

BEST AVAILABLE COPY

(1) Projects That Claim Direct Mandate Impact

Projects in this cluster include all efforts that state the project's impact expectations in mandate terms. Three subcategories exist within this cluster: Those which claim mandate impact at the Purpose level only, those which claim mandate impact only at the Goal level, and those that claim mandate impact at both of these levels. Projects in any of these subcategories can be evaluated in terms of their mandate impact without substantial preparatory work, i.e., review/restatement of the project's objectives. However, since the quality of project designs varies, it may be necessary to clarify the indicators of achievement, project assumptions, etc., before commencing evaluation. AID's recommended procedure for pre-evaluation design clarification should suffice as a method of preparing these projects for evaluation in mandate impact terms.

(2) Projects That Claim Distant Mandate Impact

Projects in this category include all efforts which, while not claiming mandate impact at either the Purpose or Goal level, can be linked to mandate impact objectives by extending the logic of the project, i.e., by adding one or more Super-Goal levels to the project's hierarchy of objectives. For fairly recent projects (those prepared since the development of AID's Handbooks 2 and 3), the information needed to define the linkage between project Goal and mandate impact objectives should be available in the DAP (now CDSS) and the project documents (PID, PRP where done, and PP). Both of the Handbooks used by AID call for spelling out the connection between Mission level objectives and project level objectives. Normally such connections are articulated in the narrative description of projects (and are often found in the early pages of a PP).

While the information required to make a connection between project level objectives and mandate impact objective tends to exist somewhere in the Mission/project documentation, these linkages need to be systematically

BEST AVAILABLE COPY

articulated (and indicators of achievement defined) before projects in this category can be considered "ready" for evaluation in mandate impact terms. AID's blue book on options in the use of its Logical Framework approach provides the basic tools AID will need to apply in articulating the logic of the Goal to mandate impact linkages in these projects. Figure II.A.1 provides an illustration of the type of objectives that will need to be defined, beyond the Goal level, in order to prepare this cluster of projects for impact evaluation in mandate terms.*

It is important to point out that this type of approach is likely to be needed for more than the 21 projects we identified as stating impact objectives below the mandate level (Table II.A.2). A manual examination of those which claimed mandate impact at the Purpose or Goal level suggested to us that a good number had left out one or more "steps in the logic" of the project, i.e. one or more levels between Outputs and mandate impact which may need to be reconstructed prior to evaluation.

The preceding paragraph discussed an approach to linking project level objectives to mandate impact as a pre-evaluation exercise. This exercise will be appropriate should AID determine that projects in this category need to be evaluated in mandate impact terms. AID does, however, have a second option, which is to evaluate these projects in terms of the impact objectives the projects have defined, i.e., project Purpose and Goal. As indicated in Section Three, the Purpose and Goal objectives identified by any projects are legitimate statements of intended *impact*. The Purpose and Goal statements in these projects define the important changes that the project is expected to render in a situation. The fact that these statements are not posed in mandate terms (i.e., do not reach to this level) should not mislead AID into ignoring the impact the projects do claim will result from project action.

---

* AID's report from ATAC on the "Goal Hierarchy" may be a useful tool for thinking about these higher level linkages, even though it may not define the specific linkages needed to connect individual projects to mandate objectives. AID's workbook from PASITAM by Peter Delp is another source of ideas about how to define these linkages.

BEST AVAILABLE COPY

NARRATIVE SUMMARY OF THE EXTENDED LOGIC OF THE PROJECT

Health status of the poor families in the valley improved.

Caloric intake, particularly protein intake, of the poor families in the valley improved.

Family income raised for most poor families in the valley (all farm families and the families that do part-time paid farm work).

Farm yields per hectare raised.

Farm families apply new seeds and production techniques.

1. New seeds available at affordable prices
2. Courses and visits on new production techniques successful

PROJECT LOGICAL FRAMEWORK			
Narr Sum	OVIs	MOVs	Asmpts
GOAL			
PURPOSE			
OUTPUTS			
INPUTS			

Figure III.A.1: In this figure, the project's intermediate impacts correspond with levels in the Logical Framework, mandate impacts reach higher.

Practical Concepts Incorporated

BEST AVAILABLE COPY

In its initial round of impact evaluations, AID should probably not make an "either-or" choice with regard to what types of impact it will measure for these projects. Rather, the Agency should attempt to learn just how far it can go in attempting to trace impact at each level it considers important in a project, i.e., Purpose, Goal and each successive higher level it defines between Goal and mandate impact. PCI anticipates that AID will find it possible to measure impact at the Purpose, Goal and at least one level above Goal for most projects in this category. In some projects which define several closely linked Super-Goal levels (the case shown in Figure II.A.1), AID should be able to measure impact at all of the levels it defines. However, in other cases where the connection between project Goal and mandate impact is truly remote, AID's ability to measure impact for numerous Super-Goal levels may be limited. Only by attempting to define and measure all of the important levels (Purpose, Goal and such Super-Goal levels as are required to link the project to mandate objectives) will AID learn just what the limitations are.*

(3) Projects That Do Not Claim Mandate Impact, Directly or Distantly

When AID attempts to assess the impact of projects that claim a direct or distant connection to the Congressional mandate, it is probable that impact evaluations, in mandate terms, will be considered legitimate by most parties involved. On the other hand, when AID attempts to evaluate projects that do not claim mandate impact in mandate impact terms, it is possible that there will be claims (from those close to the project) that AID's evaluation effort is unfair or inappropriate. Several examples of projects where such claims might be made are discussed below.

---

* Where AID's early rounds of impact evaluations define a limitation, the Agency will be able to take appropriate action, e.g., carry out such work on methodology as is needed to break through the limits it has identified, or cease attempting to measure connections that cannot, given available tools, be measured.

BEST AVAILABLE COPY

AID has a number of active projects that were designed before the "New Directions" mandate was formulated. In some cases these projects took as their target group the urban poor or the middle class. Some based their expectations concerning impact on the premise that effects "trickle down". In projects of this vintage, the Purpose and Goal statements, and even the project approach, may bear little immediate resemblance to the current mandate. AID has several options should it decide to evaluate the impact of these projects. First, it could define project impact as it was stated in the original project documents. Secondly, the Agency could choose to review this portion of the portfolio, analyze the potential of each project for yielding mandate impact, and for those that have a high probability of producing mandate impacts, evaluate in terms of AID's current objectives. The third option is, of course, to combine the first two and examine these projects in the light of their own and AID's new objectives.

The prudent choice for AID is to avoid an "either-or" situation in its first evaluations of projects in this category. If AID decides to include projects of this type in its initial round of impact evaluations, PCI would suggest that they be evaluated both on their own terms and in mandate terms (after an "up-to-date" articulation of what might have been expected of the project in mandate terms has been prepared).* Should AID decide to undertake one or more "split-objectives" evaluations of this type, it should inform all concerned parties that the process it will carry out is experimental in order to relieve whatever anxieties are aroused by the idea of evaluating a project against objectives to which the project did not formally aspire. Only after completing several experimental "split-objectives" evaluations should AID make its final choice as to how to proceed with impact evaluations for this category of projects.

---

* The Logical Framework Approach should be applied, retroactively, against whatever set of documents exist for the project (including end of tour reports, prior evaluations, etc., as well as the design documents) to "tease-out" the possibilities for mandate impact. A logic and set of indicators for this alternative set of objectives should be developed before the evaluation is undertaken.

BEST AVAILABLE COPY

Supporting assistance projects, as discussed earlier, are a second example of projects that some may consider inappropriate for evaluation in mandate terms. Yet AID should not automatically exclude all projects of this sort from its impact evaluation efforts since development impact is likely to accrue in situations where the direct result of project efforts was an increase in stability and/or perceived self-reliance and security. The problem AID may face with these projects is that, while their impact may relate to mandated objectives, their benefits tend to be systemic rather than direct benefits to AID's target groups.

In conclusion, PCI's review of the AID portfolio from the mandate impact perspective suggests that all projects be considered potential candidates, and that AID will probably find that most projects can be examined in terms of their direct or distant relationship to mandate objectives. The subset of projects which neither claim nor turn out to have bearing on mandate impact may turn out to be trivial.

2. Demonstrating the Connection Between Project Action & Impact

When AID formulates a project, it states what it believes to be the connection between project action and impact in a set of hypotheses. As stated in AID's Handbook 3, each project has identified two main hypotheses: "'If Outputs, then Purpose' is the project hypothesis. The hypothesis that project Purpose will contribute to program or sector Goal is the program development hypothesis". The transformation of Inputs to Outputs, which is the essence of the project action, is not seen as being hypothetical. Rather, this transformation is considered to be within the manageable interest of AID's implementing agents.*

The form of the two development hypotheses AID puts forward to identify the connection between project action and impact is a predictive form "if A then

* In this sense, project Outputs are the deliverables under a contract or grant to implement an AID project.

BEST AVAILABLE COPY

B". This type of a hypothesis states that "A" (which may consist of a number of elements) constitutes the necessary and sufficient set of conditions to bring about "B". The hypothesis does not break down the sub-elements of "A" and identify their relative importance (or weights), nor is that required. To validate a predictive hypothesis of this type, it is necessary to provide "A" and then determine whether "B" in fact occurs. The measurements required are those needed to verify that "A" was provided (in the amount and form stated by the hypothesis) and the measures required to verify that "B" occurred and manifests itself (to the degree and in the form stated by the hypothesis).

These two sets of measures (proof that "A" was provided and proof that "B" occurred) *and nothing more* are required to validate a predictive hypothesis of the form AID uses in designing its projects. If such a hypothesis can be validated with project data, AID has the information it needs to indicate to Congress (and itself) that there is a valid basis for asserting the "A" will yield "B" in other, similar situations.

When we validate a predictive hypothesis, we have not necessarily explained what caused "B". Explanation is something different. An explanatory hypothesis takes the form "B because of A". When we examine an explanatory hypothesis, we must deal with the issue of attribution (true cause). A predictive hypothesis can be valid even when the true cause of "B" is hidden (when it is some factor we have not identified that correlates well with what we describe and provide as "A"). The distinction is not a trivial one, since it has important implications for the level of effort required in AID's project impact evaluations. Further, it is a distinction that many evaluators avoid making, probably because the question of true cause (explanation) is more interesting than is the question of whether a hypothesis is an accurate predictor of what is likely to happen when "A" is provided.

While explanation is in the long run more useful than simple validation of predictive hypotheses, it is not clear that AID needs to fully explain how impacts come about in order to fulfill its agreement with Congress. The

BEST AVAILABLE COPY

essential need, in terms of current and future programs (as stated in the June 25 memorandum to the Assistant Administrators) is to "increase the probability that:

- ...requested appropriation(s) will result in greater and more direct benefits to identified target groups in the priority areas;
- ...(AID) will have valid information on these planned benefits to present to Congress".

The information required to carry out both of these tasks can be secured by efforts to validate the predictive hypotheses stated in past projects. Full-scale studies that seek to explain the specific causes of whatever impact AID achieved in past projects would yield more than the minimum information needed to complete these two tasks.

This same memorandum, in discussing past projects, makes two statements concerning the information Bureaus will need to prepare for the FY81 Congressional presentation evidence that:

- "...such changes (in priority areas identified in Section 101 of the FAA) are associated with, or caused by AID development assistance";
- "...this impact (on target groups) was caused or stimulated by AID development assistance".

These statements appear to offer the Assistant Administrators a choice by suggesting that either (1) proof of association or stimulation (validated predictive hypotheses) or (2) proof of cause (explanation) constitute the evidential basis for AID's presentation to Congress. PCI further notes that nothing in the FAA or in AID's May 10, 1979 report to the Policy discussion group suggests that the choice must be made in favor of explanation.

AID's documentation on evaluation, including Handbook 3, its evaluation handbook, its guidelines and the as-yet-unreleased chapter on evaluation for Handbook 3 all discuss "unplanned causes" that may explain project effects.

BEST AVAILABLE COPY

Yet none of these documents require that, as a normal part of evaluation, alternative explanations of observed effects (impact) be eliminated. The process of eliminating alternative explanations of an observed effect is a difficult one, and it may simply be a slip between intent and wording that has left this out as a requirement for AID evaluations. Nevertheless, this difference in wording does exist, and AID may wish to take advantage of that fact in its early rounds of impact evaluation.

The opportunity to use some discretion in interpreting AID's own requirements in this area, as well as the general language used to date in describing what AID will provide to the Congress, suggest that early project evaluations will meet the need if they succeed in validating the predictive hypotheses formulated by projects, i.e., by separately proving that "A" was provided and that "B" occurred.

The fact that AID can meet its requirements to demonstrate the connection between project action and impact by examining the less rigorous of two forms of a hypothesis, in the short run, should allow faster movement in the direction Congress wants AID to move than would be the case if tests of the explanatory form of AID's hypotheses were explicitly required. This does not, however, mean that AID should ignore the issue of explanation. Rather, it should start now to experiment with this tougher form of evaluation so that it is prepared, at a later date, to provide proof of *cause* as well as proof of the association between project action and impact.

In the methodologies discussion in Chapter 4, PCI reviews the type of procedures required to gather information on cause as well as to validate predictive hypotheses.

BEST AVAILABLE COPY

3. The Sustainability of Project Impact

The issues of the sustainability of project impact (or "benefit continuation")* and decisions concerning the timing of impact evaluations are related.

Sustainability, or benefit continuation, refers to the level of impact maintained, or the stream of benefits generated by a project over time. The concept explicitly recognizes the possibility that initial levels of impact may differ from the levels of impact that are observable after several years. AID itself is aware of the instances in which benefit streams failed to develop or preliminary impacts eroded. The sources of such erosion have been suggested to include lack of funds for continuing the efforts begun under a project, lack of required talent, lack of interest, etc.

The question of whether a project yields its intended impact and the question of whether that impact is sustained are different. The first asks whether a project produced the level and type of benefits it promised. The second asks whether, after the project was completed, the benefits continued to flow. These two questions are not easily answered simultaneously. AID has several options for securing the answers to each of these questions.

a. Measure Impact/Predict Benefit Continuation

This option (which DAI used in its study of PVOs for AID), involves a one-time evaluation. The evaluation is conducted at a point in time when it is expected that the impact promised by the project will be observable. The evaluation tasks are two:

- Measure whether the project yielded the impact it claimed would result from project action;
- Gather such data and perform such analyses as are useful in making a prediction concerning whether the observed level and type of benefits will be sustained over time.

---

* DAI's final report on its evaluation of PVO projects for AID uses the term "benefit continuation".

BEST AVAILABLE COPY

b. Measure Benefit Continuation/Infer Initial Impacts

This option (which AID used in its pilot ex-post evaluations) involves a one-time evaluation. The evaluation is conducted well after the point at which project impact could first be observed, e.g., three-five years. The evaluation tasks are two:

- Measure project impact as of the time of the evaluation to secure accurate data on sustained benefits;
- Estimate the level and type of impact that was yielded at the point in time when the project claimed its impact would be felt, i.e., a point several years earlier.

c. Estimate both Planned Impact and Sustained Impact

This option involves one evaluation at a point in time midway between what might be considered the "long-term" and the point in time when the project indicated its impact would be observable, e.g., 18 months. The evaluation tasks involve:

- Estimating planned impact (per (b) above);
- Measuring benefit continuation to date (per (b) above) and, if desired, predicting further continuation of the project impact (per (a) above).

d. Measure both Planned Impact and Sustained Impact

This option involves assessments at two points in time: The time period in which the project predicted its impact would be observable, and some later point that describes "long-term" for the project. The tasks during each assessment period are listed separately:

Point 1:

- Measure whether the project yielded the impact it predicted;
- Predict benefit continuation.

Point 2:

- Validate the benefit continuation prediction, i.e., measure impact at Point 2.

As the option descriptions suggest, only Option d gives AID solid information on both planned benefits and benefit continuation. The third option, on the

BEST AVAILABLE COPY

other hand, gives neither an initial nor a long-term portrait of impact. Options a and b essentially force AID to choose between having information on planned impact and having information on benefit continuation, since both use weak approaches to make statements about the level and type of benefits at the second (unobserved) point in time.

As the preceding discussion suggests, the issue of evaluation timing is in good part an issue of what AID wants to measure with a high degree of accuracy--the benefits promised when they were promised, or long-term effects from its efforts. In PCI's judgment, the best choice would be to do both (Option d), since both initial impact and long-term effects are important and we learn different things by evaluating each.

4. Identifying Lessons that can Guide Future Agency Efforts

The idea of creating space in an evaluation scope for recording "lessons learned" is not a new one in AID. The item has from time to time been listed as part of the scope of mid-project evaluations and special evaluations conducted by the Agency.

Understandably, AID wants to learn all it can from each of the impact evaluations it carries out. It is not clear, however, that including a scope item for "lessons learned" is the way to ensure that all pertinent information is recorded. What may be needed is a structured item that uses the title "lessons learned" but then goes on to describe the set of subcategories AID wants to have addressed. From many discussions with AID personnel, it would seem that the list of subcategories of interest (the thoughts that trigger a sense that "lessons learned" is a needed scope item) include:

a. Formal statements of study conclusions

Evaluation studies collect and analyze data. Often they present a large number of specific "findings" or facts and then jump immediately to recommendations. The recipient of a study report that

BEST AVAILABLE COPY

makes this type of leap is often left at sea. How do all these little facts add up to a recommendation to expand/terminate the project? One element or sub-category of "lessons learned" might well be a requirement to list the conclusions the evaluator(s) drew from the study findings.

b. Partial Information/Hunches

During the course of most evaluations, evaluators pick up information that while often partial, and usually not very rigorous, suggests conclusions about a project, or the type of facts that might have surfaced if additional study questions had been defined early on. In a formal study (e.g., a sample survey) there often seems to be nowhere to record such "soft" data. Having a specific scope item for recording this type of data would have two effects. It would in many cases lead to the recording of partial information/hunches that are often excluded by evaluators because their evidence is not adequate to justify recording the data as a finding. Secondly, it would encourage evaluators who commingle their partial data with the data that is based on solid research to separate the two.

c. Management Information

Often evaluators learn a lot about how a project was managed when they attempt to examine its progress. While AID's impact evaluations will reveal less of this type of information than a good mid-project evaluation, it may be a good idea to ask that observations about project management which have some bearing either on the continued operation of the project or on the management of similar/future projects, be recorded.

d. Information Pertaining to External Validity

The issue of external validity is an important one for AID. A set of project hypotheses have internal validity if they "work" in the project situation in which we said they would work. These hypotheses are said to have external validity if they can be applied to other situations. In planning to replicate a project design elsewhere, AID must be concerned both with whether the design worked, and whether the hypotheses have external validity. When AID evaluates a single project it is not usually able to make a definitive statement on this topic. Evaluations that assess several projects of the same type provide better information on this issue. Nevertheless, evaluations of individual projects often surface information that pertains to the question of external validity. Included in a list of such information would be data on what project assumptions appeared to be critical, what assumptions that the project failed to make proved to be important, what political/cultural factors

BEST AVAILABLE COPY

appear to have had bearing on the success/failure of the projects, etc. To the degree that AID has its evaluation teams record this type of information it increases the probability that replication attempts will be made only where there is no known reason not to make the attempt.

e. Methodological Observations

This includes what worked, what didn't work, what proved difficult about the design for the evaluation, the instruments, the data collection approach, the analysis, etc.

f. Inferences/Deductions that Go Beyond the Project Framework

Sometimes, though perhaps not often, evaluators will in the course of examining a project find that the exposure to a great deal of factual information, conversations in the area, etc., trigger a new idea about how development works, or is affected by factors that go well beyond the project they are evaluating. Such thoughts and hypotheses may be tangentially related to the specific project, e.g., cross-sector interactions that have not been identified previously. At other times the relationship to the specific project will be quite remote, concerning instead the effects of some AID or host policy, cultural patterns, or what Elliott Morss termed the "rotors" of development in his paper for PPC on "Cross-Cutting Issues".

Since there is no formal place to record such ideas within the Framework of a project evaluation, evaluators often tell themselves they'll write a monograph on the idea someday, and then they forget their inference/deduction or brainstorm because there are other pressing things to do. A final "lessons learned" item that encouraged the recording of these half-formed hypotheses could prove useful to AID.

The level of effort required to complete a "lessons learned" section for an evaluation report is not usually excessive, hence it is relatively easy to include (and not likely to be detrimental) should AID want such a scope element. The idea of creating and describing subcategories is an option that AID might consider if it seeks to gain consistent value from such a scope item.

5. The Appropriate Focus of Recommendations from Impact Evaluations

As suggested in Section Three, the primary "user" of AID's impact evaluations is Congress. On the other hand, it is not clear that Congress is the appropriate

BEST AVAILABLE COPY

audience for action recommendations that can be identified when an impact evaluation has been completed. The question of what audience is best equipped to act on such recommendations may in fact be problematic.

In projects which have been "turned over" to the host country for continued operation, the current managers of the operation may be appropriate users of certain types of evaluation information and recommendations. The AID project team will generally not be an audience, since in many instances that team will have disbanded. In situations where the AID staff or host ministry staff are considering follow-up projects, they may find study conclusions and recommendations useful. but this will not always be the case. The staff of AID/W may be interested in recommendations based on these evaluations, but it may not always be in a position to take actions that are recommended.

As the preceding paragraphs suggest, AID may need to treat the question of who should receive the recommendations that grow out of project impact evaluations on a case-by-case basis. In specific scopes of work developed for individual evaluations it should indicate who (what individuals/groups) will consider recommendations and further, what are the range of actions AID can plausibly take, e.g., AID cannot/will not become reinvolved with this project (either by providing resources or additional technical assistance).

The idea of placing "boundaries" on evaluation recommendations by identifying both uses and constraints is not new; it is both practical and prudent. PCI, in other reports, has recommended that AID identify the limits of what it will be able to do based on evaluation findings as it begins any evaluation. Taking this course for project impact evaluations is particularly appropriate. That is, AID can and should avoid a situation in which project impact evaluations make multiple recommendations on which AID cannot/will not act (because they are outside the range of what is appropriate/feasible). It makes little sense to have evaluations generate this sort of recommendation, and then find that the recommendation must (a) be monitored to determine what action AID and/or the host government took, and (b) be presented to Congress with the explanation that it is not feasible, and AID plans no action.

BEST AVAILABLE COPY

"Boundaries" or definitions of the users and range of plausible action are not forms of censorship, nor should they be used to create such an effect. Used fairly, they help make evaluations function as helpful management tools. The possibility for abuse exists, however, and AID will need to keep an eye out for instances in which it starts to overstate the definition of "actions that cannot/will not be taken", i.e., include in the list all possible actions AID will not want to take, rather than the shorter list of actions it simply cannot take.

6. Monitoring Actions Taken by AID and Host Governments

As AID itself has pointed out, the Agency will need to monitor what actions are taken based on project impact evaluations. However, this task is properly a management task for the Agency, rather than an "evaluation" task *per se*. That is, the requirement for AID's monitoring of actions based on evaluation recommendations is not a legitimate element of the "scope of work" for such evaluations. The evaluations must be prepared, received, and reviewed before action commences. It would be inappropriate to charge the evaluation team with the responsibility for follow-up of this sort. Hence, PCI concludes that this item should be deleted from the list of possible scope elements for project impact evaluations.

Monitoring and assessing the actions taken as a result of project impact evaluations is a proper function of the Agency, and the responsibility for ensuring that these tasks are completed remains. Hence, this set of tasks is discussed again in Chapter Six, which addresses the management issues associated with the implementation of a program of impact evaluation.

7. The Methodological Record of an Impact Evaluation

The quality of an evaluation is largely defined by the methodology used to carry out the evaluation.

BEST AVAILABLE COPY

The quality of an evaluation is a function of how well it was conducted, and depends on several factors:

- What questions were posed, and whether they were the right questions, i.e., the quality of the "scope of work";
- The quality of the procedures used to address the study questions (whether they were valid and appropriate);
- The quality of the information secured as a result of applying study procedures (its accuracy, completeness, etc.);
- The quality of the analyses performed by the evaluators (raw data manipulation, the use of study data as opposed to opinion/assertion in moving from findings to conclusions and recommendations, the presentation of study information in useful forms, etc.).

In a broad sense, each of these quality factors are methodological. When evaluation studies fail to report on these factors, the potential users of evaluation results must make an unnecessary "act of faith". The evaluation studies AID has mounted in the past have often forced evaluation users into this position; they have failed to require/secure complete descriptions of the methodologies employed in evaluations AID has funded.*

The Agency should avoid repeating this oversight as it defines the scope of its project impact evaluations. The general scope put forth for all project impact evaluations should include a requirement to fully and accurately disclose four elements of the study's methodology: The study questions, the procedures used to secure data, an assessment of the data quality, and a description of the analyses procedures (including not only data manipulation procedures but also the way in which data was handled in moving from findings to conclusions and recommendations).

---

* In 1976, PCI and Checchi worked with PPC/DPRE on the problem of evaluation abstracts for the DIS. PPC/DPRE sought to include abstracted statements about study methodology in these summaries. PCI and Checchi found that this was impossible due to the uneven and generally inadequate reporting of methodology in the evaluation studies AID had on hand (both in-house and contractor efforts).

BEST AVAILABLE COPY

4

BEST AVAILABLE COPY

The Timing, Duration, & Intensity of Impact Evaluations

CHAPTER FOUR

A. EVALUATION TIMING

Part of the question of the timing of impact evaluations is tied up with AID's definition of success. (Clearly, the Agency cannot consider its development assistance successful if its projects continue only for a short period after the Outputs have been produced but before funds have stopped flowing to the project.) If projects could not continue to provide Purpose and higher level benefits after they became operational, there would be many legitimate questions to be answered concerning the value of development assistance.

Secondly, there is a general question of what is meant by impact. If we consider it to be the case that, for example, the impact of a child feeding program will not be fully felt until the children involved reach maturity, we may find ourselves unwilling to declare a project successful/unsuccessful for a full generation. While it may be true that the stream of benefits from a given nutrition, education, health or population project may extend many years into the future, it is not clear that:

- AID can afford to define success/failure in a way that cannot be assessed within, say, the period of a single political administration in this country;
- AID personnel judge project success/failure in these extended terms;
- Impact evaluation (that included any attempt to associate or attribute cause) could be carried out over the time period implied by this view of impact.

BEST AVAILABLE COPY

Despite the impracticality of this view of project impact, from an evaluation and legislative standpoint, AID needs to recognize that it must deal with this view *no matter what time frame it picks for its project impact evaluations.* To those who firmly believe that impact cannot be measured until the entire stream of benefits has manifested itself, five years after termination will be just as unsatisfactory an answer as a date shortly after Outputs have been provided. This is not to say that the argument that benefits mature is invalid; in all probability it is a valid and reasonable view. It simply isn't possible for AID to address the question of impact in these terms.

Realistically, AID must consider a range of time frames that fall between the time shortly after "treatment" is provided and some point at which (a) the sustainability of benefits could be assessed and (b) attribution of impact to the project action is feasible. The issue, thus rephrased, becomes one of determining:

- How soon are project benefits measurable?
- How early can sustainability be assessed?
- What is the latest point in time at which it is feasible to make a connection between project action and impact?

PCI's conclusion (presented in Annex A to this Chapter), based on an examination of these three questions, is that AID should be able to gather data on both impact intensity and some evidence of benefit continuation if it performs project impact evaluations approximately 15-20 months after the project Outputs have been produced. If more data is needed on sustainability, AID should consider assessing impact twice. The first should occur at a point soon after Outputs are produced and benefits are expected. The second time, a low cost follow-up, could come several years after the project, and would simply assess whether the benefits found in the main impact evaluation remained. This two-part approach is discussed below.

BEST AVAILABLE COPY

Annex II.A. outlined four Agency options for addressing the issue of the sustainability of project impact, an issue that is closely related to the question of optimum evaluation timing. The options were:

- Measure Impact/Predict Sustained Benefits;
- Measure Benefit Continuation/Infer Initial Impacts;
- Estimate both Initial and Sustained Impacts;
- Measure both Initial and Sustained Impacts.

PCI's review of the paired issues of timing and the measurement of sustainability has suggested to us that AID's fourth option (of measuring both initial impact and sustainability, but not at the same time) is the optimal solution. If this option is elected, AID should measure impact and attempt to establish the relationship between cause and effect at some point within the first 18 months after "treatment", with the specific timing being dependent on when project impacts can be observed. Within this period, earlier dates where feasible may be better if AID is trying to test explanatory hypotheses. In a second, less rigorous impact evaluation (follow-up), the total picture would be rounded out by securing (1) data on the sustained versus initial level of impact and (2) data that constitutes evidence of the continued operation (provision of benefits from) the project.

Backing away from this optimal position, it is PCI's judgment that the next best option is to time evaluations so as to measure initial impacts and assess causality, without waiting for adequate evidence on sustainability. If this option is elected, PCI would recommend taking some loss on "proximity" with regard to the examination of cause and effect to gain a purchase on the issue of sustainability (using largely predictive data if necessary). This Option, #3, offers AID a good deal of information with a one-time study. The information may not be perfect, but some data on both issues will be available. As Figure IV-1 suggests a point in time might be selected, on a case by case basis, for evaluations that seek information on both immediate and sustained impact.

BEST AVAILABLE COPY

BEST AVAILABLE COPY

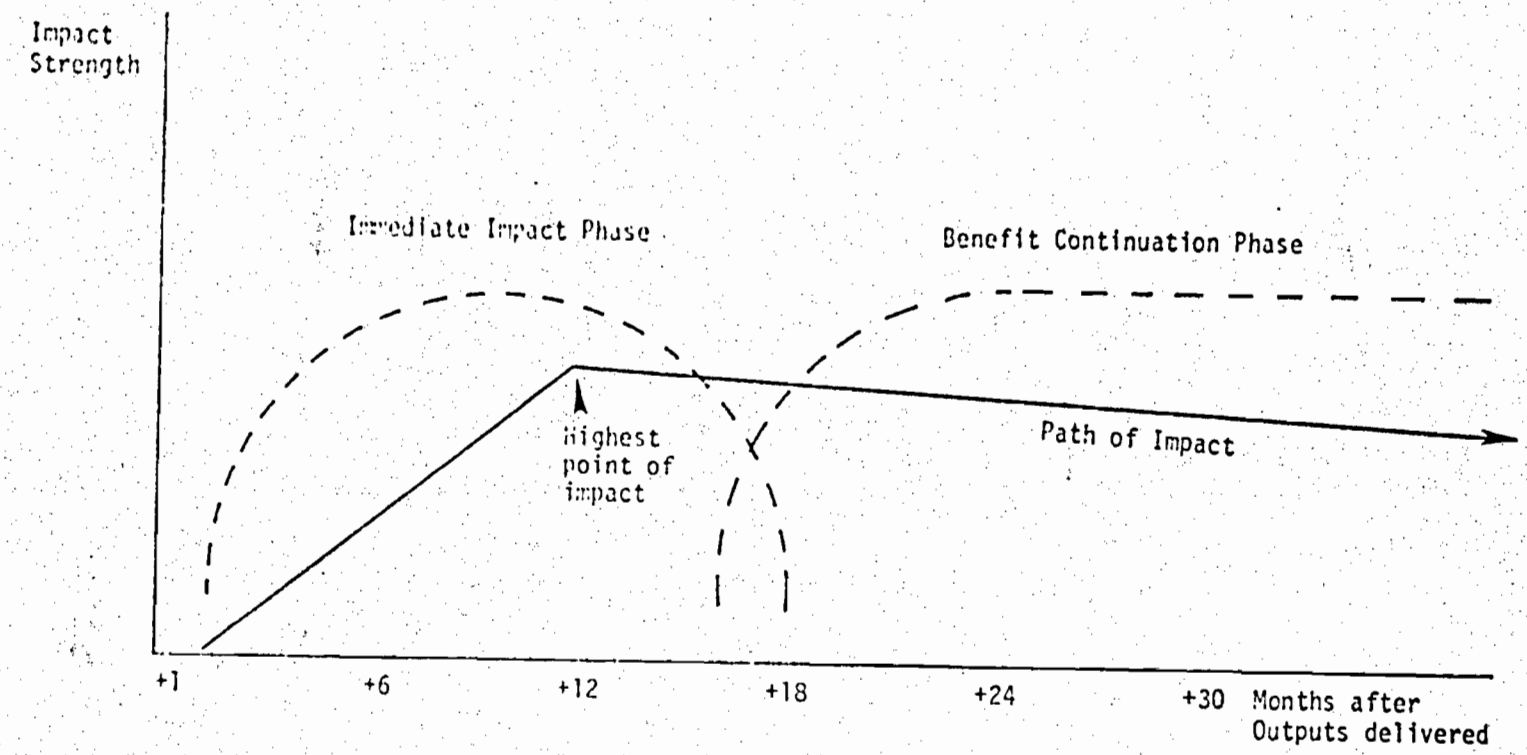


Figure IV-1: Hypothesized Overlap Between Immediate & Sustained Impact of a Project

BEST AVAILABLE COPY

The remaining two options each "cost" a lot in terms of information not secured, and are therefore both less desirable. Of the two, Option 1, which focuses on initial impacts, would be PCI's recommendation, since the difficulties of getting data on what is possible in terms of impact at a later point (three to five years out) are significant. And, as pointed out elsewhere, the impact and sustainability issues are slightly different. AID's current task is one of assessing impact.

B. THE DURATION & INTENSITY OF PROJECT IMPACT EVALUATIONS

In theory, decisions about the duration and intensity of an evaluation are derived from decisions made about what is to be measured and the standards of evidence/methodologies to be employed. However, a review of 25 AID evaluations and 25 IBRD evaluations undertaken by PCI and the World Bank independently (but using the same approach) has suggested that in practical evaluation duration may be determined on other grounds. The 50 studies examined differed dramatically in scope, methods and cost, but the vast majority took a one-year period to complete.* This data suggested to PCI that, in past evaluations, duration may well have been a function of the time available. Further, there is anecdotal data to suggest that when information is needed quickly, the duration of an evaluation is prespecified, and measurement approaches are selected based on what is feasible within the allowed time.

Evaluation intensity reflects both the scope of an effect, e.g., how many questions the evaluation attempts to answer, and the level of effort put into answering the evaluation questions. A few questions examined extensively (e.g., a sample survey) can require the same number of manmonths as an

---

* PCI carried out this review as part of the background work in connection with its effort to prepare the "Manager's Guide to Data Collection", under Contract No. AID/otr-C-1377, W.O. 14. The IBRD review was an internal effort which the Bank shared with us and with PPC.

BEST AVAILABLE COPY

intensive case study which tries to answer many questions in depth.* One of the few indicators of the intensity of past evaluations is their cost. The PCI/IBRD review of part evaluations found that costs varied widely. Further, it was found that the number of observations (e.g., interviews) did not correlate well with total cost. One or two explanations appeared possible. Either there are real cost differences by country (which change the cost per observation from $10 to $100+) or intensity is simply another way of saying that whatever money is available get used.

The actual period of time required for an impact evaluation will also depend upon whether the evaluator used existing records, such as statistics kept by the host government, or collects new field data. Data often exists on the issues of interest to AID. Too often no one looks to see what data is already available. When the effort is made to find existing records, evaluation time can be cut down significantly. With existing data an evaluation might well be completed in the period of a few months, e.g., two-four depending on the situation and the type of analysis performed. When new field data must be collected, evaluation studies take linger. Still, experienced evaluators with a good deal of practical applied research experience can gather and analyze survey data on a project group and a comparison group in a reasonable period, e.g., six-eight months under good working conditions.

On the other hand, when inexperienced personnel are used extensively, study duration can increase and the degree of intensity (e.g., number of observations, length of interviews, etc.) can exceed what is actually needed to secure the information sought by the study. There is no hard data available concerning the difference in duration and intensity found when experienced versus inexperienced personnel are used. PCI's field data collection work has suggested that it is possible to use relatively inexperienced interviewers, with no loss of time, as long as these field workers are trained and supervised by someone with a good deal of practical experience. Conversely, there are

---

* Annex III.B. to this chapter discusses the effects of different terms of reference and methodology on the duration and intensity of evaluation.

BEST AVAILABLE COPY

aspects of evaluation studies, such as the development of instruments, sampling, etc., which are more efficiently completed when placed in the hands of individuals who have a good deal of experience. Novices in these areas tend to lose time through a trial and error process. No matter what training an individual has had, some aspects of field research tend to be learned "on-the-job" in the first studies they carry out.

In addition to considering the time and level of effort required to carry out a project impact evaluation once a study team begins work, AID needs to consider the front- and back-end work that must be carried out by the Agency in order to set up an impact evaluation and put its results to use in AID programs, projects and the preparation of Congressional presentations on program impact.

BEST AVAILABLE COPY

FACTORS THAT AFFECT THE TIMING OF PROJECT IMPACT EVALUATION

Three questions that must be addressed to determine the appropriate timing of project impact evaluations are examined in this Annex.

1. When Project Benefits are Measurable

PCI has reviewed the question of when the benefits from project action can be measured from a number of perspectives. All of these lead to the same conclusions: Project benefits, for the vast majority of AID's projects, will be observable within a short period after project "treatment" has been provided. The length of the period between the provision of the Outputs and observable impact is expected to vary by the type of project and the type of impact(s) involved. But as a general rule, impact should be measurable within a few months for some projects, and for others sometime within the first 18 months after "treatment" is provided. (This generalization covers the bulk of the AID portfolio. Some types of projects are excluded from this generalization. They are discussed in greater detail below.)

The data that leads to this generalization comes from several sources. While we did not weigh all of these sources equally, it is worth noting them individually:

- Project Statements about Impact

  PCI selected a set of 16 projects, of the 1,940 it had reviewed in terms of Purpose and Goal level objectives. These 16 fairly typical projects were reviewed to determine when the project designers expected their efforts to yield observable impact. (The sample we reviewed contained nine income projects, three production projects, and four health/nutrition projects. We found that all but one project claimed that impact would be observable by the time of project termination. This single project claimed that within roughly 18 months after termination, impact would be observable. In addition,

two of the projects that indicated that impact could first be assessed at termination also suggested that some portion of the initial impacts would only be measurable after a year or two.*

- "Expert Judgment"

For the area of nutritional and health impacts, PCI examined the question of impact measurement with staff from PHA and Johns Hopkins School of Public Health. The conclusion from these interviews was that a large number of the impacts from health service and nutritional programs would be measurable almost immediately. The exception to this rule was described in terms of health services that address what are "rare events". That is, for certain impacts that are expected to accrue to only a small portion of a population, it would be difficult to know, in a statistically significant way, whether the service was effective until a reasonable sample of cases had been observed. For some effects it might take time before a reasonable sample, from a measurement perspective, built up. One example in this area would be the handling of unusual pregnancy complications.

- Evaluation Experience

In addition to the above sources, PCI reviewed its own experience in impact evaluation (particularly the studies conducted in Latin America by Dr. Samuel Daines) to determine when the health, nutrition, production, income, infrastructure and institutional development effects of these projects had proven to be measurable. The conclusions reached as a result of this review were that (1) basic health, nutrition, income and production effects are measurable within a period of four months to 18 months, depending on the impact type, (2) some effects of infrastructure projects are measurable even before "treatment" (e.g., the road), is provided. These effects appear in the form of population shifts, relocation of industry, growth of small shops and marketing institutions, etc. Additional impact in mandate terms is measurable within the general 18 month period specified above, and (3) institutional development projects, many of which strengthen rather than build new institutions are a somewhat special case. First, in these projects we must measure mandate impact *through* the institution AID assisted, e.g., in the institution's clients. Second, the degree of change in the institution AID assisted may not be great *and* the number of clients who are reached, and who themselves change during the first 18 months after the institution is acting in a new or modified manner, may be small. Nevertheless, it is PCI's judgment that where

---

* Exhibit II.A.1 to this Chapter provides the datasheets for this analysis.

BEST AVAILABLE COPY

the direction of change is in both instances (institutional and client change) positive, statements about impact (though not measures of long-term cost-benefit) can be made within 18 months.

2. When Benefit Continuation is Measurable

While the concern about the sustainability of project impact after AID's project effort is not new, attempts to devise predictors and measures of sustainability is relatively recent.*

Because there has been so little work in this area, the answer to the question of when sustainability can be measured (as opposed to predicted) must be tentative. There are however some elements that must be included in any answer:

- First, AID's assistance must be fully terminated before sustainability can be assessed;
- Second, there must be an observed stream of benefits to be sustained;
- There must be some data available that constitutes evidence that an effective operational mode for continuation exists, e.g., a new source of income, new activities undertaken by the administration or otherwise within the framework of the project in its long-term mode.

These elements to some degree suggest the time frame for first measurements of sustainability. As the preceding subsection indicated, PCI has concluded that initial impacts will be observable for the vast majority of AID's projects within 18 months after the Outputs are provided. Using this figure as a "grace period" one might conclude that 18 months after Outputs are realized is also the point at which measures of sustainability could be taken, given that (a) all AID activity has ceased and (b) some other measure of continuation is available. In some projects these two conditions will not exist at

* PCI developed an approach for assessing individual institutional viability for AID several years ago; more recently, DAI has devised an approach for assessing benefit continuation in PVO (and possibly other) projects.

BEST AVAILABLE COPY

the end of the first 18 "post-treatment" months. Either AID will still be involved (in cases where "treatment" began before termination for at least some of the target area/group) or measures of operational status (such as new funding) will not be available.

If the earliest date may be too early for some projects, the question then becomes what is a reasonable outside date by which all of these criteria should have been met. Again, guestimates vary depending on the type of project. However, these answers tend to fall mostly in the three to five year range, suggesting that for a lot of projects, the end of the fifth year would be a reasonable "late date" to assess sustainability. As indicated, these dates are "guestimates", and within (and beyond) the 18 month to five year range to determine just exactly how early benefit continuation can be measured.

3. When Can Association/Causation be Assessed

The general answer to this question, based on past evaluation efforts, including AID's pilot ex-post evaluations, is that the chances of establishing causality are better closer to the point at which "treatment" is provided than they are later. This is not to say that in many cases attribution cannot be made at later points. It does, however, suggest that there are projects where the difficulty of demonstrating an association or causal relationship increases over time.

Nutritional programs for adults and infrastructure projects are two cases in point. In the first case, there may be difficulty in assessing the effects of diet changes even close to the time when the change was made, when an adult (as opposed to infant/youth) population is involved. With infrastructure projects, e.g., roads/irrigation, changes are perceptible relatively early, but as time goes on the number of factors that can account for subsequent changes increases, often rather quickly (when new markets, etc., become available in the area). This suggests that a measurement point earlier (e.g., up to 18 months after "treatment" is to be preferred to a later point (e.g., 5 years).

BEST AVAILABLE COPY

A ROUGH ANALYSIS OF IMPACT EVALUATION TIMING FROM THE PERSPECTIVE OF PROJECT PLANS & SCHEDULES

During the course of this engagement, PCI examined 16 projects that a review of the 1,940 projects in AID's DIS database suggested were typical. The attached forms were used to record project information. By aggregating this limited data we found that as a general rule, projects in the four areas reviewed claim that impact will be observable by the time of project termination. Discussions with specialists in a number of areas also suggested that impact, in a wide variety of projects, should be observable quite quickly after the project Outputs are provided.

Mandate Impact to be Measured	Type of Project	# of Projects	When Impact is Expected
Income	Crops/seeds	3	o/a TDD
Income	Credit	5	o/a TDD
Income	General Ag	4	TDD, possibly up to a year earlier
Health/Nutrition Status	-	4	o/a TDD

BEST AVAILABLE COPY

PROJECT NUMBER: 527 0149

COUNTRY: PERU

CATEGORY: INCOME

SUB-CATEGORY: CROPS

START: 77

TOD: 80

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Increase productivity, employment and income and improve nutrition among the poor.	✓	X	X	X	The idea is clear but there is NO criteria for success.
OVI 1: % increase in farm productivity	✓ rural production	X	X	X	↓ ↓ ↓
OVI 2: % population with access to more or better food supply	✓ nutrition	X	X	X	
OVI 3: —	—	—	—	—	—
PURPOSE: 1) Increase food consumption among the poor. 2) Establish soybean and corn production on 10,000 hectares by 80	✓	✓	✓	TOD to TDD	In the project paper specific targets are set for each year after the project begins.
OVI 1: Per capita corn consumption ↑ 15% 10,000 MT soy consumed	✓	✓	✓	↓	
OVI 2: Up to 10,000 farmers rotate soybeans producing 5,000 MT soybeans per year	✓	✓	✓		
OVI 3: 5,000 highland farmers increase corn yields (using improved seeds) by 15-30% on 1 hectare plots	✓	✓	✓	↓	↓ ↓ ↓

1977 — START-UP

1978 — TOF

1980 — TDD

targets for seed production and crop production are listed for these years

TMO

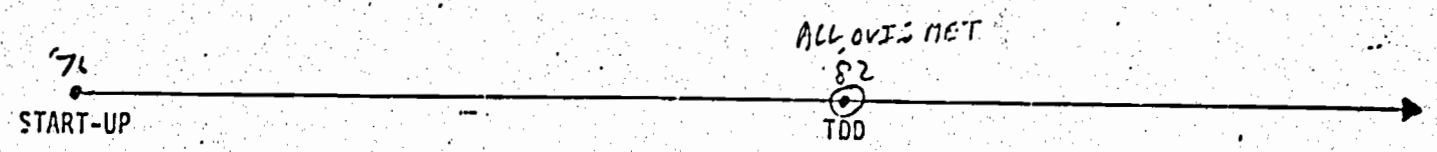
IV.A.1.-2

BEST AVAILABLE COPY

PROJECT NUMBER: 493021000 CATEGORY: Income START: FY76

COUNTRY: Thailand SUB-CATEGORY: Crops TOD: FY82

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Increase the farm income of the poor people in rural areas	✓				
OVI 1: Target area income up 15%	✓	✓	✓	TDD	(date given in paper)
OVI 2: Farmer productivity increases through yield increases associated ~~[illegible]~~ w/ use of project seed (% yield for 6 crops given)	✓	✓	✓	TDD	GOOD
PURPOSE: Cause use of improved seed on Thai farms with outreach to farms of low income.					
OVI 1: By 1982 have improved project seed yields meeting nat'l requirements.	X	✓	✓	TDD	GOOD
OVI 2: 75% of farms in target areas use seed of improved varieties	X	✓	✓	TDD	GOOD
OVI 3:					



BEST AVAILABLE COPY

PROJECT NUMBER: 730-11-130-314 CATEGORY: INCOME START: 71

COUNTRY: Vietnam SUB-CATEGORY: CROPS TDD: 74

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Assist Vietnam in the efficient production & distribution of agricultural products to earn foreign exchange, income & provide an adequate diet for people	✓				PROJECT DOCUMENTS
OVI 1: NO OVI					INDICATE A
OVI 2:					REALLY PLANNED
OVI 3:					PROJECT
PURPOSE: Develop Vietnamese Agriculture Dept. to plan & implement programs for obtaining development & production goals for crops	X				
OVI 1: Seed quality control	X	X	✓	X	
OVI 2: Specific increases in production by crop and by date (77 to 78)		✓	✓	TDD13 TDD14	
OVI 3:					

START-UP 71 TDD 74 TDD13,4 77 78

IV.A.1.4

BEST AVAILABLE COPY

PROJECT NUMBER: 6150173

COUNTRY: KENYA

CATEGORY: HEALTH

SUB-CATEGORY:

START: Jan. '76

TDD: Dec '79

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Improve quality of life & increase the productivity of rural Kenyans by reducing eye disease	✓				
OVI 1: Reduce incidence of eye disease by 2% by year	✓	X	X	✓	
OVI 2: Kenyans returned to productive & useful pursuits annually	✓	X	X	✓	
OVI 3: Kenyans treated for eye disease annually	✓	✓	✓	✓	
PURPOSE: Initiate program of blindness prevention and health education and improving rural delivery of eye care to rural poor over 3 yr period	✓	✓	✓	TDD	
OVI 1: Dev list 10, almost all of which are quite good.	✓	✓	✓	TDD	
OVI 2:	✓	✓	✓	TDD	
OVI 3:	✓	✓	✓	TDD	

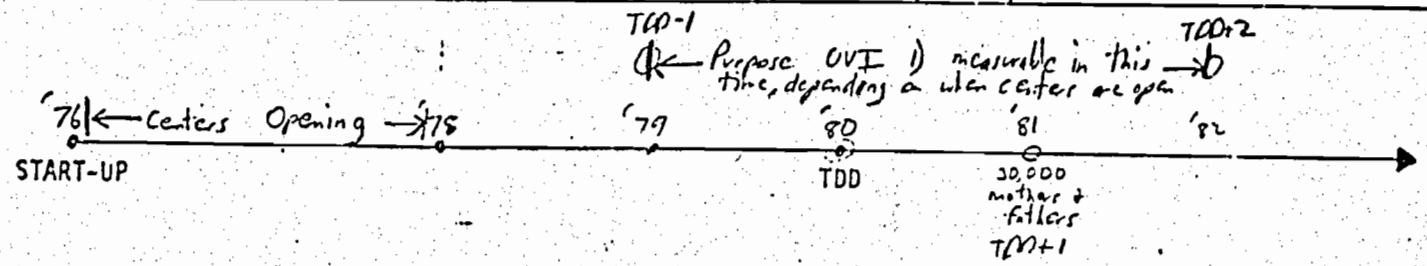
START-UP Jan. 76 → TDD Dec. 79

BEST AVAILABLE COPY

PROJECT NUMBER: 521-11-560-075 CATEGORY: HEALTH START: FY '76

COUNTRY: HAITI SUB-CATEGORY: TDD: FY 80

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Improve health of Haitian family in lowest income groups.	✓	—	—	—	
OVI 1: Reduce morbidity and mortality in under 5 age group from causes of malnutrition and infectious diseases.	✓	X	✓	X	
OVI 2: —	—	—	—		
OVI 3: —	—	—	—	—	
PURPOSE: Reduce malnutrition and infectious disease in under 5 age group.	✓	—	—	—	
OVI 1: Eliminate 3/5 of malnutrition and mortality in under 5 age group w/i 2-4 yrs. after center is opening	✓	✓	✓	TDD-1 to TDD+2	EXCELLENT - CRUX of PROJ.
OVI 2: Families will produce and consume more nutritious foodstuffs.	✓	X	X	X	TERRIBLE INDICATOR
OUTPUT OVI 1: 30,000 mothers and fathers will have participated in program by end of 5 years.	X	✓	X	TDD+1	



BEST AVAILABLE COPY

CATEGORY: IV [illegible]

START: FY 76

COUNTRY: LDCs

SUB-CATEGORY:

TDD: FY80

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Initiate appropriate systems for delivery of absorbable iron to LDC populations	✓	—	—	—	
OVI 1: Significant decrease in iron deficiency	✓	✓	✓	TDD	Though the Logframe is poorly filled out, in the paper specific goals w/ Q, Q, & T are included.
OVI 2: —					
OVI 3: —					
PURPOSE: Assist LDCs in implementing programs for alleviating iron deficiency anemia	✓	—	—	—	
OVI 1: Assess anemia in 10 selected countries	X	✓	✓	TDD-3 to TDD+1	What assessment consists of is clearly defined in the paper.
OVI 2: Develop appropriate techniques and procedures for delivering absorbable iron to vulnerable populations	X	✓	✓	TDD-4 to TDD	Demonstration projects
OVI 3: Implement programs in 3 countries	X	✓	✓	TDD-2 to TDD	These do NOT get at solving the problem, just studying it.

'76 START-UP — '80 TDD

a decrease in iron deficiency is predicted but it is [illegible]

BEST AVAILABLE COPY

PROJECT NUMBER: 498-15-130-221 CATEGORY: Nutrition START: 71

COUNTRY: International Rice Research Institute Indonesia SUB-CATEGORY: Institutional Development TDD: 75

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: 1) Increase quantity & quality of rice 2) Improve well-being of consumers	✓				There is no Logframe nor any goal which is verifiable.
OVI 1:					
OVI 2:					
OVI 3:					
PURPOSE:					
OVI 1:					
OVI 2:					
OVI 3:					

START-UP 71 — TDD 75

CANNOT EVALUATE

BEST AVAILABLE COPY

PROJECT NUMBER: 730-11-140-291.03 CATEGORY: Income START: 65

COUNTRY: VIETNAM SUB-CATEGORY: Credit TDD: 74

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Rural development w/ income redistribution and increasing the efficiency of production of agricultural	✓	—	—	—	WORTHLESS for EVALUATION
OVI 1: production ↑ due to more inputs	✓	X	X	X	" "
OVI 2: exports ↑ due to credit	X	X	X	X	" "
OVI 3: war-torn areas rehabilitated	X	X	X	X	" "
PURPOSE:					Outputs have specific Q, Q, & T. But purpose & goal OVI's do NOT!
OVI 1:					
OVI 2:					
OVI 3:					

EVALUATION TIME NOT POSSIBLE!

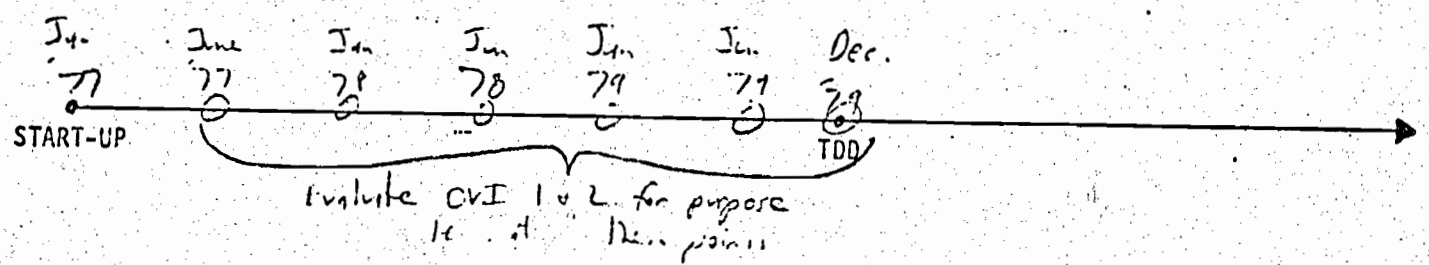
'65 START-UP — '74 TDD

BEST AVAILABLE COPY

PROJECT NUMBER: 497-13-899-225 CATEGORY: INCOME START: 77

COUNTRY: Indonesia SUB-CATEGORY: General Agriculture Institution Building TDD: 79

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Establish a sound co-op. development in Indonesia to improve living conditions & reliability	✓				
OVI 1: Increase co-op membership Establish new co-ops	X	X	X	Y	VERY VAGUE
OVI 2: Increase income of small farmers in co-op area	✓	X	✓	X	" "
OVI 3: Improve living condition of poor in co-op area	✓	X	X	X	" "
PURPOSE: Institution building in national co-op office Establish organization which will develop goals not forma co-op.					
OVI 1: Trained coop staff which will identify, plan, implement Agribusiness projects	X	✓	✓	TDD-2 to TDD	Half-year targets given for the entire project.
OVI 2: Project evaluation system for coops established	X	✓	✓	"	
OVI 3:					



BEST AVAILABLE COPY

PROJECT NUMBER: 497026500

COUNTRY: INDONESIA

CATEGORY: INCOME

SUB-CATEGORY: General Agriculture Institution Building

START: 77

TDD: 82

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Increase agricultural production, rural income, and employment opportunities	✓	X	X	X	TERRIBLE! The jump from OVI's to goal is tremendous
OVI 1: Agricultural planning will be superior to prior planning.	X	X	X	X	
OVI 2: Projects will be more efficiently planned and executed with a higher degree of project success.	X	X	X	X	
OVI 3: A large % of projects will benefit women and rural poor	✓	X	X	X	
PURPOSE: Upgrade the planning and programming capability of the Department of Agriculture	X	X	X	X	UNDEFINED
OVI 1: 380 trained participants will be performing planning and data collecting	X	✓	✓	TDD	
OVI 2: Agricultural data info. center will be established and maintained	X	X	✓	TDD	
OVI 3:					

'77 START-UP — 82 TDD

purpose OVI's 1) & 2) BVT

BEST AVAILABLE COPY

PROJECT NUMBER: 625-11-120-9176φ

CATEGORY: Income

START: 1975

COUNTRY: Sahel countries

SUB-CATEGORY: General Agriculture

TDD: 1980

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: To help achieve a higher or more secure standard of living for the people of the Sahel by increasing agricultural production	✓				
OVI 1: Increase agricultural productivity	✓	X	X	X	NO DATES
OVI 2: Higher health level	✓	X	X	X	
OVI 3: Increase purchasing power	✓	X	X	X	
PURPOSE: Establish and/or strengthen the capacity of meteorological & hydrological services in Sahel countries to provide timely hydrological, climatic, and other data [illegible]					
OVI 1: Lots of OVIs listed, none getting at impact				X	
OVI 2:				X	
OVI 3:				X	

START-UP 1975 — 1978 evaluation [illegible] — TDD 1988

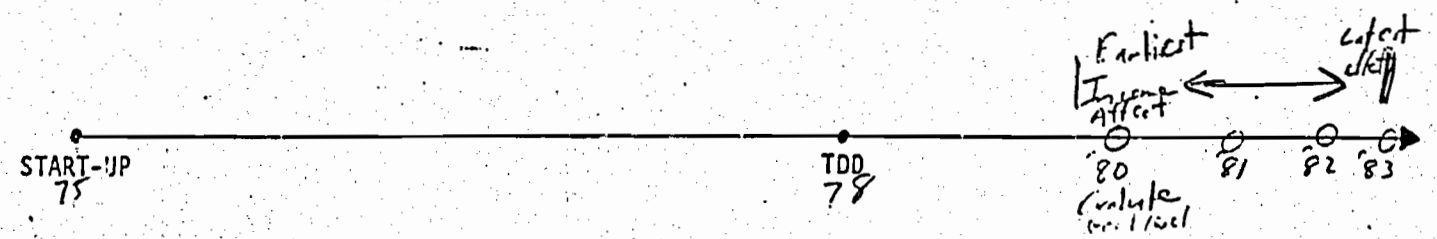
SINCE NO DATES IT'S IMPOSSIBLE TO SAY WHEN the OVIs shall only.

BEST AVAILABLE COPY

PROJECT NUMBER: 526 0113 CATEGORY: Income START: 1975

COUNTRY: Paraguay SUB-CATEGORY: Credit TDD: 1978

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: To increase the net income of the small farmer.	✓				
OVI 1: 50% ↑ in net income for farmers participating for 5 yrs (some case of 25%)	✓	✓	✓	TDD+2, to TDD+5	GOOD
OVI 2:					
OVI 3:					
PURPOSE: Organize independent, self-sufficient, credit organizations and cooperatives					
OVI 1: 40 rural + 24 urban local-ing and well-managed chartered credit unions w/ 20,050 + 14,600 members respectively	X	✓	✓	X	NO TIMES, I assume
OVI 2: Local cooperatives will be well-managed and efficiently serve their small farmer members.	X	X	X	X	TDD
OVI 3:					



BEST AVAILABLE COPY

PROJECT NUMBER: 596-22-150-068 CATEGORY: INCOME START: 1976

COUNTRY: South America SUB-CATEGORY: CREDIT TDD: 1980

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: To improve the well being of the rural poor. To increase the participation of the rural poor in the benefits fm. agribusiness	✓				
OVI 1: Health improvements in rural poor due to ability income & more nutritional products	✓	X	X	X	IMPOSSIBLE — The M.O.V. is by 'agribusiness reports'
OVI 2: All the OVIs come from increased agribusiness activity:	✓	X	X	X	
OVI 3: producing ↑; income ↑; employment ↑; salaries ↑;	✓	X	X	X	
PURPOSE: Develop non-traditional agribusiness activities which provide benefits to the rural poor. Improve institutional capability to finance agribusinesses	✓				
OVI 1: Loan funds will disperse to small farmers.	X	X	X	X	Though I could not find
OVI 2: 5,150 full-time & 2,000 to 3000 seasonal jobs for the rural poor created directly by projects financed.	✓	✓	✓	X	a date, I believe 1 & 2
OVI 3: 50% ↑ in income of rural poor benefitting by LAAD financed projects.	✓	✓	✓	X	can be evaluated by 1980; the ↑ in income would be later.

START-UP '76 — TDD '80 EVALUATE

BEST AVAILABLE COPY

PROJECT NUMBER: 5380007

COUNTRY: Caribbean

CATEGORY: Income

SUB-CATEGORY: Credit

START: 76

TDD: 81

LOGFRAME INFORMATION	MANDATE IMPACT	QUANTITY	QUALITY	TIME	COMMENT
GOAL: Increase standard of living of the small farmers in the of Caribbean Group	✓				
OVI 1: Reduction in rate of growth of agricultural imports.	X	X	X	X	VAGUE
OVI 2: Increased food production for internal consumption.	✓	X	X	X	"
OVI 3: Decrease in traditional export crops as % of total small farmer production	X	X	✓	X	"
PURPOSE: Increase productive resources & services available to small farmers.					
OVI 1: Agricultural production credit at $4.0M	X	✓	✓	TDD-1	Good, but miss impact
OVI 2: $2.5M of agricultural inputs for non-traditional crops	✓	✓	✓	TDD-1	
OVI 3: Feeder Roads of $3M	X	✓	✓	TDD-1	

START-UP 1976 — TDD-1 1980 — TDD 1981

Evaluate here as goal impact

BEST AVAILABLE COPY

THE EFFECTS OF SCOPE & METHODOLOGY CHOICES ON DURATION & INTENSITY

The "minimal" and "optimal" scopes for AID's project impact evaluations will involve different degrees of intensity in an evaluation, though the additional elements of the "optimal" scope need not increase the length of the period required for evaluation.* The additional scope elements in the "optimal" version (e.g., measurement of spread effects, unintended impact, etc.) could be carried out parallel to the assessment of the project's planned primary impact in most project evaluations. The number of evaluation team members required to carry out parallel assessments of primary and spread, etc., would be greater than the number required to assess only planned primary impacts.

With reference to methodological options, efforts to examine causality with a formal hypothesis test, or by using the tracer/*modus operandi* method, rather than to simply establish association between cause and impact, are likely to both add time and intensity to the studies. An experimental or quasi-experimental approach requires effort before the project "treatment" is applied, e.g., identification of control groups, randomized assignment, etc. Use of a tracer approach also requires time that would not be needed if only association were required by the evaluation. The tracer approach does not involve pre-treatment work, however, and to some degree it could be carried on parallel to the measurement of planned primary (and other) impacts. Thus, this methodological choice need not lengthen the study period significantly.

A second aspect of methodological effect will arise as the result of choices made for a specific project impact evaluation concerning the basic data collection approach, e.g., survey or case study, the sampling plan (where used), the

---

* This applies to all three options concerning "sustainability" that involve one evaluation period. For the two-stage option, the first period would be roughly equivalent to the duration of any of the other options, while the follow-up period would be shorter, e.g., 1/4 to 1/3 the time required for the initial impact evaluation in most cases.

BEST AVAILABLE COPY

approach to gathering field data (observation or interview) and the analysis plan (simple versus complex, and manual versus automated). (The total time required for a study is the sum of the time required for each of its parts.) Table III.B.-1 identifies in a rough way how the costs of one type of study (a sample survey) tend to divide themselves. It also suggests the way in which time is allocated to different methodological aspects of such a study. In examining the time requirements for specific project impact evaluations, AID will need to consider all of the relevant steps involved in carrying out whatever methodological approach it selects for specific evaluations.

Large projects that are spread across extensive areas will in general require more time than will smaller projects. On the other hand, any project evaluation that seeks data about subgroups (strata) within a population (of people, farms, etc.) will require more time and a larger total sample (if taken), than will studies that do not investigate subgroup effects. Sample size or the number of case studies that are involved in an evaluation affect both the manpower requirements and the level of data analysis in an evaluation, in roughly the ways identified above. Taken together, these factors suggest that one cannot make a firm rule that relates the duration and intensity of evaluation to the size/type of project. There are study designs which if applied to a small project/population could take much longer to execute than an alternative design used to study a large project/population.

BEST AVAILABLE COPY

TABLE IV.B.1-1

ESTIMATED ALLOCATION OF COSTS IN A SAMPLE SURVEY
FROM AID's MANAGER'S GUIDE TO DATA COLLECTION, BY MOLLY HAGEBOECK, *et.al.*, PCI

SURVEY STAGE	% OF TOTAL STUDY COST	BREAKDOWN OF STAGE INTO ACTIVITIES/ACTIVITY COSTS	
		ACTIVITY	% OF COST
I. PLANNING STAGE	15%	1. Develop Sample Plan	15
		2. Develop Research Instructive	35
		3. Prepare Training Materials	30
		4. Determine Data Processing Requirements	20
		TOTAL	100
		ACTIVITY	% OF COST
II. SURVEY PREPARATION	15%	1. Select Interviewers	5
		2. Select & Prepare Sites	15
		3. Establish Field Procedures	30
		4. Train Field Staff & Interviewers	35
		5. Develop & Test Data Processing Methods	15
		TOTAL	100
		ACTIVITY	% OF COST
III. FIELD IMPLEMENTATION	25%	1. Provide Coordination & Administration	20
		2. Undertake Sampling & Quality Control	20
		3. Interview & Maintain Data Collection Quality Control	50
		4. Complete Preliminary Data Processing	10
		TOTAL	100
		ACTIVITY	% OF COST
IV. DATA ANALYSIS	15%	1. Construct Model & Analysis Plan	45
		2. Define Measures	30
		3. Execute Measurement & Validation	25
		TOTAL	100
		ACTIVITY	% OF COST
V. REPORTING	20%	1. Write Report	70
		2. Present Findings	30
		TOTAL	100

BEST AVAILABLE COPY

5

BEST AVAILABLE COPY

Impact Evaluation Approaches & Methods

CHAPTER FIVE

The question of evaluation approaches and methods cannot be addressed for AID in the absence of information about:

- What is to be assessed or measured by an evaluation;
- The standards of evidence that are to be applied in the measurement effort.

Annex II.A. of this report provided an extended discussion of the types of impact AID may need to measure through its evaluations. The Annex also suggested that some standardization of measurement approaches might be possible, since 80% of the 85% of AID projects that identify mandate objectives and indicators focus on only seven (out of 27) indicators and objectives found in the introductory sections of the Foreign Assistance Act. The question of what standards of evidence are to be applied in AID's project impact evaluations, however, remains to be addressed.

The information AID wants to have available to support future Congressional presentations is evidence of impact that will not easily be dismissed. The evaluations AID has conducted in the past have rarely been seen as providing convincing proof that a hypothesis, on which one or more AID projects is based, is valid or invalid. Even after evaluations are undertaken, AID personnel suggest that the conclusions are sometimes argued or ignored. The cases in which conclusions are argued are more important for the present discussion than the cases in which conclusions are ignored, since a number of factors may explain the latter phenomenon. When AID finds itself arguing internally about the conclusions reached by an evaluation, the Agency is in effect demonstrating

BEST AVAILABLE COPY

that it does not have an agreed-upon standard of "proof" or evidence, and that the absence of an enforced standard causes problems.

The question of what standard of evidence is to be used in AID's project impact evaluations cannot be answered without reference to the external audiences for these evaluations. Whatever standard of evidence AID uses, it must satisfy both the Agency and Congress, as well as other external readers. There is little to be gained by conducting evaluations that are rejected because they fail to demonstrate in an adequate manner the impact they claim occurred. There are, however, a number of difficulties with the idea of setting and enforcing a standard of evidence for AID's project impact evaluations:

1. It is not clear whether AID has, or will utilize, an ability to discriminate between strong and weak methods, or between adequate and inadequate evidence. There is anecdotal information to suggest that this capacity is not well developed in AID. Evaluations which have been brought into question on methodological and factual points often receive circulation without comment on "flaws" that have been identified either inside AID or by parties outside the Agency;

2. The task of defining a standard of evidence to be used in AID's project impact evaluations would itself be difficult. Too high a standard might not be met in practice; too low a standard might result in the rejection of some studies as inadequate proof of project impact. The expertise required to develop such a standard might not be available either in AID or through arrangements with non-AID personnel. And it might simply take too long and cost too much;

3. Enforcement of a standard of evidence would also be difficult. AID's capacity to discriminate differences in this area would need to be developed, or non-AID personnel would be needed to undertake quality reviews for the Agency. This task too might simply take too much time and cost too much to pursue over the long run.

If AID is not prepared to develop and enforce standards of evidence for its project impact evaluations, it should probably not undertake a major exercise to specify the methodologies to be used either. On the other hand, if AID is prepared to define and enforce standards, an effort to prepare clear and usable guidance on methods will be a good investment.

BEST AVAILABLE COPY

AID can, of course, spend time developing guidelines that evaluation teams have the option of using if they wish. Alternatively, AID could simply refer evaluation teams to a bibliography of existing books and papers on evaluation methods and measures that might be applied to the assessment of specific types of impact. However, AID should recognize that unless it is prepared to review each evaluation to make sure that appropriate procedures are being used, it has no guarantee that they will be. Guidelines, with no enforcement plan, may be a make-work exercise.

AID has three realistic options:

- It can ignore the whole issue of standards;
- It can set and enforce methodological standards and standards of evidence;
- It can try to ensure quality by only using evaluation personnel who have a proven record/reputation for producing high quality information in developing country situations.

If AID chooses the first course or the third, it should not invest in a major effort to identify and prepare guidelines on each of the measures and field data collection methodologies that are appropriate for assessing project impact. At most, it should develop a good bibliography. If AID, on the other hand, decides to tackle the problem of setting and enforcing standards of evidence for its evaluations of mature and terminated projects, some effort in the area of measures and methodologies will be warranted.

Before proceeding to the discussion of the methodological aspects of AID's project impact evaluation task, a few general comments are in order.

First, let us note that impact evaluation is not a new idea. The conceptual approaches for assessing the effect, of an intervention derive from science and have been in use for a long time. In the past decade, there has been a good deal of work undertaken that addresses the application of these concepts in social science and development situations. The work has been of two sorts:

BEST AVAILABLE COPY

(1) Efforts to define what kinds of information are needed to assess impact in social programs (including numerous attempts to devise appropriate measurement schemes), and (2) impact evaluations of social and development programs. The individuals who have worked in these areas are scattered far and wide as are the projects on which they have tried their evaluation approaches.

While there is a growing literature on impact evaluation, it is not clear that the work in this area is relevant to AID's needs. What AID appears to need is an eclectic approach to measuring the effects of what it has done. Some ideas in the literature may be helpful to the Agency, by suggesting what measurement approaches do/do not work.

Despite the growing availability of information on the methods for impact evaluation that have been tried, and the problems that exist with most approaches, little of this material seems to find its way into AID/W, let alone the Missions. The only guidance that seems to be widely available in AID is that which is provided in the Project Evaluation Guidelines (an annex to the Handbook system), and the AID project evaluation handbook. Neither of these documents adequately reference and extrapolate the lessons of past impact evaluation efforts, internationally or in this country. Further, much of the information in these two documents tends to be general and theoretical (e.g., they contain classic evaluation design strategies, the Logical Framework approach, etc., but they do not offer examples of when and how AID has tried to measure such results as income changes, changes in health or nutritional status, etc.).

An AID effort to bring its staff up-to-date on the literature/lessons of impact evaluation need not be a major undertaking. In the first place there is already a good deal of information that has been summarized in previous "state of the literature" studies.* In addition, the Agency has a substantial, but

---

* PCI carried out one such review for the Peace Corps two years ago under a contract to assist that organization in deciding how its impact might be measured. Works such as this provide a good starting point for an update by AID.

BEST AVAILABLE COPY

widely scattered collection of materials that may be pertinent to such a review. This material ranges from evaluation study reports (some of which contain material on methods and study implementation problems), critiques of specific evaluation efforts, guidance and training material on evaluation prepared for its DSP program, its overseas Training of Trainers in Management seminars, etc., and all sorts of written pieces on specific information gathering approaches, e.g., employment surveys, health indicators and measures, etc. (many of which are contained in 5-25 page unbound papers and memos that can be rather difficult to locate).

The AID evaluation studies may turn out to be one of the best reference materials available to AID, since those which present solid data will tend to reflect the presence of "proven" data gathering/analysis methods. AID will get further, faster, in the area of impact evaluation if it is able to use approaches to field data collection that have proven feasible in an AID project environment. Wherever AID can avoid "reinventing the wheel", either in its study designs or for more fine-grain details such as the working of questions about employment on a questionnaire, it will save time and money.

Once AID begins to identify models/measures that can be used to evaluate the impact of its projects, it will have a second important task: To make sure that such models/measures are reused as often as appropriate, rather than lost in report appendices (thus forcing possible "reinvention" at a later point).

There is a myth in AID, as there is in many other organizations, that suggests that only quantitative data constitutes solid evidence of an effect. An unexamined preference for quantitative data could prove detrimental to AID's project impact evaluations. The changes AID is seeking to bring about in the developing countries are by and large qualitative changes. Quantifiable indicators can be developed that measure some, but not necessarily all, of the aspects of change AID is interested in. Further, in some cases a good qualitative measure will give AID a better, more direct, sense of impact than will dozens of minute statistical facts. The measures of impact, and the methods to be used in securing impact data, should be tailored to the type of effect AID wants

BEST AVAILABLE COPY

to measure. In identifying measures and methodologies, AID will need to be on its guard against advice that is based on expertise in a certain approach as opposed to advice that squarely addresses the impact measurement issue at hand.

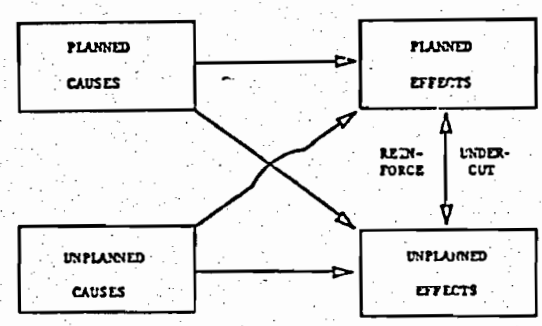
Based on all of the considerations identified above, together with the fact that AID has had a good deal of success in setting standards for the measurement of project progress at the Output level (i.e., the implicit requirement that factual evidence be secured on each of the objectively verifiable indicators/targets shown at this level in a project's Logical Framework) as well as some success in getting AID projects and Missions to comply with this standard, PCI recommends that AID tackle the problem of setting and enforcing standards of evidence for impact evaluations as well.

To facilitate an Agency effort to set standards of evidence and define appropriate methods and measures of impact, PCI has identified the main approaches that appear to be appropriate in Annex A to this Chapter. Annex B, which supplements the discussion of methodologies, addresses the question of what skills are required to carry out evaluations that employ the methods discussed in Annex A. One point made by this Annex is so fundamental that it deserves repetition: An evaluation that only examines the status of the project area/group after the project, does not tell us whether a change has occurred. A comparison of some sort is needed to establish impact.

BEST AVAILABLE COPY

IMPACT EVALUATION APPROACHES & METHODS

This Annex is structured around a diagram that was developed by PPC/DPRE, and which is familiar to most AID personnel. The diagram of project causes and effects is reproduced below:



In the following paragraphs, PCI reviews the methodologies that are appropriate for AID to consider in connection with its project impact evaluation. The order of presentation moves from planned effects to unplanned effects to planned causes and finally to unplanned causes.

1. Planned Effects

Here we are primarily concerned with project impacts, i.e., those effects which we hypothesize will occur after project services are provided. Effects at the Purpose, Goal and higher levels are included in this definition of impact.

As previously noted, projects will differ in terms of whether or not their impacts are substantively the same as those identified in AID's Congressional mandate. Some will yield mandated impacts at the Purpose level; others will

BEST AVAILABLE COPY

yield these impacts at the Goal (but not the Purpose) level. Still other projects will identify impacts that are not linked to the mandate objectives.

For all of the configurations of "impact chains" that can be identified for various AID projects, the basic question remains the same: Has there been important change? Conceptually, the question must be answered by making some sort of comparison. Information that describes a project situation (in impact terms) after project goods and services have been provided is a necessary ingredient, but it cannot stand alone. If all we have is data that describes the project situation *after* services are provided, we do not know whether the situation is any different than it was *before* the project began.

There are two general ways to compare a project situation after goods/services are provided to another situation in order to discern whether change has occurred. The first is to compare the project situation to another situation that is similar in all but one important regard: The comparison situation did not receive project services. A second approach involves comparing the situation after the project to a "baseline" description of that situation prior to the project effort. When comparisons of this sort are made, one additional type of evidence is normally secured: Proof that the goods/services that were planned in the project were actually delivered.

The conceptual requirements for establishing whether impact occurred are not hard to master. Rather it is the details of working out ways to make comparisons, measure levels of impact, and establish that the target group was exposed to the "treatment" that make the task of impact evaluation appear quite complex. The practical details of designing evaluation studies and developing appropriate measures of impact cannot all be addressed here. However, a few points and several references should be mentioned in connection with the evaluation of the primary and secondary impact of AID's project.

The question of what type of comparison is appropriate for determining whether change has occurred in a project situation cannot be fully addressed in the abstract. However, as a general rule, AID should assume that cross-sectional

BEST AVAILABLE COPY

studies (involving groups that did/did not receive project services) will produce better information at a lower cost than will before/after study designs. Before/after designs, although widely used, are inferior to cross-sectional designs both theoretically and on a practical basis. When used in conjunction with cross-sectional data, their value increases, since a mixed design can tell us both how significant a change occurred in the area/group served by the project, and something about how the process of change took place.

On the other hand, it is possible to identify, in a generic way, what evidence is required to demonstrate that project services/goods were actually provided as planned. Every project, in its Logical Framework, should identify this evidence in its indicators and targets at the Output and Input levels.

In addition to making an evaluation design decision concerning the type of comparison to be made (two-groups, before/after, or both), AID will need to decide on the proper "unit of analysis", e.g., communities, farms, households, etc., for each evaluation and the most appropriate measurement approach (e.g., survey, case study, etc.) and measurement techniques (e.g., interviews, physical measures, etc.). Recent Agency manuals and guidelines provide AID with a good deal of information on how these choices should be made, and what answers are appropriate in specific types of situations.*

a. Measuring the Planned Primary Impacts of AID Projects

Primary impacts are those which directly affect the target situation. Three clusters of planned primary impacts have been identified: Mandate impacts, intermediate impacts (which at higher levels are hypothesized to result in mandate impacts), and impact objectives that are not linked to mandate impact. Each cluster is discussed separately below.

---

* Chapters 4-6 of the "Manager's Guide to Data Collection", prepared by PCI for PPC, addresses these questions as does Chapter 12 of the "Agribusiness and Rural Enterprises Project Analysis Manual", prepared by PCI for DS/AGR.

BEST AVAILABLE COPY

(1) Measuring Mandate Impacts

The impact objectives identified by Congress in the FAA suggest that what is sought from development assistance is a *qualitative change* in the lives of the poor majority of the world's people. The ways in which the world's poor are to be "better off" are alluded to in statements about their "quality of life", meeting their "basic needs", living lives of "decency, dignity and hope", being "engaged in productive work" and "participating in/influencing decisions that shape their lives".

At one level, each of us knows what these statements mean and can subjectively detect when qualitative changes of this sort occur. While at a general level we tend to know when project situations are improving, difficulties arise when we attempt to use our subjective impressions as "proofs" of the change we sense is occurring. One of the difficulties relates to calibration: While people tend to agree on gross changes, their impressions are not so well aligned at a fine-grain level. They disagree about whether small changes have occurred, and whether a small, albeit agreed upon change, is good enough. The second difficulty relates to what they look at when they are trying to judge whether there has been a change, and how they weigh the various aspects of their observations. A third problem is one we call "observer bias": Our tendency to see what we want to see, or what we believe we are seeing.

If the difficulties of calibrating the judgments of people could be resolved, evaluations which depended upon people observing total situations to assess change would be the optimum approach for detecting qualitative differences of the sort intended by the Congressional mandate. Fortunately or unfortunately, we will probably never arrive at a point where we can get highly reliable (well-calibrated) information out of subjective impressions about the type of change AID will be trying to assess in its project impact evaluations.

The alternative for securing information on changes in the social domain is generally considered to be the development of objective indicators of change,

BEST AVAILABLE COPY

for which targets can be set and measures made. In keeping with this type of approach, Congress (and AID) has begun to identify objectively verifiable indicators for assessing achievement in mandate impact terms, and AID has suggested targets for several of these indicators by its partially quantitative description of who is/is not a member of the world's "poor majority". In these terms, AID succeeds, at a macro-level, when individuals move from one side of this arbitrary line to the other on one or more indicators.

In some projects even the mandated impacts of project action may flow in a means-ends (hierarchical) way. AID in its response to Congress in 1975 identified three of the mandated objectives as "final outcomes". These were income, health and nutrition. In this document AID characterized other mandated impacts as "means" to these "final outcomes". The mandated impacts characterized in this fashion were literacy, increased production, and changes in population growth rate. Participation, a final mandated impact, was treated both as a process and a measure of the distribution of access to "means" and "ends" in the AID report.

This view of mandate impact suggests that, in addition to evaluating whether projects have achieved impact in terms of stimulating mandated "means" (changes in literacy, production, population rates and process aspects of participation), AID will need to follow-through and assess whether these "means" resulted in "final outcomes" (i.e., changes in health status, nutritional status, or income). Figure IV.A.1 displays the various patterns (and measurement requirements) that follow from AID's discussion of mandate impact in this document. (The flows implied by the criteria AID prepared for assessing progress and commitment during 1976-77 are somewhat more complex, as are the flows implied by the Logical Framework of Congressional statements prepared during the course of this study.)

As Figure IV.A.1 and the discussions which suggested it indicate, there is an implicit requirement to assess both the outcomes of development assistance and the processes by which those outcomes are generated.

BEST AVAILABLE COPY

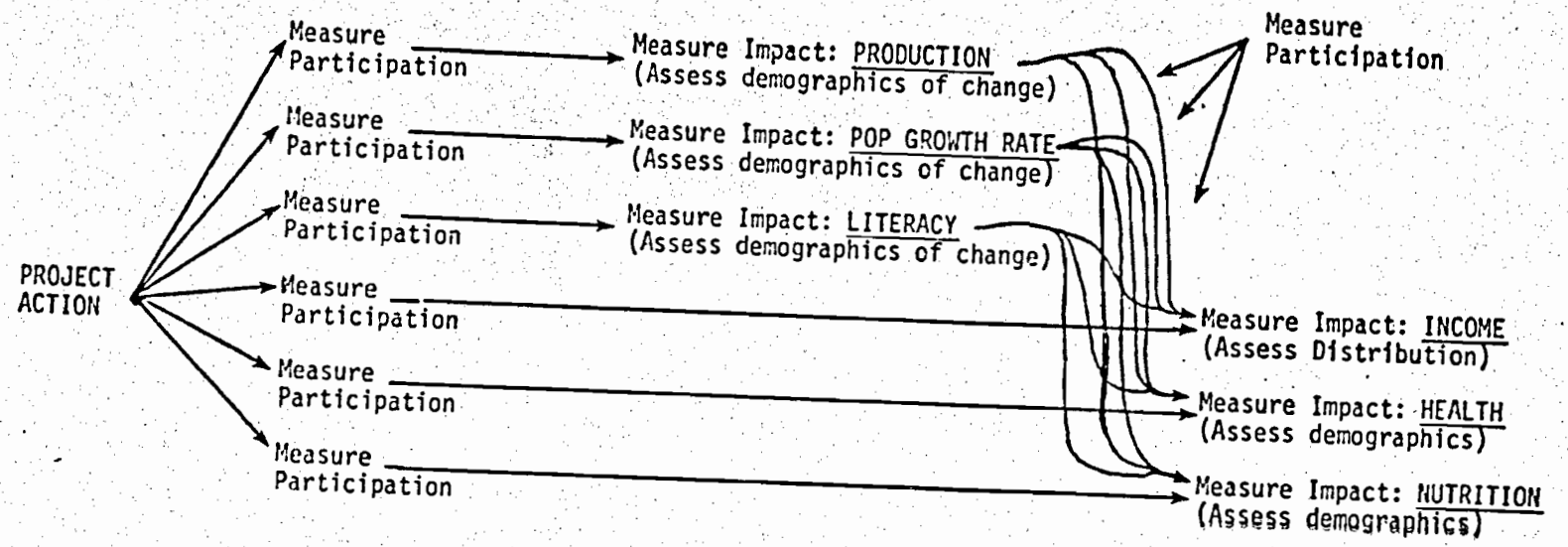


Figure V.A.1: As the diagram suggests, evaluators may find that the "chain of impacts" from a single project may require the measurement of several mandated impacts.

Nevertheless, AID's starting point for evaluating impact is, by definition, an area in which it has identified indicators and set targets. Beyond that, the Agency must begin to examine what aspects of its mandated impact are not captured adequately by the current set of indicators, and what targets/measures are appropriate for indicators that have not yet been associated with verifiable targets. A number of the specific measurement approaches that are appropriate for assessing whether change has occurred on the targeted mandate impact indicators have already been addressed in other AID contracts/products.*

Since the mandate impact objectives cut across a large number of AID's projects, there is an enormous opportunity for savings (and for strengthening AID's overall development effect) by *standardizing the definitions and the measurement approaches used in evaluating projects at the mandate impact level,* to whatever degree proves possible. Some of the definitions AID is already using are universal in character, e.g., calories. On other indicators such as income, AID may find that differences between rural and urban areas, and from culture to culture make the development of universal definitions difficult (though area or regional definitions may be possible). AID is already moving in this direction, and further standardization should help to simplify the problem of how project impact evaluations are to be conducted.

---

* Exhibits IV.A.1 and IV.A.2, which follow this Annex, offer two examples of the preparatory work in this area already completed under AID contracts and other agreements. Exhibit IV.A.1, the introduction to "An Overview of Economic & Data Analysis Techniques for Project Design", prepared for AID's DSP Program by Dr. Samuel Daines of PCI, identifies measurement approaches that can be used to secure data on income, health and nutrition status, population growth rate, production and literacy for AID's target groups. Exhibit IV.A.2, an Annex to "Feasibility & Application of Rural Development Participation: A State-of-the-Art Paper", prepared by Cornell's Rural Development Committee (Cohen & Uphoff) for DS/RAD, identifies potential measures of participation (both as benefit distribution and as a process).

(2) Measuring Intermediate Impacts & Impacts Not Linked to the Mandate

As suggested earlier, AID projects sometimes produce impacts that are either subordinate to or not linked to mandate objectives. PCI's review of the DIS printouts of AID's project Purpose and Goal statements, and ATAC's (The American Technical Assistance Corporation) work on a "goal hierarchy" for the Agency both indicate that these non-mandate impacts cluster into families. ATAC's "goal hierarchy" suggested three such families of non-mandate impact: Institution building objectives, sectoral objectives, and national development objectives. PCI's review of the DIS printouts adds a fourth category, security objectives, to this list. Further examination of AID's projects may suggest that there are additional categories and/or that categories break down into subsets which are themselves clusters of potentially measurable impacts.

For at least one of these categories, AID already has approaches for assessing performance. The institution building/institutional viability assessment work carried out by Dr. Milton Esman, MUCIA and PCI several years ago provides AID with a starting point for developing evaluation procedures for examining projects that identify "self-sustaining institutions and/or services" as an intermediate or non-linked impact. Evaluation of sectoral and national development objectives may also be an area where the "homework" has, at least in part, been done. Methods of assessing sectoral and national level development in some dimensions has been the subject of AID and World Bank analyses in the past. In addition, the security objectives of certain AID projects may be measurable with techniques developed to assess both "preparedness" and "political stability" in other situations.

The definition of fully appropriate ways of assessing intermediate impacts, and impacts that are not linked to mandate objectives, depends on further specification of what is to be assessed (i.e., a more detailed review of AID's portfolio and an analysis of the non-mandate objectives therein). Cursory review does, however, suggest that measurement techniques of the type AID may ~~require~~ have already been developed, and to some extent, tested. Since in

principle the repeated use of known measures will bring down the total cost of carrying out project impact evaluation, AID should consider multiple-time use of the measures that prove effective in the first round of impact evaluations.

b. Measuring Secondary Impacts & Spread Effect

There are a great number of secondary impacts that can result from project activity, too many to identify fully here or to gauge exactly what measurement tools may be needed. Suffice it to say that in the past (in AID and elsewhere) the range of secondary effects has been extensive, and included environmental impacts, changes in family/social structures and relationships, intersectoral interactions, soil changes, etc. AID has two options in terms of preparing to assess such secondary impacts:

- Undertake to "model" some of the most typical interactions that have been identified to date in socio-ecological investigations, i.e. use simple diagrams to define common secondary impact areas and identify *a priori* the most appropriate measures for clusters of such impacts, or
- Decide what measures are possible/appropriate on a case-by-case basis as secondary impacts are identified, and keep "lab notes" on the ways various evaluation studies went about measuring such effects.

Two types of spread effects that are interesting because of the different methods they suggest are replication (of both the processes used in a project and its substance) and reapplication or use of only the project processes:

- Spread/Replication

  This form of spread will normally need to be assessed by observing (making measurements in) a different area or with different groups of people than were measured to characterize the project's target area/group. The measures needed will essentially be the ones used to measure the strength and direction of impact on the target group; in the evaluation of spread/replication, AID will simply need to apply them to the area/group in which replication is to be examined.

(AID should note that it cannot use the same area/group to examine replication that it uses to make the "treatment"/"no treatment" comparison that establishes whether there was a change in the target area/group.);

- Spread/Reapplication-Extension

  This form of spread involves the identification/examination of permutations of the project (the reuse of project process to achieve a different objective). In evaluating this type of spread, AID's main problem is likely to be one of identifying such reapplications/extensions. (The approaches for making such identifications are discussed below in the paragraph on "unplanned effects".) Once identified, reapplication/extension forms of spread should be measurable using the types of techniques employed to assess primary impact. The difference will be that in applying the measurement processes, AID will be focusing on whether the reapplication successfully facilitated the achievement of a different objective. (How well prepared AID is to make such measurements will depend on what the objective of the reapplication/extension effort turns out to be.)

c. Unplanned Effects

As suggested in Subsection a of this section, the difficulty with unplanned impacts is identifying them. The measurement tools that are needed, once unplanned impacts are identified, will for the most part be those AID will need to identify planned primary and secondary impacts.

There are three main ways of identifying unplanned impacts (and some types of secondary effects) that AID should consider if it decides to include these impacts as part of the scope of its project impact evaluations. They differ in their comprehensiveness, feasibility and cost:

- The Perfect Picture Option: Measurement of all aspects/dimensions of a situation before and after project action is taken. This approach would capture all change, unpredicted and predicted. Except in very rare instances such an approach would, even if feasible, be prohibitively expensive. It is, however, the only approach that aspires to being comprehensive;

- The Tracer Option: Identification of as many of the changes induced by each project element (Input-Output cluster) as possible, including the tracing of second-generation effects, e.g., if an output (credit provided) led to investments in increased agricultural production, the second-generation effect would deal with the uses made of the income that resulted from increased production, etc.). Some of the changes identified using this approach will have been anticipated, others will not. The comprehensiveness of this approach depends on the persistance of the evaluator and the degree to which he/she can gain the cooperation of local informants who know what is happened as a result of AID's project.*

- The Good Listener Option: During the course of a project the people associated with the effort (both the implementation team and members of the target group) may have noticed unanticipated effects. The approach suggested by this option involves conducting open-ended interviews with some portion of this group before the evaluation formally begins. To the degree that these individuals noticed and can describe changes they had not expected, the evaluation team can build measures of such changes into the study design. The comprehensiveness of this final approach is highly uncertain since it does not formally structure or model the process of discovering unanticipated impacts.

Either of the latter two options appears feasible, and both appear to be reasonably low in cost. Because the Tracer Option has more structure, and aspires to provide a more comprehensive answer, it should be expected to cost more than the latter option.

With any of the options above, the cost of actually measuring the strength and direction of an unplanned change will vary significantly depending on whether the unintended consequence was primary or secondary in nature. If the unexpected impact affected the primary targets of the project, there will be additional measurement costs, but the measures taken will be of the same people, institutions, areas, etc. If the unplanned impacts are secondary, the costs may be higher since it is likely that a different area or group will have to be measured to determine the strength and direction of an unplanned change.

---

* A diagrammatic approach for carrying out the tracer option was described in PCI's report to AID on the horizontal expansion of the evaluation system. The term used to identify the approach for tracing impact and connecting it to project action was "connectwork".

BEST AVAILABLE COPY

2. Planned Causes

The planned cause of project impact is project action. The Outputs provided by a project constitute the "treatment" that AID hypothesizes will bring about a chain of impact effects. Assessment of whether the planned cause is associated with a set of measured impacts can be addressed in one of two basic ways. We can examine one of two propositions:

- The planned cause (project Outputs) existed, and the impact occurred. This proposition is an assessment of the association or correlation of the two elements of a predictive hypothesis;
- The planned cause is not only associated with the measured impact, it is also the true cause of that impact. This proposition requires the elimination of alternative possible causes. It assesses an explanatory hypothesis concerning the relationship between cause and effect.

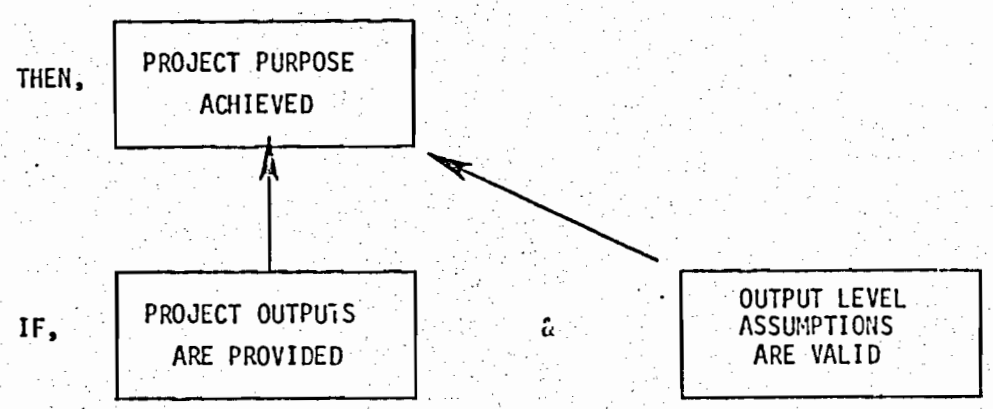
In the subsection on the scope of project impact evaluations, PCI noted that AID has not yet committed itself to "test" the second of these propositions. On the other hand, AID does discuss the value of knowing that a planned cause was the true cause of project impact. In the scope section, PCI concluded that *at minimum*, AID's project impact evaluations would be required to demonstrate the first proposition: "Treatment was provided, as planned". In this subsection we thus examine what is needed to meet the minimum scope. In the following subsection on "unplanned causes" we present information that pertains to testing the second proposition: The cause was the true cause. To the degree AID elects to examine "unplanned causes", it will in effect be moving from a minimal to a more optimal position *vis a vis* the general requirement to "demonstrate that such changes are associated with, or caused by AID's development assistance" (Shakow memorandum, 6/25/79).

The predictive hypothesis concerning the relationship between project Outputs and project Purpose (or between Purpose and Goal), involves two action elements: Direct action (expressed in the narrative summary of a project) and indirect action (expressed in the Assumption column of a Logical Framework).

Mathematically, then, the project's predictive hypothesis states:

$$[\Sigma\ (Os)\ +\ \Sigma\ (A_0)] \rightarrow (P)$$

The sum of the Outputs (which may be multiple) plus a set of valid Output level Assumptions yields project Purpose. Pictorally, the hypothesis takes the following form:



Thus, to ascertain whether "treatment" or the planned cause of the first level of project impact (Purpose) was provided, AID needs to (1) measure actual Output achievement versus planned Output achievement, and (2) validate the Output level Assumptions made by the project.*

---

* The logic of measurement described above also implies that, when AID moves from the Output-Purpose level to the assessment of higher level impacts, it will need to validate Assumptions made at these higher levels in order to establish that the preconditions (specified as being necessary and sufficient) for bringing about higher level impacts, were in place. Hence, to verify that the "treatment" needed to bring Goal level impact into being was provided, AID would need to demonstrate that (1) Purpose was achieved, and (2) the Purpose level assumptions were valid.

It should also be pointed out that the verification of "treatment" described in the text and footnote above differs from what AID literature calls "assessing the contribution of Purpose to Goal, etc.". That literature is implying a need to sort out what the relative importance of project action and project assumptions might have been. The verification procedures

If a project has been designed and implemented according to the Agency guidelines, the first of these steps should not only be fairly straight forward, but it should have been largely completed in the course of regular project evaluations. Objectively Verifiable Indicators of each project Output are specified at the time of project design. Targets (level and type) of achievement against each of these indicators is also a design requirement. AID's Project Evaluation System (PES) specifies clear requirements for identifying and recording any differences between planned and actual Output level achievement. Furthermore, AID reports that in most projects that are evaluated internally (following the PES instructions) such measurements are taken and recorded. As suggested above, information from the regular PES that suggests project Outputs have been provided is a natural trigger for impact evaluation, i.e., the "treatment" from which impact should follow has been provided.

Validation of the Output level Assumptions for a project may/may not be addressed by regular project evaluations. In most cases it is probable that the on-site team knows whether its major assumptions are valid. However, AID's personnel appear to be less systematic about recording their validation of assumptions as part of the scope of project impact evaluations. The methods for validating the assumptions will vary, depending on the specific assumptions made by a project. In past reviews of AID's portfolio, PCI has noted that the assumptions made by projects are quite varied, ranging from items concerning government action to farmer motivation and from the stability of a price to the predictability of the weather.

While the assumptions (at various levels of the project logic) may range widely, the process of validating an assumption need not be complex. In most cases, AID will simply need to verify that some type of event or behavior occurred/did not occur (depending on how the assumption is formulated). In part this

---

discussed above require only that AID demonstrate that *all of the necessary ingredients were provided as planned;* there is no implied need here to try to figure out the "percent contribution" of a well-provided Output or a valid assumption. All that matters is that both can be verified.

BEST AVAILABLE COPY

evidence can come from data indicating that the objective at the next higher level was achieved. By deduction, we can infer that all of the assumptions were either valid or trivial, if the next higher level objective was in fact achieved. In addition to our deduction, we may also find that a lot can be learned by simply asking whether the assumptions were valid. (There is no requirement that we gather quantitative proof here, only an admonition to make a reasonable effort to check on the validity of the project's assumptions.)

If the objectives of the project were not achieved, the answer may lie in the assumptions. When we find projects did not achieve the impact that was expected, it may be necessary to dig deeper into the assumptions, and in some cases attempt to ascertain just how invalid particular assumptions turned out to be.* When this situation occurs, AID will need to review the assumption(s) in question and match them against such proven measurement techniques as it has identified, in order to come up with a workable approach for determining to what degree an assumption operated counter to expectations. The types of methods that will be needed cannot be fully examined here. As in the case of secondary effects, AID can either wait and see what is needed, or attempt to "model" the probable patterns of assumptions in order to prepare for later needs in this area.

3. Unplanned Causes

The notion of unplanned or "unanticipated" cause explicitly recognizes that the project hypotheses we generate evolve from our experiences and deductions, and that these experiences/deductions are sometimes wrong. We may think that A causes B. Yet, when we begin to test that proposition, we may learn that

---

* When we know that a project did not achieve its objectives, the "lessons" AID will draw can differ substantially depending on whether the reason for failure lay with inadequate provision of the required "treatment" (project action), invalid assumptions, or (if treatment was applied and assumptions were valid) an invalid hypothesis.

BEST AVAILABLE COPY

B occurs in the absence of A as well as in its presence. We are led to conclude that another factor must be involved, an unplanned/unanticipated cause of B.

The investigations that must be conducted to determine whether factors other than A (project action) explain the occurrence of B (project impact) are generally more sophisticated than those required to test predictive hypotheses.* When we conduct this type of evaluative research, we attempt to eliminate all explanations of B, other than A, as the cause of project impact. (In some instances we may also try to determine what factor, if not A, actually caused B in a specific situation. Identifying the true cause of B, when it is not A, is, however, more difficult than simply establishing whether or not A was the cause.)

There are a number of evaluation research methodologies that are used to examine the validity of an explanatory hypothesis. The most rigorous of these is the controlled experiment. In such an experiment, we attempt to isolate the cause of an impact by allowing only one factor, our hypothesized cause, to vary. By holding everything else in the situation steady (or constant), we reduce the possibility that another factor causes an effect to very small proportions. In a laboratory situation, this type of control can be exerted. In the everyday world, we cannot bring everything to a halt in order to test a hypothesis. The experiments we run in the world take a different form. In these experiments we attempt to find two perfectly matched situation/groups of people, and, using random selection, we decide which of those situations/groups of people, and, using random selection, we decide which we measure these matched situations/people to ascertain whether the "treatment" brought about a change in the situation/group that received it.

Experiments of this sort are extremely difficult to set up. In most social science research, the experiments we conduct are less than perfectly controlled.

---

* Sophistication as used here does not imply complexity. The correlation studies used to examine predictive hypotheses are often far more complex than a simple, elegant experiment designed to test an explanatory hypothesis.

BEST AVAILABLE COPY

Either we are unable to perfectly match situations/groups, or we are unable to randomly select which situations/groups receive a "treatment". Partially controlled experiments in the everyday world are called "quasi-experiments", they give us some, but not all of the information we seek about causality. Quasi-experiments are also difficult to set up and manage, but they can be carried out, even in the developing countries.*

In addition to the experimental approaches for examining cause and effect, there are other, less rigorous ways to determine whether project action appears to be the cause of a set of project impacts. Among these Rossi lists before/after studies, post-project followup studies, and the judgment of experts. Another approach that should be included is the *modus operandi* or "diagnostic reconstruction" approach. This approach involves an attempt to trace backward from evidence of impact to its source.** The method is akin to that used by a medical examiner who attempts to arrive at the cause of death by eliminating various possibilities.

The task of eliminating alternative explanations for the impact detected in project impact evaluations is a difficult one, no matter what evaluation research methodology is used. In the first place, most of these approaches require patience and discipline to set up. Secondly, the number of alternative causes that might explain a specific impact is significant. Rossi, *et. al.*, list and discuss eight different types of "alternative causes", each of which would need to be examined with care to eliminate the possibility that it, rather than AID's project action, caused an observed impact.

---

* AID's evaluation handbook and its "Manager's Guide to Data Collection" both describe the conditions/difficulties of this sort of experimentation, as do many good texts including Evaluation: A Systematic Approach by Peter Rossi, Howard Freeman and Sonia Wright (Beverly Hills, Sage, 1979) and Evaluation Research by Carol Weiss (Englewood Cliffs, Prentice Hall, 1972).

** PCI's discussion of downward linking "connectworks" (in its report to AID on horizontal expansion of the Project Evaluation System) employed this concept. Under the term *modus operandi*, it has been written up by Michael Scriven, and published in an annual volume of Evaluation Research Society papers.

BEST AVAILABLE COPY

One final point that AID should note when it is considering whether to attempt to test the explanatory form of its project hypotheses is that the comparison groups it established as part of its effort to determine whether change has occurred will, if well selected, be usable as "control" groups as well.

MEASUREMENT OF MANDATE IMPACT INDICATORS

(Editor's Note: This section is reproduced from An Overview of Economic & Data Analysis Techniques for Project Design & Evaluation, by Samuel Daines, for AID's DSP program, Part A, pages 9-12. The term mandate in this section refers to AID's wording of the mandate objectives/indicators in its 1975 report to the Congress, cited at the end of the section.)

In answer to the question "What Do We Seek to Achieve?" the MANDATE states:

> "Determining the suitable goal must, of course, be the responsibility of the developing countries; nevertheless AID must set targets as well if we are to assess our own performance. The benchmarks used to define the poor majority suggest themselves as targets. Could not AID, in cooperation with the LDC's help move the poor majority beyond these milestones in the next decade or so?" (6)

As a target the document explains, moving all of the poor majority beyond these benchmarks may not be feasible; but as objectives, the benchmarks, income, nutrition, and health; plus production, population, and education are outlined as the mandate objectives.

The next sections in Part A focus on the problems involved in defining these six objectives in terms which are measurable so that they can become operational criteria for project selection, design, monitoring and evaluation.

4. *INCREASING THE INCOME OF THE POOR MAJORITY*

a. Defining Income for the Rural Poor

With all of its conceptual and measurement problems, income is perhaps the best single indicator of welfare. It is, however, inadequate to capture enough of what Congress and AID mean by "welfare" to be used alone. While income is not a sufficient measure, it is a necessary dimension in estimating welfare.

The focus of the MANDATE on income as a principal benchmark may have left many intellectually unsatisfied in that it fails to capture the essence of development, but it has the advantage of being operationally usable in project selection and evaluation. Compare the project selection and evaluation difficulties of operating with the following definition in comparison to the MANDATE.

> "Authentic development aims at the full realization of human capabilities: men and women become makers of their own histories, personal and societal. They free themselves from every servitude imposed by nature or by oppressive systems,

BEST AVAILABLE COPY

they achieve wisdom in their mastery over nature and over their own wants, they create new webs of solidarity based not on a domination but on reciprocity among themselves, they achieve a rich symbiosis between contemplation and transformation action, between efficiency and free expression. This total concept of development can perhaps best be expressed as the 'human ascent'—the ascent of all men in their integral humanity, including the economic, biological, psychological, social, cultural, ideological, spiritual, mystical, and transcendental dimensions." (7)

This section investigates the difficulties involved with defining income in a way which captures as much of the welfare concept as possible, is measurable for the rural poor in LDC's, and adjusts for an may potential distortions as possible.

b. Money and Subsistence Income

If income is to be a useful measurement for the rural poor it must include the value of the food and other goods which are produced and consumed by small farm families. In some cases a large proportion of "welfare" comes from these home produced products. These products may be added into income if they are valued in money terms. The income definition for the rural poor must include income in-kind. In-kind income may come form a variety of sources, from home produced food or clothing, from payments in-kind for services rendered by members of the rural family, or from barter exchanges of farm produce.

Two difficulties arise with measuring this expanded income definition. First, selecting a unit price at which to value the in-kind commodities, and second estimating the quantities. Since both of these issues are more related to measurement methodology and data gathering than definition, and examination of the techniques used for confronting them will be postponed to the accounting and data gathering sections. Largely satisfactory valuations can be made if the quantities can be estimated.

A subsistence farmer may be defined as one who consumes most of what he produces and sells very little. It is important not to confuse subsistence and poverty; many very poor small farmers are not subsistence farmers, but sell almost all of their produce. For example, in Costa Rica, target group farms sell more than 90 percent of their produce, and the level of subsistence actually rises as farms increase in size. Simply because target group farmers produce and consume large quantities of the same product does not mean that they are necessarily "subsistence" farmers.

In the example given in Table 1, from Costa Rica, the target group farmers were principally corn and bean producers and corn and beans were also important consumption items. Table 1 shows that even when small producers consume the same type of commodities they produce they may sell their own produce and purchase the same commodities on the market for consumption.

Rural income estimates must include estimates of the monetary value of in-kind consumption and transactions.

BEST AVAILABLE COPY

Table 1

Subsistence "Income" on Poor Small Farms in Costa Rica
Value of Farm Produce Consumed as a Percent of Total Farm Product

Farm Size	% of Farm Product Consumed Target Group Farms (Less than US $150 net family income per capita)
0 to 1 Ha.	4.0
1 to 2 Ha.	5.3
2 to 5 Ha.	6.4
5 to 10 Ha.	5.8
10 to 20 Ha.	5.6

Note: Farm consumption valued at average farm gate prices for comparable products in the relevant Canton.

Source: Richard Kreitman, Costa Rica Agriculture Sector Assessment Working Papers, Agency for International Development, Washington 1976, p. 100 & 107.

c. Net Income

In the event that the rural poor are landless laborers the income estimate is relatively easy; their income is largely net of expenses because they are not businessmen. They may have some expenses related to their wage incomes, transportation, tools, etc., but relatively simple accounting procedures can provide reasonable net income estimates. For the small farmer, arriving at a net farm income estimate is no small task. The difficulty is not principally because the small farmer does not keep records and does not therefore have his own net income figure calculated; even if he did, the figure would be largely useless because of the difficulty of comparing net income figures for different farmers using different accounting conventions. The only way to obtain reliable net farm income estimates is to gather crop by crop input and output data and then reconstruct farm accounts using consistent accounting conventions for all farms included. In addition one should anticipate "other income" which can be substantial to very small farmers.

If the process of project selection, monitoring and evaluation is to be serious there is no escaping the necessity of estimating net farm income of target group families. AID and Congress realized the futility of aggregate national income averages, and the importance of farm and family level net income estimates in the MANDATE:

BEST AVAILABLE COPY

"Actual income can be difficult to measure. National averages for per capita income are inappropriate, of course, since they would place whole countries in or out of the poor majority. We need more microlevel data, ideally for individuals, that permit isolating persons in a given country with income below $150. But such data are scarce in LDC's and expensive to collect, particularly for the poor whose 'income' may consist largely of subsistence output produced and consumed outside the market economy." (8)

Since the Mandate definition of net income per capita is in the US $ at 1969 prices, there are difficulties in comparing measurements over time due to inflation, and comparing local estimates with the Mandate US $ standard due to conceptual problems with exchange rates. There are obvious computational methods for handling the inflation issue, but the exchange rate was in effect in the year in which the net income estimates were made, the exchange rate used in comparing the farming and landless populations with the Mandate standard of US $150 had a significant impact on the size of the target group.

Table 2

Proportion of the Rural Population Classed as Poor by the Mandate Standard of US $150 per Capita using Each of the Existing Multiple Exchange Rates 1973

Exchange Rate	Percent of Rural Population with Net Income Capita Below US $150 in 1973	
	Farming Population	Landless Population
Coffee Rate	51.3%	49.2 %
Mixed Rate	60.1	61.3
Free Market Rate	67.0	70.5

SOURCE: Richard Kreitman, Costa Rica, Rural Poor Profile, Agency for International Development, Washington 1976, p. 28.

Net income on small farms is possible to measure, but it is unfortunately a costly and painstaking accounting task. However, if Mandate projects are to be undertaken with the intent of increasing the net income of the target group, there is no way to avoid measuring it. If net income is the objective it will be impossible to answer the three underlying project questions posed in the introduction without direct measurement of net income changes or differences in target families. The MANDATE states the necessity of using income measurements as one of the prime project selection and evaluation criteria as follows:

BEST AVAILABLE COPY

"Targets for AID assisted programs and projects should reflect unique local circumstances, but to the greatest degree possible they should be cast in terms of their contribution in the long term–5 to 10 Year–goals of improving the status of the poor. Working in cooperation with the LDC's targets should be defined in terms of "output" indicators–changes in income, health, et cetera–where possible to assure that we focus on the relative effectiveness of alternative programs and that we are able to evaluate and assess their impact on development objectives.". (9)

4. *DEFINING MEASURABLE NUTRITION OBJECTIVES*

The Mandate establishes improved nutrition as a major objective as well as a benchmark characterizing the poor. Measurable indicators of nutrition are divided into two general types:

1. Measures of nutrient intake or diet.
2. Measures of nutritional status.

If a project aims to improve nutrition in a selected target population, it would follow that its success or failure should be judged in terms of its nutritional impact. Food intake can be measured directly and the nutrient intake can be estimated from these food accounts. The difficulty of using only diet indicators is that they do not provide a complete basis for making comparisons between different sub-groups inside the poor, or for us in intercountry comparisons. Using only average caloric intake as a nutritional indicator is not sufficient, but it is an obvious beginning. Very few projects aimed at nutritional improvement have ever measured the actual impact on the caloric intake of the target group.

Two problems make the nutrient intake approach insufficient: first, there are sizeable differences in the nutritional needs of different healthy people, and secondly, the efficiency with which the body absorbs nutrients is significantly affected by different types of health conditions.

The MANDATE emphasizes the importance of disaggregating project measurements of nutrient intake as follows:

"Adequate diet requirements vary with age, sex, size, health status, occupation, and climate. Pregnant and lactating women generally require 300 to 500 additional calories daily. Sick persons, particularly children, require additonal calories especially when their illnesses inhibit absorption of what nutrients they do receive. Those engaged in active labor need hundreds of calories more as do those living in cold climates." (10)

BEST AVAILABLE COPY

Even for nutrient intake, adequate measurements would require that the target group affected must be divided on the basis of age, sex, climate, health status, etc.

More direct measures of nutritional status are available and in most cases they are easier to obtain than diet measurements. These measures are directed at the health and physical effects of malnutrition; they attempt to measure the results of nutritional intake rather than the intake itself. Three general types of measures are available:

1. Body measurements.
2. Measurements of nutrition-related illnesses.
3. Blood chemistry.

Body measurements most commonly used are height, weight, arm circumfrence and "fatty-fold" measurements. These measurements may be taken by non-medical personnel, and except for "fatty fold" and blood chemistry measurements they are relatively inexpensive to obtain. Body growth and size measurements are only reliable as indicators for project selection and evaluation purposes on children up to about 16 years of age. Beyond that age the measurements provide only sketchy indications of nutrition in earlier periods.

A wide variety of illnesses are related to nutrition. Measuring the incidence and gravity of these illnesses can be used as an indicator of nutritional status. Oedema, anemia, and goiter are among the most commonly used, but a wide variety of others may be as important in particular areas. For target populations, morbidity indicators are more difficult to use as project selection and evaluation criteria because they require clinical examinations and cannot be gathered by non-professional survey personnel.

Biochemical blood examinations may be useful, but are not really practical as project indicators. These measurements are more suited to in-depth clinical studies of nutrition.

In project terms the nutritional objective may be defined: to increase nutritional intake levels, to reduce the incidence of nutritionally related morbidity, and to improve the physical growth characteristics of the target population.

Mandate benchmarks in nutrition are in terms of calories only. Daily caloric intake of less than 2,160-2,670, depending on the country, places the average person in the target group. Table 3 from the MANDATE outlines these target group averages as they appear in the committee print.

The project cycle for a nutrition project should include direct measurement of some of the above mentioned nutrition indicators on the target population to be affected. The nutrition system is so interrelated with the income, employment, food production, education and health situation of the target group family that most successful mandate oriented projects will have important nutritional impacts. Estimating the nutrition impacts should be an element of most Mandate projects.

BEST AVAILABLE COPY

Table 3

Average Per Capita Daily Energy Requirements

Country	Calories	Country	Calories	Country	Calories
Uruguay	2,670	Mali	2,340	El Salvador	2,300
Turkey	2,520	Niger	2,340	Ghana	2,300
Egypt	2,500	Rwanda	2,340	Guinea	2,300
Cyprus	2,480	Ethiopia	2,330	Ecuador	2,290
Syrian Rep.	2,480	Malawi	2,330	Gyuana	2,280
Chile	2,450	Mozambique	2,330	Honduras	2,280
Afganistan	2,440	Pakistan	2,330	Madagascar	2,280
Trinidad/Tob.	2,430	Burundi	2,320	Dominican Rep.	2,260
Yemen Rep.	2,430	Cameroon	2,320	Philippines	2,260
Moroco	2,420	Colombia	2,320	Cent. Afr. Rep.	2,250
Bolivia	2,410	Ivory Coast	2,320	Costa Rica	2,250
Tunisia	2,400	Jordan	2,320	Haiti	2,250
Brazil	2,390	Kenya	2,320	Jamaica	2,250
Upper Volta	2,380	Mauritania	2,320	Nicaragua	2,250
Chad	2,370	Somalia	2,320	Khmer Rep.	2,320
Mauritius	2,370	Zambia	2,320	Thailand	2,320
Nigeria	2,370	Dahomey	2,310	Laos	2,220
Senegal	2,370	Liberia	2,310	Sri Lanka	2,220
Angola	2,360	Panama	2,310	Zaire	2,220
Korea, Rep.	2,360	Paraguay	2,310	India	2,210
Botswana	2,350	S. Leone	2,310	Guatemala	2,200
Peru	2,350	Tanzania	2,310	Nepal	2,190
Sudan	2,350	Togo	2,310	Vietnam, Rep.	2,170
Gabon	2,340	Bangladesh	2,300	Indonesia	2,160

SOURCE: Calculated from Annex Table: Population, food supply and demand for food in individual countries: ASSMENT OF THE WORLD FOOD SITUATION, PRESENT AND FUTURE, Item 3 of the Provisoral Agenda, United Nations, World Food Conference, November 1974, as it appeared in IMPLEMENTATION OF "NEW DIRECTIONS" IN DEVELOPMENT ASSISTANCE, AID, U.S. Government Printing Office, Washington 1975, p. 75.

5. DEFINING MEASURABLE HEALTH OBJECTIVES

Health objectives appear as important components in the Mandate. The Mandate lists four measurable indicators of health:

1. Life expectancy.
2. Infant mortality.
3. Birth rate.

BEST AVAILABLE COPY

Acceptable health status is obviously difficult to define. Projects with health objectives have found it very difficult to define in measurable terms what the final benefits are. For many years there was an effort by health project analysts to attempt to measure the outcomes of health projects in monetary terms. This approach, sometimes called the "costs of disease" approach has been largely unsatisfactory. It is encouraging that the Mandate avoids this approach and focuses the discussion directly on final indicators of health. Proponents of measuring health benefits in monetary terms argued that bad health reduces income in two ways: The productivity of the sick is lowered, and treating disease uses resources which could be producing added income.

The problem of defining measurable objectives for health projects (or mandate projects with important health impacts) is not easily solved by suggesting that the final intent is really to increase income, via improved health. There is no hint that this is what is meant in the MANDATE; it speaks of "good health" as an objective to be pursued as an end in

Even more than nutrition, impacts on objective indicators of health are likely to be long run impacts. A health project may expect to affect the life expectancy of the target population, but it may take considerable time to measure the effect even if it is successful. Lowered incidence of particular illnesses, reduced days seriously ill, and reduced infant mortality are indicators which have been used as measurable project objectives.

The last of the mandate health benchmarks is access to health services. Using access (a "means") rather than improved health (an "end") as an objective is a reflection of the measurement problem outlined above. This is the first mention in the Mandate of access objectives. The variety of health benefits, their differences in kind and degree, and the high cost of gathering data using scarce medical personnel have all worked together to make health among the most difficult areas for objective measurement. Beyond the benchmarks mentioned in the Mandate (life expectancy, infant mortality, and birth rates) and simple morbidity measures (incidence of particular diseases, number of days incapacitated, etc.) lies a set of extremely costly and complex measurements which have been used to evaluate health expenditures. Most of these benefit measurements are unsatisfying because they cannot be reduced to similar units, and even inside health projects few comparisons are possible.

Measurable objectives for health will probably be limited in most cases to reducing infant mortality, extending life expectancy, and reducing the incidence of specified diseases or the amount of time incapacitated.

Using increased access to health facilities as the final objective of a health project, without even attempting to measure its final impact on more 'final' or direct measures of health status, has many disadvantages. Most important, no reliable experience is accumulated on the relative merits of different approaches to improving health. If no measurements can be made associating project activities and final results, no serious evaluation can be made. Even though the above mentioned indicators of health status are much less complete than might be desired, they are substantial improvements over access objectives.

BEST AVAILABLE COPY

4. INCREASING RURAL PRODUCTION

Increasing production in rural areas is an objective which requires conceptual clarification before it can be measured and used as a project design and evaluation criterion. Production increases in and of themselves are not necessarily beneficial to the target group. To illustrate this point let us take the example of a small farmer who increases his production of corn by using increased fertilizer which costs him more than the added production is worth. Additionally there are many well known examples where farmers increase the production of a particular crop, resulting in a sharp decrease in the market price and an actual reduction in the net income of the farmer. Added production of a particular crop is a poor indicator of added "production," since it may be substituting for other crops or livestock products in the farm.

With all of these conceptual problems it might appear that increased production is a useless indicator of welfare for target group farmers, and that net income, and not its proxy "production" is the only adequate measure. As long as the benefit is narrowly defined as a direct farm level benefit that is largely true, however, if the project takes a broader societal view of benefits, there is a residual of benefits to the target group which is not captured by the direct farm level net income benefit which may be addressed by properly formulated measures of increased rural production.

The private profit or net income interest of small farmers is not always consistent with society's interest in maximum total welfare, or even maximum total welfare of the poor majority. Economic theory would tell us that if product and factor markets operate freely, and if there are no significant factor immobilities, the private profit or net income drive of each producer will result in the greatest possible total welfare. Unfortunately these optimal conditions are not characteristic of most underdeveloped rural settings, land and credit are seldom priced and distributed in freely operating markets, and a wide variety of communication and exaggerated transportation and other barriers prevent equitable access to the scarce markets which do exist. Public interventions of all sorts distort the system to such a degree that it is unclear to what extent the individual net income measure is a valid proxy for overall welfare.

An example of the possible divergence between the private profits interest of the farmer and societal interest in the welfare of the total poor majority may be seen easily in the case of expanded production of a particular commodity by a wide variety of small farmers. Let us assume in the case of Costa Rica outlined in Table 1, where the target group farmers were producing corn, an expansion of corn production might result in a decreased corn price, and in reduced net incomes to small farmers if measured in simple direct farm level crop accounts. However, the reduced corn price means that both urban AND rural poor who consume corn will be able to buy more corn for less money, resulting in increased welfare. Thus, if the program is judged based only on direct farm level net income impact on small farmers, the important overall welfare benefits would be ignored. It is not infeasible that these indirect benefits may outweigh farm level losses. The example given is only one of a wide

BEST AVAILABLE COPY

variety of difficulties which make farm level net income a less than fully adequate measure of overall target group welfare. This section is not the place, nor DSP an adequate setting for the examinatio of these difficulties: their conceptual complexity and the methodological difficulty of measurement in this area preclude any serious treatment. It is important to realize only that increased production is only a useful measure when it is used at the economy level and not for a particular farm or group of farms. If production increases on a particular farm or group of farms, we cannot say that their net income has increased until net income itself is measured, and therefore increased production is a meaningless farm level indicator; we might just as well measure the real thing, net income, directly.

To measure the societal benefits which are not captured by the farm level net income measures, added refinements must be made to the concept of "rural production."' These refinements may be divided into two categories, first netting gross production to arrive at estimates of "value added," and adjusting observed market prices to reflect market distortions (shadow pricing). Neither of these techniques will be dealt with in depth in DSP because of their conceptual and methodological complexity. It is important for project personnel to realize that only when these adjustments are made can measures of increased production be obtained which are useful as project selection and evaluation guides. Of all the Mandate objectives, increased production is the most difficult to measure in a way which is useful for AID programming. For most agricultural projects the size will not justify the necessary expenditure on analysis and measurement required to measure indirect benefits, and it may be tempting to select production increases as the project target instead of income because it is easier to measure or is thought to be a valid short-term or intermediate step. Speaking of rural production, the MANDATE addresses this issue as follows:

> "In some cases it may be necessary to restrict targets to certain short-term achievements... rather than final outputs like improved income. However, the program should be designed with a view toward AID's overarching goals of moving the poor beyond the poverty level and links with those goals should be explained." (11)

5. POPULATION

The population objective of AID as explained in the MANDATE is an intermediate objective presented in the committee print as a means for increasing per capita living standards. If we were to treat it like increased production as a means to increased income, health, and nutrition, we would wish to measure the final desired objectives and not simply achievements on the intermediate means. Population programs, like production increases, can be easily measured in terms of a single indicator, birth rate. The difficulties all begin when we attempt to link the population variable to the final objectives of increased income, health and nutrition. While considerable reserch has been undertaken on what factors influence fertility, little is known about the reverse, what effect altered fertility has on net income, health, nutrition and education. The mandate specifically enjoins production projects

BEST AVAILABLE COPY

to "be designed with a view toward AID's overarching goals of moving the poor beyond the poverty level" and further requires that "links with those goals should be explained." (1) In the case of population no such linkage is asked for; the MANDATE simply states.

> "Increasing the size of the pie by providing more food and health services is essential to improving per capita living standards; substantial progress will only be possible if population growth abates." (12)

Taken at face value, the MANDATE discussion of population bypasses the issue of linkage of population projects to the three "overarching goals" (income, health, nutrition) and allows project designers to look at the intermediate objective of reduced fertility as if it were a final objective.

6. EDUCATION

The MANDATE outlines six objectives, three of which are presented as final objectives or desired "ends" (income, health, nutrition), and three which are presented as intermediate objectives or "means" to achieve the final desired improvements (increased rural production, population, education). It is important to note that while education might have been argued as a good independent of its impact on income, health, or nutrition, the MANDATE makes it clear that this is not the case for AID programs.

> "In a world of plenty, "education" may connote literacy and wide learning, truth for its own sake as well as a means to progress. In a world of want education must unfortunately of necessity be something far more restrictive and practical–as a means to improving living standards rather than an end in itself." (13)

Unlike population, considerable serious research has been undertaken to attempt to understand the impact of increased education on the final objectives of income, health, and nutrition. While the research has been extensive, the results are far from clear, and are usually unable to clarify project design and evaluation issues. Given the current state of the art, it will probably be some time before design and evaluation can be guided by measurement of project impacts on final AID objectives. Recognizing this difficulty the MANDATE defines "minimum practical education" as the proxy for the education objective.

7. PARTICIPATION

The congressional committee documents and AID's replies emphasize another objective of assistance under the rubric of "participation." The MANDATE explains the sense in which this objective is taken in the following terms:

BEST AVAILABLE COPY

WHAT IS THE MEANING OF "PARTICIPATION"?

Development progress for the poor will require time-consuming systemic change. Programs most likely to succeed, and which receive highest priority emphasis under the congressional mandate and AID policy, are those involving the active and effective participation of the poor in all facets of the development process. This is clearly not a simple proposition to implement, especially as it has important political implications. We shall refer to participation frequently in the following section of the report, but to briefly describe the participation approach to our staff we have used the following:

- *Economic benefits are widely and significantly shared by the poor* with the objective of narrowing the relative income gap between rich and poor, for example, the co-op which benefits small farmers.
- *Decisions concerning the activities to be carried out are made, preferably, by those benefited (for example, the poor)*, or if not, at least with effective consultation and substantial acceptance by those benefited.
- *The activity in which they participate is, ideally, a learning experience for benefited persons*, which increases their technical skills and/or their capacity to organize for common purposes and for greater access to the benefits of development.
- *The poor make a significant contribution* in effort and resources to the activities from which they benefit, for example, through personal savings, or serving as members of local planning or project implementation committees, or through actual project implementation.
- *The participation and contribution of women* should be explicitly taken into account under the above-mentioned considerations, for example, any of the above or other examples when the participants are women.

Only the first dimension, that of participation of the poor in the BENEFITS, outlined in the MANDATE will be dealt with directly in this manual and segment of the course. The other four dimensions, which relate to participation in the PROCESS will be addressed in other segments of the course.

FOOTNOTES TO PART A: DEFINING MEASURABLE "MANDATE" OBJECTIVES

(4) Agency for International Development, *The Congressional Mandate: Aiding the Poor Majority*, 94th Congress Committee Print, Washington, D.C., 1975, p. 65.

(5) Agency for International Development, *Implementation of "New Directions" in Development Assistance*, Report to the Committee on International Relations, 94th Congress Committee Print, Washington, D.C., 1975, p. 6.

(6) Agency for International Development, *The Congressional Mandate: Aiding the Poor Majority*, 94th Congress Committee Print, Washington, D.C., 1975, p. 68.

(7) Denis Goulet, "An Ethical Model for the Study of Values," *Harvard Educational Review*, May 1971, 41:206-207, as cited in Development Alternatives, Inc., *Strategies for Small Farmer Development*, Washington, D.C., 1975.

(8) Agency for International Development, *The Congressional Mandate: Aiding the Poor Majority*, 94th Congress Committee Print, Washington, D.C., 1975, p. 65.

(9) *Ibid.*, p. 64.

(10) *Ibid.*, p. 66.

(11) *Ibid.*, p. 71.

(12) *Ibid.*, p. 73.

(13) *Ibid.*

(14) *Ibid.*, p. 74.

BEST AVAILABLE COPY

GETTING SPECIFIC ABOUT PARTICIPATION: ANALYSIS FOR PROJECT DESIGN, IMPLEMENTATION & EVALUATION

(Editor's Note: This reprint by Norman T. Uphoff, John M. Cohen and Arthur A. Goldsmith was published as Appendix 2 of the Cornell Rural Development Committee's State-of-the-Art Paper on Feasibility and Application in Rural Development Participation as part of the work under its cooperative agreement with AID's Office of Rural Development in the Development Support Bureau.)

Rubric Rather than Concept

Because participation is essentially a descriptive term, comprising numerous different activities and situations, there is much room for confusion about its causes and effects, its amounts and distribution. It is necessary to specify what is meant by participation in any particular situation, if we are to speak usefully about it in any specific rural development effort. This we have elaborated elsewhere and the following is a summary of that exercise.[1]

Our framework is based on the crucial assumption that "participation" is not a thing that exists in certain quantities, and that can be measured like a dam's capacity or a farm's agricultural production.[2] While some studies treat participation

---

[1] John M. Cohen and Norman T. Uphoff, Rural Development Participation: Concepts and Measures for Project Design, Implementation and Evaluation (Ithaca: Cornell University, Rural Development Committee, 1977). In the second half of this monograph, the framework is applied to the Chilalo Agricultural Development Unit (CADU) project in Ethiopia assisted by the Swedish International Development Agency, pp. 176-315.

[2] Participation is much like such blanket concepts as power and energy. We use them to describe and particularly to make comparisons, but any analysis is best done in terms of specific kinds: economic power, physical power, social power, moral power, and so forth, or electromagnetic energy, thermal energy, nuclear energy, gravitational energy, and so forth. With regard to energy and power, there is clearly some convertibility among the forms, and even some established units of measurement for comparing forms of energy. When it comes to participation, as suggested below, there is observationally some connection among the different kinds, but no accepted measures or demonstrations of conversion. For each kind there are possibilities of measurement, especially with respect to changes over time, so

BEST AVAILABLE COPY

as a clearly defined concept capable of measurement, we have chosen to treat it as a rubric under which a number of clearly definable elements can be assembled.[3] Although these elements can be related under a framework called "rural development participation," it would be quite misleading to try to define the framework concretely to allow summary measurement by specific empirical indicators. In this sense, participation is an overarching concept best approached by looking at its more specific components.

At the heart of our elaboration of rural development participation is the distinction between dimensions and contexts of participation. Briefly, dimensions of participation concern the kind of participation taking place, the sets of individuals involved in the participatory process, and the specific characteristics of that process. The context of participation focuses on the relationship between the rural development project's characteristics and the patterns of actual participation that emerge. In addition, the context concerns the project's task

---

we are not without some basis for empirical examination. But the state of art is fairly rudimentary. The first step toward establishing more reliable and useful knowledge is specifying more concretely what are the elements or units to be analyzed.

[3] In this sense we treat it in theoretical terms rather than observational terms. See: Abraham Kaplan, The Conduct of Inquiry (San Francisco: Chandler, 1964), p. 54. Theoretical concepts cannot be defined in the strict sense because of the openness of their meanings. We think it often more useful to speak of participation in terms of an adjective, describing whether certain kinds of development activity are participatory or how participatory they are. This does not suggest quite as reified or homogeneous a thing as when one speaks of participation as a noun.

BEST AVAILABLE COPY

environment, specifically the historical, ecological, and societal factors that strongly affect emerging patterns of participation. Figure 1 provides an overview of the dimensions and contexts of rural development participation that will now be elaborated.

Dimensions of Rural Development Participation

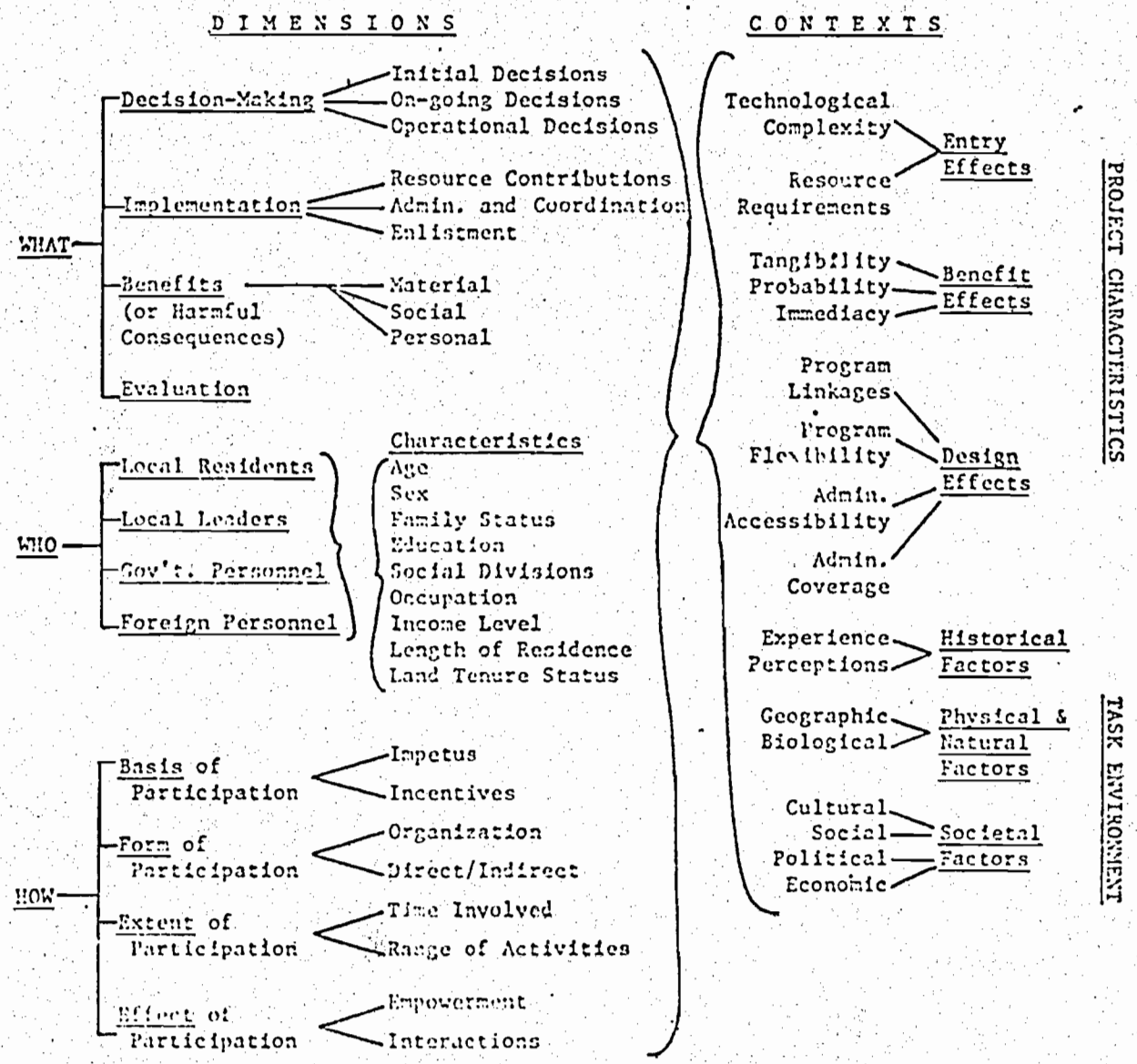
The framework we have devised delineates three dimensions of participation. These answer the questions: What kinds of participation can take place? Who participates in them? And how does the process of participation take place?[4]

We discuss each of these below, to show how it is possible to be specific about participation. The first dimension--what--is the basic one to distinguish, since participation in decision-making can be different from participation in implementation, for example. For project designers and evaluators, the second dimension--who--may be most salient, since this deals with

[4]The only study we have found that deals fairly comprehensively and analytically with rural development participation takes a somewhat similar view. Robert Chambers, Managing Rural Development: Ideas and Experience From East Africa (Uppsala: Scandinavian Institute of African Studies, 1974), pp. 84-113. It suggests three ways of analyzing participation: who, what institutions or channels, and what objectives or functions? The first is the same as our second dimension noted above. The other two appear somewhat more descriptive, nor encompassing as much variation in analytical terms as the other two dimensions we identify, but they will be taken up in the course of the discussion since they raise important questions about participation's role and effectiveness in rural development.

BEST AVAILABLE COPY

Figure 1: Basic Framework for Describing and Analyzing Rural Development Participation



BEST AVAILABLE COPY

intended beneficiaries or target groups,[5] whose participation is to be provided for. The third--how-- adds a qualitative dimension to consideration of participation, one which may often be more difficult to ascertain, but which remains important for purposes of diagnosis or evaluation. It may well be crucial whether the participation by a certain group of persons is voluntary or coerced, continuous or intermittent, effective or ineffective. What kinds and conditions of participation are relevant in a given situation will depend, of course, on certain other factors, such as the project's task and environment.

i. What Kinds of Participation?

The kinds of participation that warrant major concern are: (1) participation in decision-making; (2) participation in implementation; (3) participation in benefits; and (4) participation in evaluation.[6] We find the first three kinds of participation

---

[5]One may well take exception to the term "target group" or "intended beneficiaries" in the context of supporting participatory development, since these are "top-down" concepts. They involve quite possibly no participation by the people so identified in determining whether or not they are or should be regarded as a group to be served in specified ways by a project or policy. Unfortunately, no less paternalistic term has come into use and we have no more precise alternative to propose. We would welcome some improvement in the terminology.

[6]The 1968 conference at M.I.T. on implementation of Title IX came up with the first three. David Hapgood, Role of Popular Participation in Development (Cambridge: MIT Press, 1969), passim. After a close look at the literature we confirm these and add the fourth. Participation in evaluation occurs even less frequently than the others, but it deserves increased attention if development efforts are to be progressively improved, and its underscoring is consistent with efforts being made in the development community to introduce systematic evaluation activities.

BEST AVAILABLE COPY

are now reasonably well defined by development assistance agencies,[7] and see no grounds for objecting to the fourth. Indeed, while evaluative participation occurs even less frequently than the others, it deserves increased attention if development efforts are to be progressively improved. Our underscoring of participation in evaluation is consistent with efforts being made in the development community to introduce systematic evaluation activities.

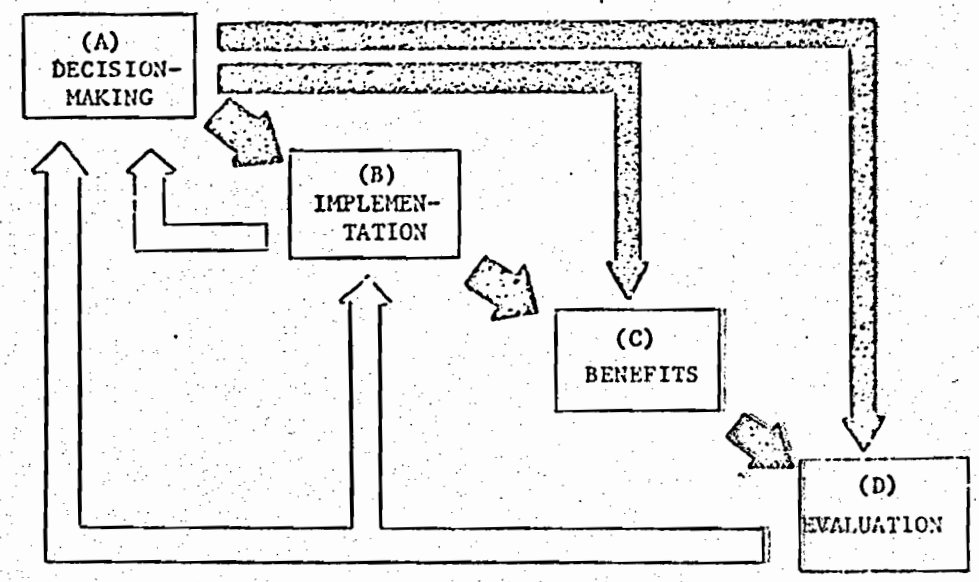
Together, these four kinds of participation constitute something of a cycle for rural development activity. In practice, there is seldom a consistent or complete cycle of interactions, such as shown in Figure 2, and participation in these different activities is often quite limited or unequal. Yet they constitute a tangible set of things to focus attention on and represent the major ways in which participation in rural development can be assisted and assessed.

---

[7] For example, in a report prepared for the FAO's Rural Organizations Action Program in 1977, the first three kinds of participation were specified as elements of popular participation: (1) decision-making regarding development; (2) contribution to development efforts; and (3) sharing in development benefits. Bernard Van Heck, The Involvement of the Poor in Development Through Rural Organization: Framework for Research-cum-Action (Rome: Food and Agriculture Organization of the United Nations, June 1977), p. 25. For similar language in USAID documents see: Implementation of "New Directions" in Development Assistance (Report prepared by USAID for Committee on International Relations on Implementation of the Foreign Assistance Act of 1973, 94th Congress, 1st Session, July 22, 1975), pp. 7-8: (1) Decisions concerning the activities to be carried out are made, preferably, by those benefited (for example, the poor), and if not, at least with effective consultation and substantial acceptance by those benefited; (2) the poor make a significant contribution in effort and resources to the activities from which they benefit, for example, through personal savings, or serving as members of local planning or project implementation committees, or through actual project implementation; and (3) economic benefits are widely and significantly shared by the poor with the objectives of narrowing the relative income gap between rich and poor, for example, the coop which benefits small farmers.

BEST AVAILABLE COPY

Figure 2: Four Kinds of Participation



Participation in decision-making is what political scientists most often refer to when they think of participation, whereas administrators are likely to focus on implementation participation. On the other hand, economists have in the past stressed participation in benefits as the most important thing. No one has been much concerned about participation in evaluation, just as evaluation itself has been neglected. Participation in decision-making and implementation pertains to the "inputs" of rural development projects, providing authority, information, labor and other resources, while people's involvement in benefits and evaluation of development activity concerns "outputs." (When considering "benefits," however, it is important to recognize harmful consequences can flow from participatory activity, an issue returned to shortly.)

BEST AVAILABLE COPY

Decision Making: More specifically, this kind of participation centers on the generation of ideas, formulation of options, evaluation of options, and making of choices about them, as well as the formulation of strategies for putting selected options into effect. For this reason we discern three types of decisions; (1) initial decisions: (2) ongoing decisions; and (3) operational decisions.

Initial decisions begin with the identification of local needs and how they will be met through a particular project. Often this is the most crucial stage of the project, when some of the most critical decisions are made. In most cases these very early decisions remove a large number of options from the decisions that follow. For this reason, initial decisions need to be carefully distinguished and focused upon.[8]

Probably the most fundamental decisions concern what is defined as a "problem" and which "problems" need to be addressed most urgently. These initial decisions, generally described as "project identification," may or may not involve local participation. Ideally, local people and their leaders should be involved at this early stage, where they can provide vital information on the local area and prevent misunderstandings on both sides about the nature of local problems and the strategies for their resolution. Among the initial decisions in which local people can be involved are whether the project should start, where it should be

[8] In a comparative study of 36 rural development projects intended to benefit small farmers, Development Alternatives, Inc. undertook to analyze who participates (farmers, project planners, or both) in "idea generation" and "idea refinement" (seen as having two stages--communication and experimentation) before there is "decision-making" about a project. This latter phase is not analyzed in any detail with regard to participation. DAI, Strategies for Small Farmer Development: An Empirical Study of Rural Development Projects (Washington, D.C., DAI, May, 1975), I, pp. 95-107.

BEST AVAILABLE COPY

located, the ways it should be financed and staffed, the paths by which local people will participate in the project, and the contributions they are expected to make.[9]

---

[9] The range of possible decisions can be elaborated using the basic conceptual framework set forth in Figure 1. (1) Where Decisions: Decisions on the location of specific project activities or sub-project units can be very important, and may be open to local participation. These decisions are most crucial in projects involving construction such as farm-to-market roads, schools, health clinics, irrigation canals or cattle dips. But they may affect the location of services where establishing certain centers is involved. (2) How Decisions: Decisions on the ways and means of project implementation are necessary in all projects, though local people and even local leaders or lower-level government personnel may not be included in them. Some of the kinds of decisions that may be analyzed to see who participated in making them are: (a) strategy--the specific approaches to be taken to meet the project's goals and to help resolve the problem that led to the formulation of the project; (b) financing--whether the project shall be financed in whole or in part with local resources, and if so, whether in cash, in labor, or in material contributions; (c) staffing--whether personnel shall be drawn in part from the community or shall be entirely from "outside"; what will be the terms of service; to whom shall staff report; (d) standards--what quality or coverage shall be sought, such as all-weather or dry weather roads, universal enrollment or selective enrollment in programs; and (e) enforcement--what sanctions shall be invoiced to ensure contributions to or participation in the project, who shall be responsible for enforcement, and so on. (3) Who Decisions: Decisions on who will participate in the other phases of the project deserve consideration because they are central to its operations and effects. One should be interested not only in who decides the where and how design features of a project but also in who decides which persons or groups will be involved in: (a) implementation--who is required or expected to perform what tasks in carrying out the project; (b) benefits--who is eligible for the benefits which the project is expected to confer; and (c) evaluation--who, if anybody, is expected or authorized to make appraisals of project effects and success.

BEST AVAILABLE COPY

Ongoing decisions occur after the initial decisions have been made, and some evidence suggests participation in them may be more critical to project success than participation in initial design decisions.[10] It is possible for local people not to participate in initial decisions, but to be asked to join in ongoing decisions once the project has arrived. Here opportunities exist for searching out new needs and priorities the project might respond to. Moreover, local people might be given an opportunity at this stage to participate in decisions to continue or terminate particular aspects of the project, if not the project itself. Finally, local people can be involved in decisions on the structure and content of activities already in progress, as well as new decisions about the project's goals and the means to achieve those goals.[11]

---

[10] See Section B of the Executive Summary of DAI's report, Strategies for Small Farmer Development. Project success was measured in terms of increases in farmer income and agricultural knowledge as well as in self-help capacity and probability of project benefits becoming self-sustaining. Local action taken by farmers to complement outside management and resources accounted for half the variation in overall success rankings, and farmer involvement in decision-making in the implementation phase was one of the two factors found most significant in promoting overall project success.

[11] These include: (1) a continuing search for other needs and priorities that the project might respond to; (2) an evaluation of whether a project should continue or be terminated; and (3) a continuing elaboration of the structure and content of established activities, such as: (a) where decisions--there may be decisions on relocation of service centers or on the location of new facilities (roads, schools, canals, and so on); (b) how decisions--there will certainly be decisions on implementation with respect to financing, staffing, standards and enforcement; and (c) who decisions--there will also have to be continuing decisions--or review of decisions--on who will be required or expected to contribute what and how to implement projects, as well as on who will be eligible for benefits.

BEST AVAILABLE COPY

Operational decisions relate to specific local organizations established by the project, or linked to the project to involve local people in delivering services. Here the focus is on voluntary associations, cooperatives, traditional associations, women's clubs, and other organizations involved in the substantive activities of the project. Our framework suggests the need to focus on membership composition, meeting procedures, leadership selection, and influence of such organizations. More specifically, we would address such questions as whether membership is universal, what groups are represented or underrepresented, what continuity or turnover in membership exists, and what obligations are placed on members. Moreover, our concern with operation decisions would lead us to focus on the procedures for calling meetings, the way meetings are publicized, whether people attend them regularly, the amount of influence or authority held by those who attend, and the degree of independence under which the meetings are held. We would also look closely at the leaders of such groups, to explore their social and economic background, the process by which they are selected, the terms of their office, the ways in which they are rewarded, and processes by which they can be replaced or removed. Finally, and most importantly, it is essential to give attention to the influence of such organizations on projects, their ability to affect the activities of project personnel, the amount of effort they undertake to influence the project, and the degree of success they have in reaching their selected goals.

Implementation: Rural people can participate in the implementation aspects of a project in three principle ways: (1) resource contributions; (2) administration and coordination efforts; and (3) program enlistment activities.

BEST AVAILABLE COPY

Resource contributions can take a variety of forms, such as the provision of labor, cash, material goods, and information. All such inputs are vital to projects seeking to incorporate local resources in the development enterprise. Through such participation local people may lend their labor to dig wells, give land for a school, contribute timber for construction of the health station, donate tools for work on a local road, donate money for financing community grain storage bins, or provide crucial information on crop yields, tenure arrangements, pest problems, sources of nutrition and so on.

The relation among the three dimensions of participation is illustrated clearly with regard to resource contributions. It is particularly important to know who is contributing, how their contributions are made, whether they are voluntary, remunerated or coerced, the degree to which they are provided on an individual or collective basis, and whether they occur intermittently or continually. These are particularly important questions, since resource contributions can often be unequal and exploitative.[12]

Participation in administration and coordination of projects is a second way rural people can be involved in implementation. Here they can participate either as locally hired

[12] We have found that one way to insure careful analysis of these interwoven participatory dimensions is to establish tables representing a particular kind of participation with, for example, the "who" categories on one axis and the "how" factors on the other, with the table representing a particular kind of participation. We have presented some 30 illustrative tables for one project in Cohen and Uphoff, Rural Development Participation, p. 204.

BEST AVAILABLE COPY

employees, or as members of various project advisory or decision-making boards. They can also be members of voluntary associations that coordinate activities with the project. Having local people help administer and coordinate a project not only increases their self-reliance, but also gives officials valuable inside information on local problems and constraints affecting the project.

Finally, perhaps the most common form of participation in implementation is the enlistment in programs. Distinguishing between enlistment and participation in benefits is essential, because enlistment does not necessarily assure benefits. Such a distinction is justified also by the fact that harmful consequences may result for rural people who have enlisted in project programs.[13]

Benefits: Enlistment in a project can lead to at least three kinds of possible benefits: (1) material; (2) social; and (3) personal. While participation in benefits is usually a relatively passive phenomenon, we would emphasize participation in benefits is a desirable goal, which can often be realized through participation in decision-making, implementation, and evaluation. Perhaps the only danger in focusing on this kind of participation is that it can sometimes be quite high and lead observers to overlook the fact that participation in other important aspects of the project, such as decision making, have either not occurred or been quite limited.

---

[13] For example in a project innoculating a herd against diseases like bruncellosis, it is the reduction of morbidity in one's cattle herd that is the benefit, and not the innoculations themselves. That is, some services yield only uncertain benefits. Planting HYV seed does not assume a better harvest, for seeds may not be suited to the particular environment or they may encounter weather that nullifies their yield potential.

BEST AVAILABLE COPY

Material benefits are basically private goods.[14] They can perhaps be summarized as an increase in consumption, income, or assets. Increases in consumption can result from higher food grain yields, and income benefits can come from selling surplus production. Increased assets can be seen in the acquisition of land, livestock, implements, improved farm dwellings, savings, and so on. As with all other aspects of participation discussed in this paper, breaking down aggregate data, by analyzing the participants and the participation process, is essential.[15]

---

[14] For a good discussion of the difference between "private" and "public" goods, see Mancur Olson, The Logic of Collective Action (Cambridge: Harvard University Press, 1965). While it is a valid and often revealing distinction, in examining projects and the relationship between public and private goods appears less significant than how either kind is distributed.

[15] Here particular attention should be given to quantity, distribution and quality measures, such as: (a) higher average or total income for a particular group, or similar data on consumption, possibly just for food but preferably measures for all consumption; (b) more equitable distribution of income or consumption within a community or population, particularly a narrowing of the "gap" between highest and lowest incomes within a community; and (c) more security of income and consumption, i.e., less risk that these will be reduced or interrupted. With regard to assets capable of producing an income flow over time attention should be given to: (a) greater per capita or total stock of assets for a particular group; different kinds of assets may be assessed separately or summed in terms of monetary values; (b) more equitable distribution of assets within a community or population; and (c) more security of assets, e.g., legal recognition and enforcement of tenants' previously de facto right to cultivate their land.

BEST AVAILABLE COPY

Social benefits are basically public goods, usually services or amenities such as schools, health clinics, water systems, improved housing, and better roads.[16] Increasingly, as rural development projects are designed to be more "integrated," and as efforts are increased to improve the "quality of life" for poorer sections of the population, participation in such benefits will have to be assessed. Particular attention should be given to the amount, distribution, and quality of these services and amenities.[17]

---

[16] We would consider education separately in recognition of the special place it occupies in increasing productivity, welfare and power for the poor majority. See Norman T. Uphoff and Warren F. Ilchman, The Political Economy of Development (Berkeley: University of California Press, 1972), p. 329 ff.

[17] Examples of the kind of objective measures needed are: (1) greater availability of certain services to a community or a particular group, measured in terms of total amount or as some ratio to the specified population. In measuring the effects of a water supply project in different neighborhoods, one could compare number of water taps per 1,000 population or figuring a certain maximum usage rate (e.g., one tap can serve up to 500 persons per day), one could estimate how much of the specified population was being served; (2) wider access to certain services for a given community or group, accounted for in terms of reduced cost and/or time to utilize the services. Availability and access are certainly related, but efforts to measure access would look at more water taps as distinguished from a larger reservoir capacity in a water supply system; reduced average distance to health clinics; more bus stops or more frequent bus service; and (3) improved quality of certain services, such as shift from curative to preventive medicine (lowering morbidity rates), purer water (reducing incidence of water-borne diseases), or bus service more reliable and punctual (fewer breakdowns and more predictable schedules). Another example could be inspection of markets to ensure accurate scales and measures in addition to providing more stall facilities. In the case of amenities, quality may be a considerably less important criterion than availability and access in service-starved rural areas. But this judgment in the choice of measures depends, as in all such choices, on the nature of the project and on its context.

BEST AVAILABLE COPY

Personal benefits are usually greatly desired, though often not attained, on an individual basis, coming rather to members of groups or sectors as these acquire more social and political power through the operation of a project. We term them "personal" benefits to distinguish them from "material" and "social" benefits, but this does not imply that they are necessarily "individual" in their causes or effects. Among several possible project-generated benefits of this sort, three appear particularly important: self-esteem, political power, and sense of efficacy.[18]

The focus on benefits should not cause us to overlook the many harmful consequences that can follow from participation in a project. These might include seeds that do not germinate, cross-bred dairy cattle that do not survive, or more extreme results, such as the erosion of the local culture, or the eviction of tenant farmers. While harmful consequences are obvious to careful observers, they are often not studied as benefits are.

---

[18] Self-esteem comes from improvement of an individual's status according to whatever criteria of "worthiness" prevail in the community or among his or her associates. It probably has to be measured in each case by locally-determined standards, such as land ownership, literacy, possession of belongings like a wristwatch or bicycle, freedom from debt, and so on. Acquisition of political power is even more difficult to assess. It affects how able a person is to avail himself or herself of the other benefits discussed here. Essentially it relates to the enhancement of a person's ability to influence authoritative decisions, whether through the electoral process, official administrative channels, or private negotiations with officials. Sense of efficacy, an increase in the individual's recognition that he or she can play a role in the development process, can be a result of almost any kind of rural development project. If a project can enhance this orientation within various rural groups, its contribution to developing will be multiplied beyond its direct benefits.

BEST AVAILABLE COPY

As suggested earlier, we concluded that to analyze participation in benefits from a particular project, consideration should be made of whether harmful consequences have also occurred. If so, any differential rates of participation will be significant data to determine. The main concern will be who is participating in adverse outcomes of the project. Once this is known, one will want to try to establish why, in case remedies can be found for the outcomes and can be built into a re-designed project. For this reason, broad participation in the evaluation of projects is very important.

Evaluation: Because little is written--or actually accomplished--on participation in evaluation, it is difficult to conceptualize how this kind of participation might best be analyzed and measured.[19] Still, rural people can participate in project evaluation through three major activities.

Direct or indirect participation can occur with project-centered evaluation. If there is any formal review process, one would want to know who participates in it, how continuously, with what power to get action on suggestions, and so on. People

---

[19] Two particular methodological forces exist. Donors and local people may have very different perceptions about what evaluation means or what is being evaluated. Moreover, they may well use different criteria. In addition, evaluation literature confirms that negative evaluations are generally more spontaneously offered because people suffering a loss are more likely to complain, other things being equal, than those gaining something are to express approval. On the other hand, when evaluations are solicited, positive evaluations may misleadingly predominate because of a common disinclination to express disapproval under such circumstances (often reflecting differentials in power between the person asking an opinion and the person being asked). Such considerations should be kept in mind.

BEST AVAILABLE COPY

may be consulted informally only, in which case we would want to know answers to the same kinds of questions. Most probably, however, unless evaluation is specifically provided for in the project design, local people or local leaders will not participate in evaluating the project. Government personnel may themselves join in annual budget reviews that fulfill a certain kind of evaluative function, but local-level officials are generally not involved.

More likely is that whatever local involvement in project evaluation there is will occur through political activities of one sort or another. Where there are elected officials, such as members of parliament, local residents and village leaders can voice complaints and suggestions through their ballots. Participation in elections at local, regional or national levels can possibly enable local evaluations to be fed into policy processes, though such inputs are likely to be gross, reflecting simply dissatisfaction or satisfaction with what the project has accomplished.

Even where such "democratic" political processes are not available, local people and local leaders can engage in lobbying--possibly through some organization like a cooperative or peasant league--to communicate their views to the project or the government. Alternatively, but not necessarily more effectively, people may demonstrate or protest to force officials to attend to local grievances. Indeed, some projects have stimulated sustained conflict and violence. Such activities should be studied to determine if they provide indirect evaluation of the project, or reflect other issues in the task environment. If the former, the content of the protest as well as the characteristics of those making it should be the subject of inquiry.

Less direct would be participation in evaluative activities that aim at influencing public opinion, to build support to

BEST AVAILABLE COPY

continue or possibly modify a project. Usually such efforts seek to use the media, for example through a "letter to the editor," to promote a favorable or unfavorable opinion of the project or to suggest some improvement. This is a diffuse approach, but might be one possible form of participation in evaluation and better than no such participation at all.

ii. Who Participates?

The participation of most concern to development agencies and governments these days is that of the "rural poor" or the "poor majority." But if the poor are considered an aggregated mass, assessing their participation is difficult, for they are a large and heterogeneous group, differing in occupation, location, land tenure status, sex, caste, religion, or tribe. To talk about "the participation of the rural poor" is to compound one complex and ambigious term with another, even more complicated and amorphous. If we want to deal usefully with the problems of the rural poor, we need to begin making some analytical distinctions among them.[20]

[20] We will not here delve into the difficult question of the outer parameters of the "rural poor." We appreciate that the distinction between "rural" and "urban" is generally hazy, as there is much movement back and forth between what are called, often arbitrarily, rural and urban areas. The measurement of "poverty" has become confused with conflicting definitions of the term and various poverty lines being drawn, based on alternative minimum consumption levels, nutritional requirements, imputations of subsistence income, etc. Since we are concerned with identifying who participates in what activities, we can leave for others to decide where some boundary is to be drawn. The measures derived from our framework would give a descriptive continuum and show concentrations and absences of participation, to be assessed according to the objectives set for the project.

BEST AVAILABLE COPY

Our framework begins with a more differentiated scheme than others have suggested.[21] The merit of scheme, we think is that it can be used to analyze the entire rural community, as well as important sub-groups, such as the rural poor. We would distinguish four general types of participants whose characteristics warranted specific attention. Depending on the setting and the goals of the project, certain characteristics of participants would be more significant than others. We would distinguish: (1) local residents; (2) local leaders; (3) government personnel; and (4) foreign personnel. Each of these sets of persons can be further subdivided, according to background characteristics that are essential to the analysis of individual participation.

The first two sets of people are those who have local roots; the last two are, to varying degrees, outsiders. Local residents is a residual category. It is large and heterogeneous, made up of self-sufficient farmers, tenants or landowners, farm laborers, herdsmen, craftsmen, and so on. This group of people is usually the target of rual development projects.

Local leaders have a long run commitment to the area in which they work. Usually they are local elites, such as landowners, major merchants, and professionals. The definition of

[21] For example, Chambers identifies two groups whose participation should be of concern: (1) local people; and (2) government staff who are liable to transfer elsewhere. Participation by both sets of persons he calls joint participation. Managing Rural Development, p. 85. Alvaro Chaparro develops a very specific set of classifications for the FAO: (1) small owner-farmers; (2) tenants and sharecroppers; (3) landless agricultural laborers; (4) rural artisans (some categories, mostly part-time farmers or fishermen); (5) hunters-gatherers, fishermen or shifting cultivators (may be self-provisioning tribal people); (6) forestry workers; (7) nomads, semi-nomads or refugees (non-settled people); (8) females and children (as special categories of rural poor). "Participation of Rural Workers' Organizations in the Development of Agriculture and Rural Life in Latin America" (Paper Prepared for Food and Agricultural Organization of the United Nations, Rome, ESM: TU/LA/76/2, August, 1976).

BEST AVAILABLE COPY

local leader will vary from area to area, but basically there are three types: (1) informal leaders, such as clan chiefs, religious figures, influential professionals, and local notables; (2) associational heads, elected or appointed to a formal organization, such as a cooperative league president, the head of a voluntary association, or the leader of a local trade union; or (3) local office holders, such as headmen, elders, mayors, or tax collectors. Sometimes the nature of officeholding, given requirements to uphold local interests, makes it difficult to distinguish these people from government personnel.

Government personnel are assigned to an area for a certain period of time. Even if from the local area, their career rests with the bureaucracy at the center and is not usually determined by what happens in the local area. Government personnel are typically subject to transfer, and often have higher education and social status than do locals. It is important to look at the participation levels of these people to evaluate what role they may be playing in promoting, controlling or blocking project activities.

Some consideration should also be given to foreign personnel. While often not important, they can at times play a crucial participatory role. This category would include foreign donor employees, heads of private voluntary organizations, missionary personnel, expatriates, or immigrants who live and work at the local level.

Personal background characteristics are important to various kinds of participation. A wide range of such characteristics exists, and judgment is needed to ascertain the most significant in the given case. We would suggest the following are among the most important of the background characteristics: (1) age and sex (with special attention to male/female differences); (2) family status (household head, other members); (3) educational level (functional literacy, formal schooling); (4) social divisions

BEST AVAILABLE COPY

(if relevant, according to ethnicity, religion, caste, language or region of origin); (5) occupation (typically agriculturalists); (6) level of income and sources; (7) length of residence and distance of residence from the project, service or activity; and (8) land tenure or employment status (tenant with or without security, casual vs. permanent laborers, and so forth).

Each of these background characteristics can be subdivided and amplified. The occupation characteristic, for instance, can be divided into agriculturalist and non-agriculturalist. In some cases, further distinctions among agriculturalists may be necessary. Here important distinguishing features may well be: (1) size of holding; (2) ownership status; or (3) percentage of income from agricultural production only. Likewise, one could distinguish between (1) large-scale land owners; (2) small-scale owner-cultivators; (3) tenants (either renters or share-croppers); and (4) agricultural laborers. Even tenants might be subdivided into those without permanent leases and those who have considerable security on the land. And, non-agriculturalists might be subdivided in a similar way into (1) businessmen; (2) artisans or craftsmen; (3) professionals; (4) day laborers or domestic servants; (5) students; and so on into a large number of possible classifications. This kind of breakdown could be applied to any of the suggested background characteristics. The background characteristics used will depend on the type of project one is evaluating.

Several principles are involved in deciding what information to gather for assessing who participates in what activities. First, not all of these characteristics are relevant for all projects. Obviously, age and sex would be crucial for a family planning program, while one might want to know whether certain ethnic minorities or immigrant groups or casual laborers and

BEST AVAILABLE COPY

their families were utilizing health facilities.[22] Second, often the data are not readily available for making some of these distinctions, though some fairly simple, even observational data can be used to make meaningful categories.[23] Finally, where persons are using services such as schooling, credit, or clinics, the data on participant characteristics can be gathered. Rather than complete enumerations, simple sampling (like three random days a month) can be done. Where surveys are conducted,

---

[22]The Chilalo Agricultural Development Unit (CADU) in Ethiopia when it first started extending agricultural credit was not particularly aware of the emerging pattern of distribution, with tenants getting only 4 percent of the total and landowners with more than 100 acres getting one-third. Within three years, by paying attention to who was participating in the program, the latter category had been eliminated from the scheme and tenants' share was raised to 36 percent, while small farmers (under 25 acres) received 57 percent of the credit. See Table X-II in Cohen and Uphoff, Rural Development Participation, p. 248.

[23]In his analysis of extension services in Kenya, David K. Leonard wanted to determine who was being served. To do this he divided farmers into three groups constituting a continuum of practices from commercial to subsistence. Farmers who planted both hybrid rather than indigenous maize and had a cash crop such as pyrethrum or tea he put in one category; those who did either, went into a middle category; and those doing neither fell in a third. This proved to be both an easy and a revealing way of classifying persons for analysis. The "progressive" farmers (10 percent of the total) who received 57 percent of the extension visits, were 44 times more likely to be served than the "traditional" farmers (47 percent of the total) who got only 6 percent. The middle group, 43 percent, got 37 percent of the visits. See Reaching the Small Farmer: Organization Theory and Practice in Kenya (Chicago: University of Chicago Press, 1977).

BEST AVAILABLE COPY

most of the information on characteristics can be obtained.[24]

iii. How is Participation Occurring?

The how dimension adds something qualitative to the analysis of participation. It generates insights into such questions as why participation takes place, continues or declines, and why it has the particular patterns which it does. The amount, distribution and trends of participation can be assessed basically by looking at the who and what dimensions. But one should not be oblivious to the way in which participation is occurring, such as: (1) whether the initiative for participation comes mostly from above or from below.[25] or (2) whether

[24]For example, a study of the effects of land reform efforts in Nepal in which Uphoff participated in 1973 compared households according to ten occupational/land status categories, from insecure (unregistered) tenants and agricultural laborers to large landowners and business personnel. The survey with four enumerators covered 1,711 households (93 percent) in the area where pilot reform measures had been introduced 10 years before. In five weeks' time, extensive land tenure, agricultural production, economic income and political participation data were obtained. On agricultural innovation, for example, it was determined that secure (registered) tenants used four times more modern inputs per acre than insecure ones (because of the combined effect of lower income and less security); large landowners (more than 75 acres) were lower in their expenditure on modern inputs per acre than all but the insecure tenants. When asked whether land reform had improved their status, those owning land all responded about 80 percent "yes," as even the largest holdings were broken up. Among tenants, however, only 69 percent of insecure tenants said "yes" while 95 percent of secure tenants did. Revisit to Eudhbare (Kathmandu: Centre for Economic Development and Administration, 1973).

[25]This distinction is analyzed by Huntington and Nelson under the heading of "mobilized" versus "autonomous" participation in No Easy Choice: Political Participation in Developing Countries (Cambridge: Harvard University Press, 1976), pp. 7-10.

BEST AVAILABLE COPY

the inducements for participation are more voluntary or coercive.[26] It may be relevant to analyze and compare over time, (3) the structure and (4) the channels of participation, whether it occurs on an individual or collective basis, with formal or informal organization, with direct participation or indirect representation.[27] Further consideration should often be given (5) the duration and (6) the scope of participation, whether once-and-for-all, intermittent or continuous, and whether it covers a broad or narrow range of activities. Finally, it will usually be useful to consider (7) empowerment, how much capacity people have to get intended results from involvement in decision-making and implementation.[28]

---

[26] It is suggested by Sharon Perlman Krefetz and Allan E. Goodman that one should distinguish between voluntary and mandatory participation as an important "condition" of participation. "Participation for What or for Whom? Some Considerations for Research," Journal of Comparative Administration, 5, 3 (1973), pp. 368-438. They would also query whether participation was determined by the participants themselves or by others, the distinction we outlined in terms of initiative.

[27] Chambers suggests that the most convenient categories for considering experience, lessons and prescriptions are "the institutions through which local participation occurs." The channels he lists as important in East Africa are local government authorities, development committees, community development committees, self-help groups, public meetings, and local interest groups such as churches, women's groups and political parties. Managing Rural Development, p. 85.

[28] Some examples might prove helpful here: (1) initiative--villagers contribute labor to help build a village health clinic on the basis of their recognition of the need for such a clinic or in response to a request from the project director or another government official; (2) inducement--individual participation in a cooperative association because a person is persuaded by a friend of the need to act cooperatively, because he can get fertilizer at a subsidized price, or because the law requires all farmers in the area to belong to it; (3) structure--

BEST AVAILABLE COPY

To focus on who instigates participation is important. Does the initiative come from the grassroots or from the national center? More specifically, does it flow from the people themselves, from local leaders, from project staff, from local or national officials, or from foreign personnel? This question addresses the distinction between top-down and bottom-up initiation. Different patterns of initiative can mark different aspects of the project, and patterns of initiation can change over time. For example, the Minister of Education may tell the local people to raise funds to build a school, but the people may go further and offer to build it themselves. This combines both top-down and bottom-up initiative in the same project.[29] As projects progress, one frequently finds an increasing number of initiatives coming from the grassroots level.

Incentives to participation are particularly important to consider. Usually one looks to see if the participation is

---

participants in a community come to work on a school construction project as individuals or work together as members of a formal association characterized by leaders, rules and permanence over time; (4) channel--individual attends cooperative meetings and votes on all issues or is represented instead by a person who votes for all members in his area; (5) duration --self-help project requiring two days of labor from each person to build a bridge vs. a cooperative society requiring monthly attendance at meetings for an indefinite period; (6) scope--road building projects involving just labor inputs vs. participation in farmers association urging members to attend meetings, contribute savings, listen to radio programs and have wives and children participate in women's and youth clubs; and (7) empowerment--farmers able through their associations only to suggest what crops the extension service would promote vs. being able to determine what crops would be promoted.

[29] The basis for participation is likely to be stronger where there are mutually reinforcing impetuses. See discussion of a successful project with high participatory levels and mixed idea origins in: DAI, Strategies for Small Farmer Development, I, pp. 95-101.

BEST AVAILABLE COPY

voluntary or coerced, and participation initiated at the grass-roots tends to be more voluntary. Actually, a subtle continuum runs from voluntarism to coercion. Often it is difficult to distinguish where an action is on that continuum, for participation usually results from a combination of inducements, both voluntary and coercive.[30] In addition, one must look at the relationship between an initiative and the particular group of persons affected by it. Frequently, the degree of coercion will shift, depending on what particular group is being used in a given activity. For example, local leaders may voluntarily join the project, and call on the government to force landless workers to labor on it. Coerced participation is generally regarded as inconsistent with democratic values, yet there might be cases where it is required, such as ensuring compliance with a range management scheme to pool cattle, adhere to herd

---

[30] There are various ways one can characterize these different kinds of inducements. At one end of the continuum there are voluntary actions undertaken irrespective of any particular reward, because people think such actions are right, correct or socially useful. Introducing rewards for actions people would or might undertake anyway begins to change the basis of compliance to that which is voluntary but rewarded, and moving along the continuum one comes at some point to actions people would not have done otherwise but which they will do if rewarded. Beyond this, at some point one gets to things which people do not want to do but which they must do because of the threat (and possible use) of sanctions against them. In some circumstances, the positive inducements and negative sanctions may be mixed--in what we call "carrot-and-stick" situations. But along the continuum, the former diminish until we find only coercive measures left. For the sake of simplifying exposition, if not always analysis, we identify three major positions along the continuum. They correspond to the three types of compliance relationships dealt with by Amitai Etzioni in A Comparative Analysis of Complex Organizations (New York: Free Press, 1961): normative, remunerative and coercive.

BEST AVAILABLE COPY

size limits and follow rotational grazing requirements. Finally, impetus to participate and motivation for participation can combine in several different patterns, such as volunteered participation initiated from below, to enforced participation initiated from above. And, the type identified for a given project can be enriched by using the rest of our framework to ask which groups are involved in what kind of activity.

The organizational pattern greatly affects the process of participation. One of the first questions here is whether a person enters into participation as an individual or as a member of a group. For example, can any farmer get credit from the project, or must he belong to a cooperative to be eligible? Another issue relates to the complexity of the organization. Do legally-established leadership roles and well established rules govern activities? Are there clear standards for evaluating the performance of leaders? Or are the roles traditional and the procedures unwritten? If organizations are too complex it may be very difficult for local people to participate in them. More important, local elites may be able to "capture" such organizations to promote their own ends. Indeed, complexity can be designed into a project to keep participation under close control.[31]

The process of participation is also affected by whether

[31] One [illegible] that the existence of formal organization, [illegible] leadership roles, regular meetings and [illegible] participation of some significance is occur[illegible] should be taken to determine whether the orga[illegible] units, or certain aspects of it, serve to [illegible]te the kind of participation which a project should [illegible]

BEST AVAILABLE COPY

one participates directly or is represented by someone selected by him and others.[32] Direct participation probably allows the individual greater control. Unfortunately, direct participaion is often difficult to achieve because of the number who might be involved,[33] especially in rural areas marked by inadequate infrastructures, where people have great difficulty finding time to journey to far-away meetings. In general, indirect participation is more likely with decision-making activities, and direct participation in either benefits or implementation.

The time required of the participant affects the amount of participation that occurs. The longer and more regular, the participatory experience the greater the likelihood that a formal participatory organization exists.[34] Projects should

---

[32] This points attention to the channels, if any, between a rural person, the project and the larger community. Direct participation is exhibited on projects where rural people attend meetings themselves, work personally in a school-building project, or borrow credit from a cooperative society. Indirect participation occurs where farmers are represented in deliberations by a spokesman, whether elected, appointed or hereditary, where they contribute money to have skilled workers build a school, or where tenants receive credit from their landlords because they have no freehold land to offer as collateral for a loan.

[33] The scale of the organization operates to limit direct participation the larger the membership or the more complex the undertaking.

[34] A singular involvement is illustrated by the gathering on a particular market day of all or most of the rural people living in a certain district in order that they might help decide whether and where a new elementary school would be built with funding by a foreign donor. Possibly there might be follow up meetings to involve the rural people in building the school or making decisions about admissions, school fees, etc. Regular, continuous participation would occur if a parents' association were formed as a concomitant of building the new school and monthly meetings were held, dues paid, etc.

BEST AVAILABLE COPY

carefully monitor changes in the frequency of participation, giving particular attention to the emergence of more regular and continuous patterns of involvement, as well as trends in the opposite direction.

The intensity of participation in a given project is frequently related to the range of project activities involving participation.[35] Here one should consider the number of possible activities persons being analyzed could participate in. And, it is important analytically to determine whether project procedures make participation in one activitiy a precondition for other activities.[36] Careful attention should be given to the number of activities people are participating in, as well as the effects of that range on their overall participation. Multiple participatory activities may lead to inadequate participation in all of them. On the other hand, the multiple activities may reinforce each other, returning more concrete benefits to participants, and also raising their awareness of the importance of actively engaging the wider society.

Careful evaluation should also be given to participants' power. The degree of empowerment ranges widely, and it is important to know whether or not participation is a formal action, with little meaning, or an activity that allows the

---

[35] In an undertaking characterized as "integrated rural development," one might find members of a farm household involved in a cooperative society, a savings group, an adult literary program, youth clubs, women's organization and so on. At the other extreme, participation might be restricted to selling farm produce to a marketing society.

[36] For example, a farmer may have to belong to the project's cooperative society before he or other members of his family can attend adult education classes.

BEST AVAILABLE COPY

individual greater control over his or her life.[37]

From our reading, we would give particular attention to the structure and channels for participation. Participation can be individual and unorganized, but usually needs organized expression and support to be effective and sustained. One of the hypotheses most worth examining is the extent to which organization conditions the amount, kind and success of participation, recognizing that these may vary for different tasks or for different groups.

We believe these different characteristics will illuminate the possibilities, dynamics, and consequences of participation, if applied appropriately to the assessment of who participates in what rural development activities. They may be combined in interesting and different ways,[38] or compared to each other in

---

[37] Empowerment could be high for relatively inconsequential choices, such as where to build a cooperative's new storage shed, and yet minimal for controlling the actions of the cooperative's director, who may be making adverse decisions affecting the farmer's earnings. On this, see Mathew D. Edel, "The Colombian Community Action Program: Costs and Benefits," Yale Economics Essays, 9, 2 (1969), pp. 3-55. This aspect of participation is sometimes characterized as "genuine" versus "symbolic" participation. See Krefetz and Goodman, "Participation for What or for Whom?"

[38] For example, Mary Hollnsteiner presents six modes of participation that represent degrees of empowerment partly because they reflect differences in structure and channels as we have defined these terms: (1) unofficial and indirect representation of the people by "public citizens" who discuss and endorse programs planned outside the community; (2) appointment of local leaders to positions in the government bureaucracy, coopting them into planning and administration; (3) choice of final plan from among predetermined options provided to community; (4) ongoing consultation with community, starting with plan formation; (5) representation of people's spokesmen on official decision-making bodies; and (6) community control over expenditure of funds in programs affecting its

BEST AVAILABLE COPY

various matrices.[39] Such analysis can illustrate the complexity of the how dimension. No single assessment about participation can be made, because, for example, there is no way to conclude that participation of narrow scope but high empowerment is "less" participation, than that of bread scope but little empowerment (combination B and C in the figure below). One can probably say that A is more participatory than B or C, and either of them more than D.

Figure 3: Two Dimensions of How Participation Occurs

SCOPE	EMPOWERMENT: High	EMPOWERMENT: Low
Broad	(A) High Empowerment Broad Scope	(C) Low Empowerment Broad Scope
Narrow	(B) High Empowerment Narrow Scope	(D) Low Empowerment Narrow Scope

According to most views of participation, that which is initiated from below, voluntary, organized, direct, continuous, broad in scope and empowered would be the "most" participatory. But judgments can differ about anything diverging from this ideal-typical form. This is not only because people have different values and expectations about participation, but because different aspects of participation are relevant to different

area. "Local Initiatives and Modes of Participation in Asian Cities," Assignment Children (UNICEF), 40 (1977), pp. 19-27.

[39] See for example our Table IV-2, Cohen and Uphoff, Rural Development Participation, p. 106.

BEST AVAILABLE COPY

projects. Indirect participation through representatives may be appropriate and satisfactory in some situations, but not in others. For maintaining irrigation canals, periodic participation may suffice, whereas continuous participation may be needed for distributing water. In analyzing the how dimension we want to alert persons to the ways participation by certain groups (who) in given activities (what) can vary. Even if no quantitative value is attached to these aspects, one should be sensitive to what they are and particularly to changes, such as shifts from bottom up to top down initiative, or from voluntary to more coerced performance.

Contexts of Participation

i. Project Characteristics:

Briefly, we have identified ten major characteristics in rural development programs and projects. These also are elaborated at length elsewhere,[40] and will only be listed here with a brief illustrative example of their effects. (1) Technological Complexity: a cooperative society organized for a comprehensive agricultural package program that requires complex accounting and marketing practices, may exclude less educated persons from leadership roles, and thus bias participation toward "local" farmers or professionals. (2) Resource Requirements: a mechanized cultivation project may require participants to have a minimum amount of land to be eligible for tractor services, thereby excluding persons with less land; similarly, a credit scheme requiring collateral may exclude tenants from participation. (3) Tangibility of Benefits: farmers already actively participating in a cash crop economy may join a project that offers to increase yields 25 percent, while

---

[40] Cohen and Uphoff, Rural Development Participation, pp. 112-138.

BEST AVAILABLE COPY

more "traditional" farmers may not join unless they can see yields double. (4) Probability of Benefits: larger owner-cultivators may be more likely to participate in a cotton-growing project than are smaller cultivators, because the former will still have enough land to grow food for their families, in case the new cotton crop fails or the price drops. (5) Immediacy of Benefits: farmers are more likely to be willing to participate in building a bridge that helps them get their crops to market, than in a reforestation project, the benefits of which will take a long time to result. (6) Equity: a poultry project requiring participants to invest capital is likely to have more restricted participation in all phases, than is one providing week-old chicks free or cheaply to all who want them. (7) Program Linkages: women's participation in a population control program is likely to be greater when connected to a child care program; similarly, farmers' use of fertilizer is likely to be affected by its being connected to a credit program. (8) Program Flexibility: an adult education program that does not provide for local input could lose participants, if it failed to respond to local demands for a voice in selecting or removing teachers, or designing lessons. (9) Administrative Accessibility: a project to promote local handicrafts is likely to attract more participation if artisans can meet with and get suggestions acted upon by those who control the staff and funding for the project. (10) Administrative Coverage: a low ratio of extension workers per farmer will likely limit the farmers' participation, as the program's services will not be widely available; it is possible, on the other hand, that too high a ratio may discourage farmers' participation in decision-making. Because these project characteristics appear self-explanatory, we will not elaborate them here. They need, however, to be considered since project design can leave little scope for participation, or allow only certain kinds of participation to emerge.

BEST AVAILABLE COPY

ii. Task Environment:

As with project characteristics, the project's physical setting, the history of the region, and the social systems operating in it, have powerful and subtle affects on participation patterns. We have identified six sets of factors, which should be immediately recognized by most social scientists working at the rural level. (1) Physical and biological factors: a long rainy season may make it impossible to hold regular cooperative meetings throughout the year, because roads and paths are impassible; or poor soil fertility for upland farmers may mean they must work harder than lowland farmers, leaving no time for participating in farmer organizations. (2) Economic factors: land tenure conditions may obligate tenant farmers to landlords, who are opposed to the formation of new cooperatives. (3) Political factors: local government units may be more an extension of central government authority than representative of local population, and lack a tradition of exercising local authority; or the national center may fear grassroots mobilization and give only superficial support to participatory activities. (4) Social factors: farmers may live in isolated homesteads making cooperatives difficult to organize; similarly, cumulative social cleavages can make it difficult to form cooperatives not controlled by wealthy, landed, and dominant groups. (5) Cultural factors: in certain communities, rules may not let women leave house compounds, let alone join home economics clubs; or general attitudes of family loyalty and inter-family competition may inhibit farmers' associations. (6) Historical factors: prior experience with a project whose rice seeds failed to germinate may make farmers unwilling to adopt new high yielding varieties; or an experience of previous embezzlement of self-help funds raised by the community may lead many local people to distrust new community development efforts.

BEST AVAILABLE COPY

There is no definite list of critical environmental or task factors.[41] Rather programs and projects need to be sensitive to the relevant ones and their likely effects. In this way participatory goals are more likely to be reached and obstacles to participation circumvented.

Participation for What?

It is important also to consider the purposes of participation. Because it involves normative judgements, assessing

---

[41] Examples of the kinds of factors that should be considered are: (1) Physical and biological factors: climate, weather fluctuation, rainfall, soil fertility, water supply, elevation, terrain, vegetation patterns, insect and animal pests, population size relative to land resources. (2) Economic factors: land tenure and ownership patterns, agricultural production patterns, land rents, occupational patterns, crop and livestock resources, income and expenditure levels, savings, investment and credit; employment possibilities, level of industrial development, markets and transport, physical infrastructure. (3) Political factors: centralized vs. decentralized structure of government, competitive vs. single party system, tradition of local government (or none), linkages if any of central elites to rural areas and problems, prevailing ideology, orientation toward participation by rural people. (4) Social factors: settlement patterns, nuclear vs. extended family structure, clan, ethnic or voluntary association memberships, caste or race divisions, social stratification and class, cumulative vs. cross-cutting social cleavages, local institutions for conflict resolution, rural-urban differences, patterns of migration. (5) Cultural factors: values relating to place of agriculture in people's lives, sex roles and division of labor, orientation toward future and toward change, attitudes toward group activity and cooperation, patterns of political and social deference, attitudes toward role of women in local and national society. (6) Historical factors: past relationships between this area and the national center (cooperative or hostile), traditional rivalries between towns within area, past experience with central government initiatives for rural development, levels of technological sophistication in area.

BEST AVAILABLE COPY

purposes is even more likely to lead to disagreement than are the more descriptive dimensions, discussed above. Many different purposes for participation can be listed, such as Chambers offers in his book.[42] The difficulty with constructing an analytical framework for purposes is that their assessment, and even their factual basis, shifts depending on whose perspective one takes. The government, for example, may view taking credit for new maize varieties as a means for reducing national food deficits, or for helping stabilize the regime; farmers may see the same action a way to augment family income and food consumption.

As with all objectives, they may be intended or unintended, stated or unstated, achieved or unachieved. Considering the purposes for which participation is undertaken or advocated is useful, but analysis of purposes cannot be as rigorous as for the three dimensions outlined already. One of the key questions to ask, is whether the authorities' objectives for getting people to participate are the same as, [illegible] compatible with, the purposes the people would accept as the[illegible] Where governments want things from or for the people that the people do not want for themselves, ambiguities and obstacles affecting the intended participation are likely. We can imagine sinister purposes

[42] Chambers says "local participation can be analyzed in terms of objectives and functions." The values ascribed to participation in its various forms include: making local wishes known; generating developmental ideas; providing local knowledge; testing proposals for feasibility and improving them; increasing the capabilities of communities to handle their affairs and to control and exploit their environment; demonstrating support for the regime; doing what the government requires to be done; extracting, developing and investing local resources (labor, finance, managerial skills, etc.); and promoting desirable relationships between people, especially through cooperative work. Managing Rural Development, pp. 85-86. These categories are, unfortunately, more descriptive than analytical.

BEST AVAILABLE COPY

such as encouraging critical participation from the public to identify malcontents (some thought this was done during the campaign in China to "let a hundred flowers bloom"). On the other hand, some may try to use the opportunities for participation provided by a government to effect its downfall.

Some distinctions about purposes of participation need to be made, the most important of which relate to agreement and openness. Specifically, consider action is important of whether the government, project staff and local people agree on the purposes of participation. And one needs to know if the stated purposes of participation are operative, irrelevant to the situation, or masking concealed intentions.

Questions can be raised about participation's purposes in terms of who is supposed to benefit from it--participation for whom?[43] According to Holm, the Village Development Councils in Botswana could operate only under tight political-administrative control, and the plan gave no funds for local projects, so popular participation was mostly an instrument for bureaucratic domination of the village modernization process.[44] This may or may not be true, and may or may not be common elsewhere, but is certainly a valid question to consider in evaluating participation. Some would dismiss activity such as Holm describes, as not constituting "participation." But we think it more appropriate to make qualitative assessments of the kind of activity involved and its effects. An analysis which identifies who is participating how in decision-making and implementation, and who is participating in benefits and evaluation should illuminate substantively concern with "participation for whom and for what?"

---

[43] See Krefetz and Goodman, "Participation for What or for Whom?"

[44] John Holm, "Rural Development in Botswana: Three Basic Political Trends," Rural Africana, 18 (1972), pp. 80-92.

BEST AVAILABLE COPY

THE SKILLS REQUIRED TO IMPLEMENT PROJECT IMPACT EVALUATIONS

The skills required to implement AID's project impact evaluations will vary to some degree with the scope and methodology options chosen by the Agency. In this subsection, the skill requirements for various options are outlined, and their availability, inside of the Agency as well as through arrangements with other organizations, is examined.

The skills needed to carry out impact evaluations for AID's projects cluster into several categories, including:

- Issue/Evaluation Question Identification;
- Study Design;
- Instrumentation/Procedures Development;
- Data Collection (including document research, sampling, interviewing and other skills);
- Data Analysis;
- Technical Writing and Presentation;
- Dissemination/Utilization Management.

1. Issue/Evaluation Question Identification

For every cluster of similar projects, and to some degree for each project, the general evaluation scope developed by the Agency will need to be customized. This customization will necessarily include the specification of issues/questions that are to be addressed under each of the categories identified in the general scope. What is required at this point in a project impact evaluation is a combination of background (in the technical area and in the

BEST AVAILABLE COPY

specifics of t. ojct situation) and skill in articulating AID's concerns in a set of "researchable" questions. If "scopes of work" in past evaluations are taken as an indicator of current AID ability to perform this task, PCI would have to conclude that while AID has adequate background (and perhaps is the only source of needed background information) to prepare specific questions/issue for project impact evaluation scopes, it lacks skill/discipline in stating what is required from an evaluation as a set of "researchable" questions.

Many scopes today tend to list a large number of factors that are to be "looked at" in the course of an evaluation, but fail to identify clearly the information that is required, and why it is needed (the action possibilities). Improvements in this area are desirable, but it is practice rather than outside expertise that is needed to make the change. These improvements are needed irrespective of what scope/methods choices AID makes.

2. Study Design

The skill requirements in terms of study design will be affected by AID's scope/methodology choices. If the Agency elects the "minimal" scope position, and interprets the requirement to demonstrate an association between project action and impact as necessitating verification of a predictive (rather than explanatory) hypothesis, certain skills may not be needed, e.g., experience with experimental and quasi-experimental research designs.

Evaluations carried out under the "minimal" scope will need to be designed so that AID can detect whether impact has occurred. The skills required to set up a valid comparison, and to measure (either two groups, or one group at two times) on the relevant indicators of performance must come into play at the point of evaluation design. Ideally, evaluation designs prepared at the time of project approval, whether two-group or longitudinal comparisons, are appropriate. At this point, AID should also identify the indicators of impact that will need to be assessed.

BEST AVAILABLE COPY

The skills involved in identifying what type of comparison is appropriate for a specific project are not complex. An understanding of the project situation (and what potential comparison areas/groups exist) is needed, as is a sense of the importance of gathering data on specific individuals versus clusters of people. Since AID evaluation plans do not at present identify what comparisons are appropriate for assessing impact (except in occasional and pilot projects), there is little evidential basis for drawing a conclusion on the availability of this skill in-house. However, since the issue is not a terribly complex one, it seems fair to assume that AID's staff could assume this new responsibility without external assistance.

If AID decides instead to assess its projects' explanatory hypotheses through experimental or quasi-experimental evaluation designs that involve the identification of "matched groups" for comparisons and the randomization of group members to "treatment" and non-treatment status, PCI's conclusions concerning the availability of requisite skills in-house must also change. There are very few people anywhere who have proven that they are skilled in setting up and managing experimental designs in the social domain. While a large number are familiar with the theoretical requirements of experimental and quasi-experimental designs, the number with practical experience is limited indeed. And while it is only fair to assume that some of those who have the theoretical background to design and manage these studies would succeed if given the opportunity, it is probably also fair to assume that there would be some casualties, too.

The skills involved in identifying indicators of performance and ways to measure whether targeted levels of performance were achieved are more widely distributed in general, and within the Agency. For a number of years, AID's staff have been involved in identifying and targeting indicators of various types of performance. The quality of the work in this area varies. Those with more practice, and more opportunity for practice, may be the source of the best examples (at least this is the claim made by individuals who have taken the PD&E course and been interviewed, formally and/or informally, some time later). AID's practice with the specification of measures (or Means of

BEST AVAILABLE COPY

Verification) has either not been as adequate as its work with indicators, or the skills required are simply different ones. In any case, the product is not generally up to the level of AID's average performance in indicator specification. Too often Means of Verification in a project design summary simply says "survey" or "host records", with little guidance about what might really be involved in collecting data on the various project indicators. On the other hand, the skills involved are not too complex to transfer into the Agency on a wide-spread basis. (At least this is the assumption being made by the DSP program, in the current "Manager's Guide to Data Collection" and in other documents that are being prepared by and for the Agency staff.) While the specification of measurement approaches, at a general level, appears to be a skill that, with practice, can be acquired by most of AID's design and evaluation staff, familiarity with the detailed aspects of measurement may not be so wide-spread or easily acquired. These aspects of the skill composite needed to carry out project impact evaluations are discussed further below.

3. Instrumentation & Procedures Development

The instrumentation and detailed measurement skills needed for the project impact evaluations will be a function of the specific impacts to be measured, e.g., nutrition, income, etc., and of the degree of rigor implied by the scope/methodology choices AID makes at a general level. As suggested elsewhere, AID can cut costs and skill requirements to the degree that it is able to standardize (repeatedly use the same) measures of various impacts. Where this is possible, the basic instrumentation/measurement dilemmas may need to be faced only once or twice (depending on the degree of regional and other variation in how a concept can be measured).

The second aspect of this step in evaluation implementation will depend on choices AID makes concerning the degree to which it expects evaluation answers to be representative of the groups/areas examined. When a high degree of representativeness is required from a study that deals with the impacts of a

BEST AVAILABLE COPY

project on a large group or area, skills in identifying units of analysis, developing list or area frames from which samples can be drawn, etc., come into play. (Parallel skills are required when a high level of reliability or repeatability, or strong statements about the validity of study answers are needed.) In summary, the more AID wants in the way of quality, the more it will need in the way of skills at the point of design and instrumentation for its evaluation studies. High level skills in these areas tend to be scattered.

Here and there, in AID, in its RSSA, PASA, and contractor rosters, and in universities, there are individuals who are highly skilled in each of the requisite domaines. However, AID should not think that those who are good at developing sampling frames are not necessarily good at finding out whether a question that appears meaningful in Washington is also meaningful, in the same terms, in Bamako or Santo Domingo. The first requires a knowledge of demographics and statistics; the second an understanding of developing countries in sociological/anthropological terms. Some care on AID's part, and perhaps a voluntary self-rating by its staff and the outside personnel that work with/for AID, will be needed to establish (on a preliminary basis) who's good at what aspects of preparing evaluation designs and procedures. In some cases, AID may find it desirable to pull together skills from different sources at different points in an evaluation study, or to put together individuals from different organizations to make up highly skilled evaluation teams. Alternatively, it will simply need to specify more clearly the practical skills that are required when it asks a single organization to pull together a team and manage an evaluation for AID.*

---

* It is worth underscoring the fact that practical applied skills, rather than degrees, are what AID will need, i.e., documents such as RFPs will need to focus evaluation point distribution on "experience in" credentials, as well as adequate academic background, if AID is to be assured that the outside personnel it recruits for evaluation work can provide the desired quality product.

BEST AVAILABLE COPY

4. Data Collection

As in the case of instrumentation, described above, the specific skills that are needed will depend on both the subject area of the evaluation and on AID's scope/methodology choices. The common denominator again will be practical field experience in such areas as interviewing, sampling, field staff supervision, case study development, the identification of local informants, etc. The list of skills AID may need is long, and once again expertise in these areas is not necessarily centrally located in the Agency, or to be found in a single outside organization, particularly in light of the fact that many evaluations will need to be carried out in the local languages of project areas.

Whereas a number of the prior steps/skills can be supplied by AID itself, field data collection skills are an area where AID should seriously consider the use of host country personnel. In many of the countries in which AID works there are local research organizations (both independent firms and units of government agencies and universities). Where these exist, AID will both realize economics, and probably improve the strength of its data, by working with local groups, either directly or as sub-contractors to such universities, firms and other U.S. agencies as AID may recruit to implement its project impact evaluations. A general effort that involves both Missions and those who work with/for AID in each of the countries in which the Agency operates, should be made to identify qualified local groups that can participate in these evaluations.

5. Data Analysis

The data developed by AID's project impact evaluations will be both qualitative and quantitative in most cases, and an evaluation team should contain individuals who are skilled in each type of data. (If only one skill is present, e.g., ability to manipulate quantitative data, some of the "lessons" from an evaluation may never fully emerge.)

BEST AVAILABLE COPY

The analyses to be performed for most of the project impact evaluations need not be highly complex. In all cases, AID will need to know whether there was impact in the project area/group. This will usually involve the tabulation of data from the project area/group, and from some comparison area/group. Fairly simple statistical tests are the normal approach for making such comparisons with quantitative data. (Complex tools such as regression analysis are not normally needed to answer the types of questions AID will be asking.) With the qualitative data generated by these studies, AID should be trying to establish what patterns of social/cultural/political interaction are involved in the production and distribution of benefits. For this type of analysis either narrative or graphic (schematic) analysis is often used. In selecting individuals within the Agency or outside it to carry out the analyses of project impact data, AID would be well advised to look at samples of past work done by its candidate evaluation specialists. Real analytical ability (the skill of working with data and synthesizing it so that conclusions and recommendations can be drawn) may not correlate highly with other easily identified skills.

6. Technical Writing & Presentation

When evaluation reports are poorly written they may receive less attention than their substance deserves. AID itself knows what writing styles, types of graphics, etc., are most readily understood by its staff and Congressional personnel. Investments by AID in the way the final product reads may be merited in some cases, and may actually add to the value of the work.

7. Dissemination/Utilization Management

Once evaluations are received by AID there are a number of steps/skills that are important in translating evaluation findings into action inside the Agency. These skills, which are further discussed in Section Five (under "roles and responsibilities"), can make the difference between a report that goes on

BEST AVAILABLE COPY

the shelf.and one that leads to better projects as well as a better convergence between AID and Congressional perceptions concerning the merit of specific (types of) projects.

BEST AVAILABLE COPY

Impact Evaluation Coverage

CHAPTER SIX

AID has two basic long-term options with respect to evaluation coverage for its mature and terminated projects:

- Universal coverage;
- Partial coverage, using one of a number of approaches, including:
  - A decision driven approach;
  - A nomination approach;
  - A criteria selection approach;
  - A sampling approach;
  - A combination of two or more of the above.

Annex A to this Chapter provides a technical discussion of these Agency options for evaluation coverage in the long-term. PCI's review of these options has led PCI to two conclusions.

There seems to be no reason why AID should make a choice between universal and partial coverage at this point in time. The decision can, and probably should, be deferred. Initial partial coverage can be converted to universal coverage at a later time, if that proves to be desirable once AID has acquired additional experience with impact evaluation.

Secondly, there appears to be no real reason at the present time to make a long-term decision concerning what type of partial coverage approach to use. In addition, there are some short-term reasons which suggest that a combination "nomination-criteria" approach will help AID get started. Experimentation with a sampling approach would also be useful, before it was actively considered

BEST AVAILABLE COPY

as a candidate for the long-term. And finally, AID might find if it is to make impact evaluation a meaningful activity inside the Agency, it may need to respond, at least in part, to the regional bureau and Mission sense that selection should be driven by decisions that need to be made about future programs and projects.

In addition to the conclusions reached as a result of the review of the specific coverage options open to AID, the full range of information developed in the course of this study leads us to conclude that:

- Universal coverage presently, and for the foreseeable future, exceeds the level of coverage required to answer the important questions that are being raised about AID's impact;

- The level of coverage AID needs, over time, has identifiable characteristics:

  1. It is the minimal level needed to determine the internal and external validity of the hypotheses about impact on which AID's projects are based. At present there are clusters of projects that utilize the same hypotheses; we expect that there will be clusters in the future. Not every project will need to be evaluated to determine the internal and external validity of the hypotheses that support these project clusters;

  2. It is a level AID can manage. To PCI this means a level which AID can plan, execute using standards of evidence that allow AID to trust the evaluation findings it secures, and which AID can absorb and utilize.

The characteristics of an optimal coverage level, identified above, are discussed further below.

PCI's review of the objectives in the FAA against 1,940 AID projects has demonstrated that a large number of AID projects focus on a few indicators of change. Further, the intermediate work we did in the course of this project using the automated databases (as well as the work we did for PPC/E/S on "Sectoral Mapping") strongly suggests that the number of approaches AID uses to address a given problem (or objective) is finite. Even if there were five basic approaches that AID projects used to address each of the seven objectives/indicators identified by 80% of the projects claiming mandate impact,

BEST AVAILABLE COPY

the total would only reach 35. Double the number of approaches or the number of planned impacts and you still have only 60 or 70 truly different project models.

From everything PCI could gather by reviewing AID and Congressional documents, it is project models (sets of hypotheses that AID and the Congress can believe, with some evidential basis, will yield predictable impacts in new situations) that AID needs to evaluate, not each and every Agency project. If one conceives of the project model as the "unit of analysis" for impact evaluations, the idea that there is a limited and manageable volume of evaluation needed to secure important impact information becomes meaningful. Solid evidence concerning each of AID's main project models is the concept behind PCI's statement that AID needs a minimal, well focused level of impact evaluation coverage.

The second characteristic, AID's ability to manage a program of impact evaluation is also essential. PCI's sense of the situation is that AID will not make significant advances in its understanding of project impact by taking on more evaluation studies than it can plan, execute in a way that secures believable evidence, and absorb into its policy, program and project planning activities. Further, we do not see evidence at the present time that AID is well prepared for any of these tasks.

Over the summer, PCI discussed a number of steps AID could take to (a) begin to get impact data on its projects and (b) simultaneously build its capacity to plan, execute and absorb impact evaluations so that over time it could, if it chose to, increase the volume of impact evaluation without sacrificing basic quality standards (such as a requirement that every evaluation compare data on the project area/group after the project to some legitimate referent). The "shopping list" of approaches that were discussed and examined by PCI to determine whether they were feasible and would achieve the twin objectives noted above included:

- Standardization of the "scope of work" for project impact evaluations in a way that could be broadly used, with only minor needs

BEST AVAILABLE COPY

for customization. (The work we have undertaken in conjunction with the development of Annex II.A. suggests to us that such a standardized scope can be developed.);

- Standardization of the ways AID measures the key impacts with which it is concerned. (Our analysis of the degree to which AID projects cluster around a few key indicators and objectives in the FAA, and work of the type done by Dr. Daines in preparing what is Exhibit IV.A.1 of this report, as well as full-scale impact evaluations in Haiti and elsewhere suggest that this second step is also possible. The work done by Cornell on participation (Exhibit IV.A.2) provides a workable starting point for one of the less well-developed measurement areas.);

- Start by evaluating projects that AID expects will demonstrate impact, benefit distribution and a connection between impact and development assistance. A short-term strategy of working with "best cases" is not simply a way of trying to "look good". It is a rational approach to practicing impact evaluation skills and at the same time getting quickly to the point where at least some of AID's project models have been validated. If AID picks projects that it knows did not provide all of the Outputs, or where a key assumption failed, it is almost assured of finding either no impact or that its Outputs and assumptions weren't necessary. AID will eventually select projects that yield these findings; and there is just as much if not more to learn from an evaluation that proves a project model is not valid. However, from a practical and political standpoint, AID just doesn't need to start with these cases;

- Start by evaluating projects that someone, somewhere knows how to evaluate. There are some impact evaluation designs in AID's past evaluation efforts. Bring them out, copy them, modify them for new situations...but don't, if at all possible, pick situations where there is no existing set of impact measures or proven approach for getting data. The point is the same as above. AID simply does not need to start with its most difficult methodological problems;

- Involve AID staff in the technical aspects of the evaluations, as many aspects as possible. PCI found that not only are AID staff members unfamiliar with the technical details of evaluation work from a practical, fieldwork perspective, but also found some evidence to suggest that they are not sufficiently familiar with what evaluators do to be able to tell the difference between a strong study and a weak one. Hands-on involvement of the AID staff, at the technical as well as "paper processing" level, is the best way to rapidly alter this situation;

- Try to involve top-notch evaluation specialists in the U.S. in the AID effort. Even if these people cannot devote the time necessary to carry out impact evaluations, they can provide highly valuable

BEST AVAILABLE COPY

advice, and give AID a better purchase on the activities it will be undertaking, by reviewing scopes of work, evaluation designs, instruments, measurement approaches, analysis plans, and draft evaluation reports for the Agency.

As in the case of a number of other characteristics of the evaluation of mature and terminated projects, the coverage issue appears to PCI to be tied to the question of what Congress expects from AID. If, at present, the choice between quantity and quality is real, and it seems to us to be so, PCI suggests that AID choose quality.

As indicated elsewhere in this report, PCI is fully aware of the pressure on AID to begin producing impact information. Nevertheless, we have recommended that AID start slowly and increase volume only when it knows that it has quality under control. One of the main reasons for advancing this note of caution is that Congress, at this point in time, may well be more knowledgeable about impact evaluation than AID is. While AID's Mission staff may feel that a short visit by an "expert" who spends a few days or weeks observing a project and then provides feedback to AID is just what is needed, it is not at all clear that this type of information will meet a Congressional definition of "evidence".

In our initial discussion of what PCI and AID's views on coverage meant in terms of absolute numbers of evaluations per year, we found ourselves to be quite far apart, with PCI taking the low side on a discussion of numbers. Under the present circumstances, a discussion of numbers for the short term appears to be a moot question. However, we do feel that as AID undertakes the 20-30 evaluations it intends to perform, the Agency's chances for success will be improved if it:

- Picks projects that it has some reason to believe yielded impact;
- Uses measurement approaches and evaluation designs that have been tested;
- Involves its staff in the technical aspects of evaluation, and seeks the guidance of evaluation specialists as it prepares scopes, designs, instruments, analysis plans, etc.;

BEST AVAILABLE COPY

- Avoids "short-cuts" which may prove to be embarrassing later.

It is entirely possible, and perhaps appropriate, for AID's method of choosing which projects it will evaluate in a given period to go through several phases. If no rigid decision were made, several variants might be useful at different times, e.g.,:

- Phase I: The Initial Year(s):
  - Criteria selection used to locate projects that have produced Outputs and reportedly have impact;
  - Nominations from Missions considering large investments based on evaluation findings included in set of evaluations on "service" grounds;
  - Demand driven selections responding to Congressional inquiries.
- Phase II: The Middle Year(s):
  - Criteria sampling used to locate evaluable examples of main "project models";
  - Some random sampling to determine whether areas of importance are being missed;
  - Some response to nominations from Missions and elsewhere and some decision driven selections from the top.
- Phase III: The Long Term:
  - Nominations the primary mechanism, particularly requests that relate to the development of new programs/projects;
  - Some criteria or random approaches retained at low level to ensure that difficult projects/evaluations are not overlooked.

BEST AVAILABLE COPY

THE ADVANTAGES & DISADVANTAGES OF COVERAGE & SELECTION OPTIONS

The two basic coverage options, universal and partial coverage, are reviewed in this Annex, as are several methods of selecting projects for evaluation under the partial coverage option.

1. Universal Coverage

AID's automated databases indicate that in any given year there are roughly 700 active projects (for FY77-FY81). The annual turnover, new starts and terminations, involves approximately 200 projects. This suggests that the universe of projects which could be evaluated in impact terms (if all projects were evaluated at the same point in their lifespan) might range from 150 to 300 per year. If, on the other hand, we consider all completed projects as a cluster from which selections can be made, the universe is substantially larger, suggesting that AID would need to undertake more than 300 project impact evaluations per year to work off its backlog of completed projects. Unless AID is going to be satisfied with either very superficial (e.g., one- or two-day "flying expert") assessments, or can put forward an extremely large evaluation budget, the idea of universal coverage appears untenable. Even if AID could afford the level of investment required, it is questionable that the manpower to carry out such a program and the absorbtive capacity of the Agency with respect to evaluation findings could be expanded to carry out and utilize such a program.

Further, it is not at all clear that universal coverage is what is required from a technical and information standpoint. From an information point of view, AID needs to acquire information that allows the Agency to state with a known degree of certainty that a project had/will have a predicted set of impacts, and further that the distribution of project benefits were/will be

BEST AVAILABLE COPY

within a predictable range. In order to make such statements, AID potentially needs information on each of the means-ends (project action-impact) relationships its projects hypothesize.

PCI's review of the AID portfolio suggests that the number of individual means-ends hypotheses put forward by projects is significantly less than the number of projects, i.e., clusters of projects use the same hypotheses. While AID might well want to evaluate more than one project per cluster, universal coverage of the larger clusters appears excessive from a technical or information standpoint.

Despite the fact that universal coverage is expensive in terms of manpower and other resources, as well as being in PCI's judgment excessive, AID has actively considered this option in its internal discussions. To PCI it appears that the discussion of universal coverage may relate to AID's perception of what Congress expects. As pointed out in Chapter 1, PCI found no evidence that Congress expected AID to elect this option if another was appropriate for meeting the requirement to demonstrate impact and the relationship between impact and development assistance.

It is worth noting that AID's option concerning universal coverage would not disappear if it elected a less ambitious course at the start of its impact evaluation effort. A final decision on universal coverage can be delayed until after AID has begun a program of impact evaluation and discovered the problems that cannot be anticipated now as well as learned a good bit more about how many evaluations the Agency can manage and absorb in a given period of time.

2. Partial Coverage

Anything less than universal coverage is automatically a form of partial coverage. However, PCI distinguishes between two types of partial coverage. The

BEST AVAILABLE COPY

first type is percentage or formula-based partial coverage; the second is decision-driven partial coverage. Both are discussed below.

There are several ways AID might go about setting a level of partial coverage. The options for determining what percent of projects or the absolute number that are to be evaluated in a given year are related to the ways in which a portion of the portfolio might be selected. Three basic approaches to selection could be used, alone or in combination. They are:

- Nomination (or volunteer) approaches;
- Criterion selection approaches;
- Random selection approaches.

In addition to these three basic options, AID could utilize any one of a number of combination options, e.g.:

- Criteria selection of topics, random sampling inside each;
- Nomination of countries, and criteria selection within each;
- Random sampling of countries, and nominations within each;
- etc.

There are advantages and disadvantages to each of the basic selection approaches identified above, and for each of the many combined selection approaches AID could use. The advantages and disadvantages of each are briefly reviewed below.

Nomination, or volunteer approaches, are not representative. Use of this type of approach will not yield selections that tell AID very much about its "across-the-board" impact. In general, nomination approaches tend to bring forward the best cases. Yet even this generalization is not totally reliable. Normally a nomination approach prompts people to show off their best examples, but it is also possible that for reasons that are sometimes personal, an example of a failure would be nominated.

BEST AVAILABLE COPY

Criteria selection approaches, like nomination approaches, are non-representative in character. However, when one knows fairly clearly exactly what type of projects and situations it would be valuable to examine, a criteria selection approach is more reliable than either nominations or random sampling in ensuring that the topics or projects that should be evaluated will be evaluated. For AID a criteria selection approach would involve more pre-selection work than either of the other approaches. First, agreement would have to be reached on what criteria to use. Second, projects would have to be screened to determine which ones met AID's criteria. (The type of screening process used would depend on the criteria. Automated screening might be possible for some or all of AID's criteria.)

There are a number of criteria that might be appropriate for inclusion in a criteria selection approach for choosing projects to be evaluated in terms of their impact, including project size, type, country, period of time since termination, etc. Other factors might also be incorporated such as:

- Whether the project is trying to impact current goals;
- Whether AID plans to fund this type of project again;
- Whether the project was, by reputation, a success or failure;
- Whether the project used tried-and-true approaches or experimental/innovative ones;
- etc.

In addition, there are three factors that would need to be included if AID chose to limit its selection of projects for impact evaluation to those in which it could fully expect to be able to measure impact. The three basic technical tests projects would need to pass in this case would be:

- The project Outputs have been produced;
- There were no major failed assumptions reported;
- Impact was expected to be observable by the time the evaluation would take place.

BEST AVAILABLE COPY

If AID applied these technical criteria to its project portfolio in an effort to find "cases" of each of its major means-ends hypotheses that were ready for impact evaluation at the present time, it would probably find that such situations could be identified for only a portion of the hypotheses it was able to identify for validation. In order to examine the hypotheses for which no "case" presently exists, AID, using this technical perspective, would need to wait until the requisite conditions were in place. Coverage of the hypotheses AID makes about impact would, if technical selection criteria were applied, necessarily begin with those hypotheses where project "cases" existed that met the pre-evaluation criteria listed above.

Random sampling as a selection approach would guarantee AID a representative set of projects, and offer the Agency the best chance of making statements about its impact "across-the-board". This type of approach might well lead AID to select projects for evaluation that it would not have considered worthy of the effort on some other grounds. If AID used this type of sampling approach, with or without stratifying AID's portfolio first, the Agency would have to be prepared to follow through and evaluate projects it didn't really care about evaluating, since as soon as it rejected one of the projects selected by this method the claim to "representativeness" would be undermined.

Use of a random sampling approach, like use of the technical approach outlined above, would require some pre-selection screening. AID would need to determine, at minimum, which projects should be considered mature, i.e., projects in which impact could be observed. The apparently simple rule-of-thumb that terminated projects are mature projects appears to be a less than satisfactory solution. PCI's review of selected AID projects during this engagement indicated that in some projects there is an expected lag between the production of Outputs and Purpose achievement or Goal level impact. Random selection based on termination date would not necessarily yield only projects that were ready for impact evaluation. An option here would be to extend the date to "termination plus X" with X being a sufficient number of months beyond termination to ensure that Goal level impact would be observable in all

BEST AVAILABLE COPY

projects selected. To exercise this option, AID would need to review project design papers to identify X before setting up and applying its random sampling procedures.

There are advantages and disadvantages to using a random sampling approach. On the positive side, AID would be fairly certain that its selection procedure was not overlooking or avoiding either one-of-a-kind projects (that might be successful and serve as models for other areas) or major failures. Random selection would also tend to ensure that AID didn't inadvertently concentrate its evaluation efforts in one sector or geographic region. On the negative side, this process would almost certainly lead AID to evaluate some projects that it already knows it will never repeat.

This description of advantages and disadvantages of random sampling has ignored the possibility that projects which do not produce their Outputs or in which assumptions failed, may still achieve Purpose and Goal level impact. Using only situations in which the validity of project hypotheses can clearly be examined tends to rule out an effort to learn what Outputs aren't necessary and what assumptions aren't important. Since this information would be of value to AID it may be important to try to ensure that it is secured either by using partial if not random sampling in selecting projects for impact evaluation, or by trying to identify this type of situation through AID's regular project evaluation system.

A "decision-driven" approach is a special case of a nomination approach, per the above discussion of methods of choice. This approach would limit the number of evaluations needed to satisfy requirements for impact information that become clear as the result of the decisionmaking process in Congress or in AID. It is the type of approach Missions use in selecting projects to evaluate for effectiveness/impact, e.g., projects which have a bearing on a potential new project. It is also the type of selection approach identified by the bureaus other than PPC in discussions during the course of this contract. A decision-driven approach has no fixed number of evaluations per year.

BEST AVAILABLE COPY

Volume, in a decision-driven approach, is limited to what is needed to answer important questions in situations where the payoff can already be anticipated. It is the approach that would be likely to result in the highest information utilization pattern. It also appears to PCI to be the approach that would entail the least number of project impact evaluations, since the number of times that decisionmakers in AID stop their deliberations to secure evaluative information is quite limited at the present time. Even Congress has to have data focused on only a few types of projects at any given time.

BEST AVAILABLE COPY

Roles & Responsibilities for Impact Evaluation

CHAPTER SEVEN

The responsibility for project impact evaluations that meet AID's need to provide Congress and other external organizations, as well as its staff, with information on project benefits, their distribution and the relationship between impact and development assistance ultimately lies with the Administrator. The Administrator has a variety of options for delegating this responsibility. In Washington, it could be delegated to PPC's Office of Evaluation, to the Regional Bureaus, to technical offices in DSB and elsewhere or to AID's audit staff. In the field it could be made either a Mission level or project level responsibility.

In examining the options before AID, PCI found it useful to break down the issue of responsibility into the components that are relevant for management of a program of project impact evaluations. The four categories that appear to cover the areas of responsibility that need to be assumed are:

- Choice (of topics and projects);
- Methods (selection and certification);
- Implementation;
- Utilization of evaluation findings.

Annex A to this Chapter provides PCI's review of the theoretical possibilities for assigning these responsibilities to centralized and decentralized offices in the Agency. Annex B discusses the focus and activity that is carried out in other organizations under the heading "ex-post" or "impact" evaluation; the main example reviewed in that Annex is the World Bank

In the following paragraphs, PCI combines the data from the theoretical review of where the Administrator could place various responsibilities with other information developed in the course of this study (as well as data from other engagements with AID) to define what we judge to be a plausible and appropriate set of evaluation roles for various AID offices.

A. CHOICE OF TOPICS & PROJECTS

While AID's approach to the selection of projects for evaluation may, in the short run, be tailored to respond to a series of pressing demands, over the longer run the selection of projects for impact evaluation should respond to the Agency's need to ensure that (a) all of the major project models are examined, (b) the approach takes into account Bureau and Mission needs for decisionmaking information and (c) no important evaluation areas go unattended. As noted elsewhere, this suggests that a combination of criteria, nomination and sampling approaches be used, with varying emphasis over time.

In the early stages of impact evaluation, a combination nomination (or decision-driven) approach that is responsive to Congress and criteria selection approach is what AID appears to need. The knowledge of what types of project Congress wants AID to evaluate is, in effect, centralized knowledge. At the same time, there is a real need to demonstrate the utility of impact evaluation inside of the Agency. Knowledge of which impact evaluations would be most useful to the regional bureaus and the Missions is decentralized.

An approach that seems to offer the proper balance between centralized and decentralized selection, from the start, appears feasible. If PPC/E could prepare a list of the topics (types of projects, special concerns, etc.) that Congress is particularly interested in having AID evaluate, regional bureaus and Missions could identify specific projects in each category that, from their point of view, warrant impact evaluation. This type of approach to the short-term problem of identifying specific projects for evaluation would allow

Missions to indicate where impact information could be useful in making decisions about "repeater projects" and other elements of their programs. It offers AID a good chance to make early use of the findings from such evaluations.

In the middle years of its impact evaluation program, AID/W will want to continue to respond to Bureau/Mission interests by maintaining a decentralized element in the selection of projects for evaluation. AID will, however, still need to continue the use of criteria that help ensure that major project models are examined and that some projects are selected randomly to ensure that AID is not missing important information by trying to focus on an examination of its basic project models. The proper balance between nomination and other selection approaches will need to be worked out as AID proceeds. Ideally, AID will find that it can use criteria to select sets of projects and nominations to make choices within sets.

A combination of centralized and decentralized elements in the selection process suggests that PPC/E will need to act as a coordinator of the selection process. In the end AID may find that it needs to exceed the number of impact evaluations required to test its major project "models" if it is to be fully responsive to Missions and bureaus. PPC/E's role would be one of determining how Missions and bureau nominations fit into a "larger picture".

In the long-term, the system of program of impact evaluation should be an open one that "belongs" to the whole Agency. Eventually, the need to test project "models" will begin to decrease, and the decisionmaking focus in selection can take over almost completely. This point in time is several years away, and AID need not make a final decision now; however, it is conceivable that the bulk of selection could be decentralized, with PPC/E maintaining the responsibility to ensure that "new models", not now on line, receive proper impact evaluation consideration.

BEST AVAILABLE COPY

B. METHODS

As discussed previously, PCI recommends that AID develop standards of evidence to be used in impact evaluations. In addition it will benefit the Agency to develop "standardized" measures of impact and approaches for collecting impact data. In PCI's view there are two Agency bureaus that can make substantive contributions to this effort.

The question of standards of evidence is a centralized one. The types of standards PCI has discussed, e.g., the necessity of a comparison, the need for facts, and for proper procedures when "representative" data is required are considerations that will affect project impact evaluations no matter what type of project is involved. The standards of evidence AID uses must be accepted throughout the Agency and applied wherever impact evaluation is undertaken. In AID it is often impossible for a bureau of office to dictate a general framework for the Agency. However, when the Agency has sought to develop such frameworks, it has often done so by asking PPC to draft an approach and secure agreement to the approach from other bureaus. To PCI, this appears to be an appropriate procedure for the development of a set of evaluation evidence standards. PPC/E is the logical office to receive the assignment to develop draft standards of evidence, and to manage the process of securing bureau agreement.

The question of enforcing standards of evidence is slightly different than the question of drafting such standards. One of the problems today in envisioning a logical home for this function is the lack of requisite skills inside the Agency. In essence, the function here is a quality control function. Administration of the quality control function will require a great deal more knowledge of legitimate methodology and measurement procedures than any one AID office currently appears to have. AID's options in this area appear to include:

- The creation of a new office, with largely new staff (some from elsewhere in the Agency and some new hires) that can review and

BEST AVAILABLE COPY

judge the quality of evaluation scopes, designs, workplans, instruments, data collection and analysis plans, drafts and final reports;

- Externalizing the quality control function, either through a process similar to that used for proposal review by organizations such as the National Science Foundation (i.e., having experts outside of the Agency on retainer whose sole function is to review, judge and critique evaluation plans and report) or by finding a unit outside of AID (e.g., inside another organization) that can perform this function.

For the short and medium term, the NSF approach appears to PCI to be the best way for AID to introduce quality control. A number of the senior evaluators used by AID (through contractual and RSSA arrangements) as well as a number of the "deans" of evaluation who are attached to major U.S. universities could be drawn into an initial pool of experts who perform this function on retainer. Centralized management of this function is the most sensible approach. Thus, PPC/E would be the logical office to develop and manage such a network, and to feedback to the Agency the comments of the quality control reviewers at the evaluation planning stage as well as at the point where evaluation reports were to be disseminated.

If the concept of standards of evidence is to have meaning, PPC/E would need to have the authority to recommend, based on an external quality review of an evaluation design, that implementation not proceed if the design were found to be inadequate, i.e., not likely to produce evidence of impact, benefit distribution and/or the connection between impact and development assistance. PCI anticipates that such situations might arise. There are those in the Agency who feel that it is useful to have an "expert" come to the Mission, get a feel for the situation, and critique the Mission on a specific project. PCI recognizes that Missions value, on occasion, having someone provide such a critique. However, "evaluations" of this sort do not as a rule develop the type of factual evidence that we are recommending AID require in its impact evaluations. Enforcement of standards, in a case like this, might mean that PPC/E would suggest to those involved in the planning of such an "expert critique", that the Mission or project fund the activity, and that it not be

BEST AVAILABLE COPY

included in the Agency's listing of its impact evaluations, or endorsed as providing evidence that meets the Agency's needs for impact information.

Enforcement at the point of a final report is a different problem. In this type of case AID would presumably have carried out a quality check on the evaluation design, and endorsed it as appropriate for securing information on impact that would meet AID's standards of evidence. If an evaluation failed to pass the quality check in the draft final report stage, what AID would be saying in effect was that either (a) the evaluation team had not utilized the design it proposed or (b) the design didn't work. The responsibility PPC/E would bear at this point would be one of issuing the final report with comments. Those comments would either warn future evaluation teams about the perils of a specific design/methodology for securing data, or it would publically comment on the fact that the evaluation team had not "played by the rules" in changing its evaluation approach mid-stream. The comments of the quality review panel on this, or any, evaluation report should, in PCI's judgment, be circulated with final versions of impact evaluation reports both when they are positive and when they are negative.

The issue of measurement approaches and specific field data collection (of existing or new data) and analysis approaches is much more topic-specific than the issue of standards of evidence. The knowledge base for work in this area is in some cases quite technical. PCI anticipates that PPC/E, or any other single office, will find it difficult to stay fully aware of all of the measurement approaches that can/are being used in such fields as health, nutrition, production, participation, etc. At the same time, we are aware of the fact that a number of the projects in the Development Support Bureau have components that address just such methodological questions. These two facts suggest a need for coordination and a sharing of information between DSB and PPC/E. To the degree possible, PPC/E as a "general manager" for the impact evaluation program should be looking to DSB to help identify known measures, existing designs for field data collection on specific indicators, individuals and organizations with experience in measuring specific aspects of project performance (sometimes in connection with project design or applied research instead of evaluation *per se*), etc.

BEST AVAILABLE COPY

PCI is aware of the fact that there is currently some degree of tension between DSB and PPC/E about their respective roles in evaluation. A clarification of roles, with PPC/E taking the lead on defining standards of evidence and managing the impact evaluation process and DSB identifying specific measures and proven field data collection approaches as well as helping locate qualified individuals and institutions, should provide both bureaus the opportunity to contribute where they have comparative advantage.

In the short run, DSB could be of great assistance to the Agency effort if it pulled together the knowledge about measurement and field data collection already available in its offices. A four step activity, suggested below, might offer DSB an approach for providing this assistance:

1. In each of its offices, identify the types of impacts that are expected from each of the types of projects (technical codes) backstopped by that office;*

2. Develop articulations of the linkages between project inputs and impacts. Representations of the general case, for each type/impact combination in a rough draft "objective tree";

3. Identify from existing products/literature in the office, the measures that at the present time are viewed as being most appropriate for detecting change at each level of each objective tree;

4. Identify, using information available in the office, any known examples of field data collection approaches for securing this measurement data, and annotate each of the examples found in terms of the quality of the data collection/analysis methodology used.

C. IMPLEMENTATION

Managing the implementation of impact evaluations is a large job. It requires the sustained attention of at least one Agency officer, even if implementation is carried out by a non-AID team. If an AID team is used, the resources the

---

* The fact that there is not a perfect correspondence between Office names and primary technical codes is a problem DSB should be expected to solve. As far as PCI can detect, all of the technical codes (types of projects) are covered somewhere in DSB, and ferreting this out as well as getting internal DSB offices to work together ought to be left in the Bureaus' hands.

BEST AVAILABLE COPY

Agency must devote to implementation will increase. The tasks of preparing (or customizing a general, "standardized") scope of work, identifying a team (including all the procurement steps required if it is an outside team), review of the design, workplan, etc., and the supervision during the evaluation as well as report review take time.

PCI's rough estimate is that an AID staff member can reasonably undertake the "management"/administrative tasks for three impact evaluations at a time. At 12 manmonths is the typical length of an impact evaluation, then each AID officer would work with three impact evaluations per year; if six manmonths turns out to be the typical duration, a manager would handle six impact evaluations a year, at capacity.

In addition to the management of impact evaluations, AID is concerned with the actual source of implementation teams. As discussed previously, competent teams, particularly qualified team leaders, are expected to be a scarce resource. However, once a team leader has been selected, implementation responsibility will generally lie with him/her for the actual conduct of a specific evaluation.

In considering the assignment of implementation responsibility in AID, the primary concern is thus the locus of the responsibility for "managing" or monitoring impact, including all the preparatory steps involved in getting a team ready to begin its work. It is PCI's finding that, except in the case of evaluations that will provide information for decisionmaking about "repeater" projects, the Missions are not sufficiently interested in impact evaluation at this time to take on the role of evaluation management during implementation. Further, PCI found that no single office in AID/W has the staff complement that appears to be required to carry out the short term program of 20-30 impact evaluations over the next year. PPC/E has the biggest single staff that could do this work. In addition, there are bureau evaluation offices who have skills equivalent to those possessed by PPC/E's evaluation staff. To carry out 20 evaluations in the period of a year, at the

BEST AVAILABLE COPY

"management" rate of four per AID monitor, a total of five full-time equivalents is required. For 30 evaluations the total is 7.5.

To be realistic, AID needs to double this total, and assign only two impact evaluations per AID monitor, since virtually everyone in the Agency spends some portion of their time on other things, no matter how hard they and their offices try to keep from burdening individual officers with a diffuse set of tasks. By doubling the number of AID officers involved in managing/monitoring implementation (without including the work done by evaluation team leaders) the figure becomes 15 full-time equivalents.

For 15 full-time equivalents, in the short run, AID's best approach appears to be distribution of the responsibility around the Agency, with a centralized overview/management role placed in PPC/E. All of the regional bureau evaluation officers should be involved along with PPC/E in carrying out this set of tasks. To the degree that more staff is needed, other Agency evaluation officers, e.g., PHA, DSB, etc., should be assigned implementation management roles. The securing of time from regional and other bureaus for work on AID's agency-level effort to carry out a large number of impact evaluations must necessarily be the task of the Administrator, since important bureaucratic lines must be crossed.

By pooling its resources, in the short-term, the Agency should be able to carry the implementation management task for the program it has defined without adding staff immediately. Over the long-term, PPC/E will need a larger staff complement if it continues to carry a central role in this area. The size of that staff should be based on the projected volume of impact evaluation together with empirical data concerning the validity of PCI's estimate of one month administration for every three months of actual evaluation work. Alternatively, if AID decides now that it will eventually want to work toward a more decentralized (nomination approach) to the selection of projects for impact evaluation, staff increases in the regional bureau evaluation offices (to handle implementation) might be appropriate. In this case the PPC/E staff size would relate only to its needs to coordinate the whole program, i.e., the

BEST AVAILABLE COPY

overall selection coordination role, setting standards, identifying for DSB areas that need measures/methods work, managing the quality control process and coordinating the dissemination and use of evaluation findings in policy, program and project planning as well as Congressional presentation.

D. UTILIZATION OF EVALUATION FINDINGS

In addition to carrying out evaluations of project impact and reporting their findings to Congress, AID has a responsibility for using evaluation findings and reporting on the actions taken by the Agency and host governments to apply the lessons learned from successful/unsuccessful projects. There are two important ways in which AID should anticipate using the information from its impact evaluations. In the Missions, information on impact should be used in planning "repeater" projects for that Mission. On an agency basis, the greatest knowledge and potential for use will come when AID has (a) enough information to suggest that its project "models" will produce impacts in a new situation comparable to those it produced in a past situation, i.e., when AID knows more about the external validity of the hypotheses/assumptions on which its projects are based, and (b) information on the relative effectiveness of various project "models" in yielding specific impacts.

Use of these latter types of findings will be made above the project level, and in some instances in different Missions at or above the Bureau level. The lessons of project evaluations carried out in a specific country should have implications well beyond that country. To take advantage of what is learned, AID will need a central mechanism that both disseminates what is learned and promotes the agency-level equivalent of a "project evaluation review". In the regular project evaluation system a Mission evaluation officer brings together a meeting chaired by the Mission Director to review each evaluation and determine what actions should be taken based on the evaluation results. At an agency-level, the important participants in such a review would be individuals and offices that are involved in determining AID policies,

BEST AVAILABLE COPY

defining its programs and designing both the strategies and technical approaches used in projects. In specific terms, this means top personnel in PPC, the regional bureaus and the other central (program related) bureaus.

In terms of policy level use of impact evaluation findings, it may be possible to hold such reviews in AID/W. PPC staff and bureau staff, as well as the Administrator's office would need to be involved. PPC/E could play the role of the "evaluation officer" in setting up such reviews, or, if an evaluation was monitored by another bureau, it could delegate the role to that bureau, as long as the same group was assembled for the review.

The problem of an "evaluation review" that stimulates utilization at the program and project design level, as well as in the host government, is different, and more difficult. Project formulation, if not full design, is carried out in the Missions, as are the initial stages of program formulation. AID/W often knows that is being planned only after the first steps have been taken, i.e., through the CDSS and PIDs. Logically then, impact evaluations will need to be reviewed in the Missions and "actions" identified there, if these evaluations are to be utilized at this level. Because AID/W is often unaware of what Missions are planning, it will find it hard to identify which evaluations should be reviewed by which Missions. In effect the Missions will need to have access to all of the impact evaluations, and make their own decisions about which ones are important enough to warrant Mission review.

Mission reviews of AID impact evaluations could probably be carried out in much the same way that evaluations conducted by the Mission are reviewed, e.g., the existing procedures for setting up an evaluation review should be adequate. The recording of how AID uses of plans to use impact evaluation findings could at this level be the responsibility of the Mission evaluation officer. A summary of what the Mission sees as action implications could be forwarded to AID/W following a review. Monitoring of whether these actions are followed, it appears, could be done in Washington, via review of new PIDs and CDSSs.

BEST AVAILABLE COPY

AID cannot guarantee host utilization of impact evaluations findings, but it can involve them in the Mission level process it carries out to review impact evaluations and identify the actions those evaluations suggest. Missions, in the course of recording what they feel is actionable based on an impact evaluation, could also be asked to record whether host representatives attended the review and to note what action implications they identified.

Mission comments on impact evaluations should receive wide circulation in AID/W since these reviews are likely to generate action items that require the involvement of AID's central bureaus, as well as the regional bureaus. In addition, PCI sees a potential value to circulating the collected action comments on each impact evaluation to all Missions. Ideas from one Mission will stimulate thinking elsewhere, and may yield additional utilization benefits.

PPC/E, as general manager for the program of impact evaluations, should have the responsibility for collecting and collating comments from the AID/W and Mission evaluation reviews and circulating the compiled comments back through the Agency, i.e., to all AID/W offices that are involved as well as to the Missions.

BEST AVAILABLE COPY

THE OPTIONS FOR LOCATING A SERIES OF EVALUATIONS RESPONSIBILITIES

This Annex reviews the Administrator's options with respect to the assignment of responsibility for:

- Choosing which projects are to be evaluated;
- Selecting appropriate impact measures and evaluation methods;
- Implementing project impact evaluations;
- Monitoring and facilitating the utilization of evaluation findings.

1. Choice

Even if AID were to opt eventually for universal coverage, it would face the immediate problem of deciding which projects were to be evaluated in the first, second, and third years of AID's program of impact evaluation.

Three basic alternatives for selecting projects were identified:

- Nomination (a "volunteer" type of selection system);
- Criteria selection using one or a number of criteria options;
- Random selection.

Each of these three methods suggests a different approach for assigning selection responsibility.

BEST AVAILABLE COPY

a. Nomination

A volunteer or nomination approach could involve everyone in the Agency. Under such a selection system, a central office would need to be designated which received nominations. The responsibility for recording project names and other data such as date volunteered, project location, size, start and end dates, objectives, etc., would fall to this office. But the actual selections would be made by those who nominated specific projects for impact evaluation. Nominations could be made by field staff or by offices that saw a value in having information on a specific geographic area to type of project; the "decision-driven" approach is the same.

b. Criteria Selection

There are a variety of selection criteria that AID could use in setting up a selection system.

Responsibility for the management of a criteria selection approach can be divided into two steps: Definition of the criteria and selection of projects using the criteria. Since the criteria would be applied to all projects in the Agency, the responsibility for defining the criteria would necessarily be a central one. This central responsibility could be given to a single office, or it could be made the province of a special committee of the Agency. The responsibility for selecting projects using the Agency-wide criteria could be delegated in several ways. It could be a central responsibility (given to an office or a committee) or it could be decentralized. If this latter responsibility were decentralized, it could be placed at the Bureau level or at the Mission level. It could not however be placed at the project level.

c. Random Selection

The locus of responsibility for management of a random selection procedure would necessarily be a central office or committee if the procedure were to

BEST AVAILABLE COPY

be applied across AID s entire portfolio. If, on the other hand, the portfolio were stratified into geographic units such as regions or countries, then it could be administered by the Regional Bureaus or Missions, respectively.

2. Methods

As indicated elsewhere, AID has two basic options in this area. One involves monitoring quality by overseeing the methodologies used in project impact evaluations. The second allows methodology to vary widely, and concentrates AID's quality control effort on the selection of the people who carry out evaluations. PCI recommended the first of these two options because it was more likely to yield good results in terms of solid evidence of impact and because it takes advantage of AID's opportunity to keep both costs and errors down by standardizing, to the degree possible, both the measure of impact and the procedures used to gather impact data.

Whether AID follows PCI's recommendations for achieving quality through methodology controls, or elects instead an approach that tries to achieve uniform quality by monitoring evaluation personnel, it will still need to decide where to place the responsibility for the selection of methods/people and quality control. There are two elements to this task: (a) setting the standards, i.e., defining the measures/methods or the qualifications of personnel who can carry out AID's evaluations, and (b) administering those standards.

The first of these responsibilities will necessarily be a central responibility that must be carried out by an office or committee. Only by centralizing responsibility for this task can AID be assured that all of the project evaluations, worldwide, will have a roughly common level of quality. While this is a task that must be carried out centrally, it is not likely that any single office or a small committee will have all the knowledge required to decide what measures/methods of personnel qualifications are appropriate for evaluations of projects in the wide range of sectors and subsectors treated by AID.

BEST AVAILABLE COPY

At the level of impact indicators, measures and general methods, an office or committee may have the needed expertise, but it will need help when it comes to specifying exactly how to measure or pre-qualify personnel to measure such elements of a macro-impact like "basis needs" as adequate shelter, caloric intake, etc.

To move from the level of specification provided by this report to a final set of measures/methods or pre-qualifying specifications for evaluation personnel a central office or committee charged with the responsibility for this task is likely to need the help of either a series of topical subcommittees inside AID or an outside organization familiar with measurement/methods and the broad topical range the AID program encompasses.

The second responsibility, that of applying the standards set for methods or people by a central office or committee, can be either a centralized or decentralized responsibility. Here again AID has the option of using a central office or committee as the responsible agent for ensuring that AID's standards are employed. It also has the option of making the responsibility for applying the standards a bureau level or Mission responsibility.

If the responsibility were centralized, the office or committee involved would need to examine each and every project impact evaluation effort. If measures and methods were the quality control mechanism, then this office/committee would need to review, at minimum, the scope of work, the evaluator's design and workplan, the raw data and preliminary analysis, and the draft and final reports. (An option here would be to have the office/committee receive assistance in this detailed task from some permanent quality control group. Several options exist including an internal group, a contractor, PASA or RSSA charged with this review task, or an approach that involves a series of external reviewers, such as is common in proposal reviews in such organizations as the National Science Foundation.) If on the other hand, AID's quality control point was personnel, then under a centralized scheme the office/committee involved would need to review the credentials of proposed evaluators against both the general criteria and the specific scope of work.

BEST AVAILABLE COPY

On a decentralized basis, these tasks could be carried out by the bureaus or the Missions. However, such assistance with quality control as external reviews by evaluation specialists around the U.S. would be difficult to manage on a decentralized basis. The other options for quality reviews if measures/methods are the focus, such as an internal AID quality control group, a contractor, PASA or RSSA are still possible. But if Missions were the locus of responsibility there would be delays introduced in the process if such services were used. If AID used a personnel based method of controlling quality, it would not be more difficult or take longer if the responsibility were placed at the bureau of Mission level.

3. Implementation

There are several subtasks to be considered in this category including the preparation of scope of work (or customization of a general scope, if AID adopts PCI's recommendation in this area), selection of an evaluation team (including managing a procurement if needed), monitoring each main step in the evaluation, troubleshooting, monitoring expenditures on the evaluation over the life of the effort, and finally, reviewing, accepting and disseminating the report.

AID has the option of keeping all of these tasks together and assigning responsibility for the cluster to a central/office committee, to bureaus, to Missions, or to the projects that are to be evaluated. (In some cases, this would involve naming a person in the Agency to represent the project if for example it had terminated or AID no longer provided assistance in the country involved.)

The other option is to split up the subtasks, perhaps as follows:

- Scope of work preparation;
- Selection of the team;
- Monitoring substance, expenditures and troubleshooting;
- Report review, accpetance and dissemination.

BEST AVAILABLE COPY

Again, any of the options for locus of responsibility identified above could also be the locus of responsibility for a subtask.

Implementation responsibilities are likely to be the most time consuming of the four types of responsibilities identified. Thus, selection of a locus of responsibility may need to relate more to the availability of staff who can handle the technical aspects of implementation and travel with evaluation teams when needed, than to other factors. In part, this aspect of implementation argues for dividing the responsibility into subtasks, based on expertise, e.g., AID/W personnel who have dealt with a large number of evaluations already may be better prepared to draft/customize scopes of work and manage a team selection process than field personnel. On the other hand, field personnel may be in a better position to monitor the milestones in a workplan, review progress against expenditures and troubleshoot as needed, during the field portion of project impact evaluations. Reports, if they are to pass through a central of bureau level quality check before being accepted/disseminated, would more logically become a central responsibility.

4. Use

Once an evaluation has been carried out, and the report has been accepted and disseminated, AID has the option of taking the extra steps required to translate evaluation findings into policy, program and project level changes. In its Project Evaluation System, AID requires that evaluation findings from formative evaluation be translated into changes in the project activity to the degree needed. It has had a good deal of success with this approach. It does not place the same requirement on those who initiate "special evaluations", and the record here appears to be mixed in terms of the degree to wh. evaluation findings lead to changes in policies, programs and/or projects.

In its documents and in the interviews PCI conducted, AID indicated that it expected to be the main beneficiary of evaluations findings. It even recommended additions to a minimum scope for project impact evaluations that would

BEST AVAILABLE COPY

help the Agency use evaluation to improve policies, programs and projects. If AID is to accomplish this objective, it appears that a special responsibility will need to be assigned for ensuring that evaluation findings are used (or if not used, that proper consideration was given to using them). This final responsibility corresponds to the task AID outlined in its memorandum of June 25, 1979 on information for the Congressional presentations. That memo cited the need for AID to assess whether it and the host countries made use of the lessons learned from successful/unsuccessful projects. In PCI's discussion of scope, we suggested that this was not an element of the scope of an individual evaluation, but rather part of the larger management task. The identification of a locus of responsibility for the use of project impact evaluations is, in our judgment, the appropriate point to reintroduce this aspect of AID's work in preparing for future Congressional presentations, as well as fulfilling its prophecy that the Agency will benefit from project impact evaluations.

There are several subtasks involved in ensuring that evaluation findings are translated into action, including monitoring dissemination to ensure that the people who can make changes are informed, monitoring the form and substance of what is disseminated (e.g., calling for executive summaries in situations where sheer bulk will be likely to deter use, etc.), bringing together the parties who can react to evaluations and act on them (i.e., calling evaluation review meetings in AID at the policy, program and project level as appropriate and engineering a process that makes the same sort of meeting occur in the developing countries), capturing on paper the decisions made at those meetings and identifying who is in charge of carrying out the actions that follow from those decisions, and finally, monitoring whether the responsible individuals carry out the actions they have been assigned (and learning what results come from those actions).

This task, like many others, could be assigned at a number of levels in AID. It could be a Washington or a field responsibility in the project and program areas. It must be a field responsibility from the host action perspective. And it can only be a central responsibility in the area of AID policy, when a worldwide policy is involved.

BEST AVAILABLE COPY

LESSONS FROM THE EXPERIENCE OF OTHER ORGANIZATIONS

PCI's review of the evaluation systems of federal agencies that manage domestic programs and organizations that, like AID, manage development assistance programs has indicated that there is no model of an "ex-post" or "impact" evaluation system in existence that systematically provides the type of impact information that AID is seeking.

The U.S. Department of Health, Education and Welfare and its associated institutes have the best experience with impact evaluation on the domestic scene. They have found impact evaluations to be difficult and expensive. Ex-post evaluation is not a preferred approach because it too often involves post-measurement without adequate baseline data. The problem of attribution with both ex-post and impact studies is a major one.

The World Bank has an ex-post evaluation system, but that system is an extension of the organization's audit function, rather than being an impact measurement system. Nevertheless, this system has certain management features that have potential applicability in AID. The World Bank has also begun work in the area of impact evaluation. It's first formal impact study was released in June, 1979. This study, however, is a single effort and like the individual impact evaluations undertaken by other organizations, it is a resource to AID. The World Bank appears to be considering the development of a systematic program of impact evaluation, but it has not yet defined the coverage of such a system or the way in which it would operate.

The World Bank Evaluation System reflects a management style composed of the following elements:

- Centralized review to ensure objectivity;
- A post-project audit system that focuses on lessons to be learned for future Bank operations in the same country and sector;

BEST AVAILABLE COPY

- Regard for impact evaluations as research rather than as an integral part of the management system. Concomitantly, the Bank endorses the importance of impact evaluations, but performs them sporadically.

With regard to substance, the Bank's Evaluation System concentrates on the following:

- A comparison of each project's results with the original project objectives;
- Reestimation of each project's financial and economic rates of return.

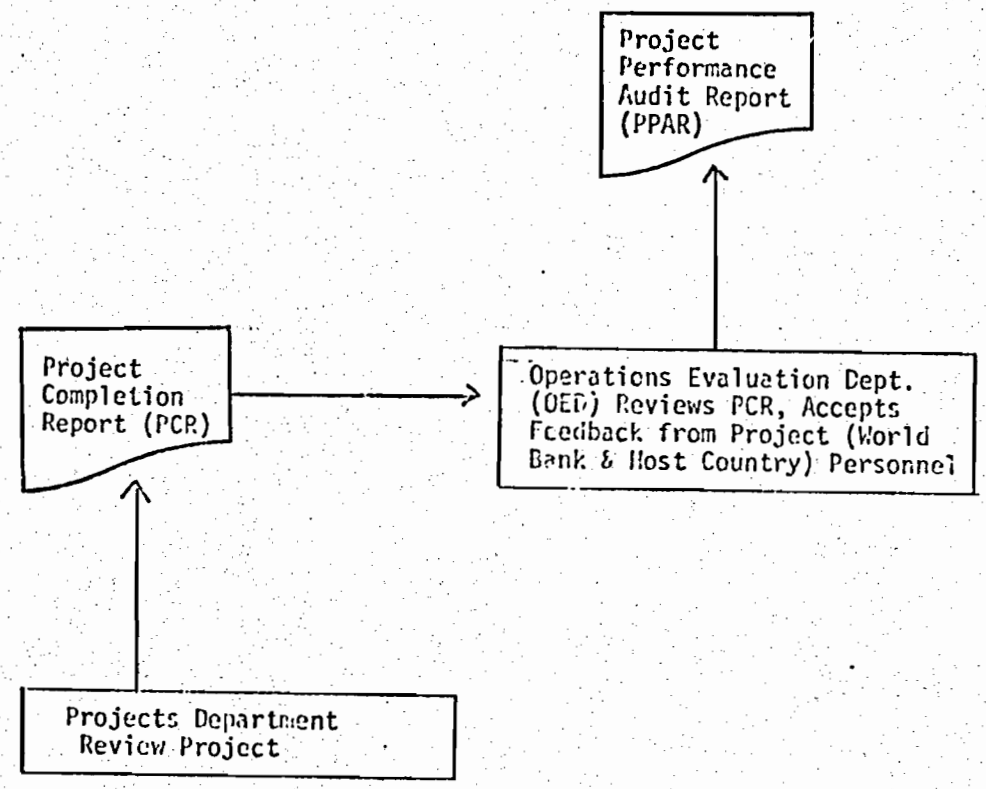
The Operations Evaluation Department (OED) is the central unit responsible for the Bank's Evaluation System. However, the various Projects Departments concerned with implementation of projects are also part of the evaluation process. The Projects Departments produce Project Completion Reports (PCRs) at the end of every loan project. Approximately 12 months later, the OED publishes a Project Performance Appraisal Report (PPAR). The PPAR is typically composed of an audit memorandum and the PCR. The audit memorandum assesses the PCR for objectivity and comprehensiveness. The OED staff base this assessment upon a review of project documents and interviews with Bank staff. In some cases the evaluator will travel to the host country and interview the borrowers. A typical PPAR requires four person-weeks of work.

In all cases both World Bank project staff and the borrowers comment on the PPAR while it is in the draft stage. The PPAR is then delivered directly to the Executive Board of the Bank. The report may include comments on the design and size of the project, technical and institutional issues, and whether the timing of the project was appropriate. A flow diagram of the PCR/PPAR process is displayed in Figure VI.B.1.

BEST AVAILABLE COPY

FIGURE VII.B.-1

WORLD BANK PROJECT EVALUATION PROCESS



BEST AVAILABLE COPY

The Project Completion Report and the Project Performance Appraisal Report are frequently critical of the project in question. The issues most frequently raised are:

- Cost overruns;
- Delays in project implementation;
- Design problems.

These issues have been consistently identified in the Bank's periodic summary of project evaluations: Annual Review of Project Performance Audit Results. As stated in the fourth review:

> "The purpose of this Review is to summarize the experience with these 109 projects in a way that will indicate the principal characteristics of their performance, note any recurring patterns, and point to aspects which merit attention in the design and implementation of future projects."*

The OED produces two other types of studies. They are:

- *Evaluation Studies:* To examine the impact of sets of projects by country or sector;
- *Operational Policy Reviews:* To examine the application of specific policies and procedures governing the management and administration of the Bank's programs.

The Bank's total operations in one country over a 20-year period were evaluated in 1972. Its findings caused a controversy among some Bank officials and host country personnel. No more such evaluation studies are planned.

In summary, the World Bank has established an evaluation system geared toward learning from experience and improving the design of future projects. However, the system in no way claims to measure project impact or the distribution

---

* World Bank, Annual Review of Project Performance Audit Results (World Bank: Washington, D.C., November, 1978).

BEST AVAILABLE COPY

effects of projects. The system provides project personnel and borrowers an opportunity to express opinions about the project. It could possibly serve as a foundation for an impact evaluation system.

BEST AVAILABLE COPY

BEST AVAILABLE COPY

The Place of Impact Evaluation in AID's Evaluation System

CHAPTER EIGHT

In this chapter, PCI addresses a number of characteristics of impact evaluations for mature and terminiated projects in the context of AID's existing systems for evaluation, design, manpower planning and staffing, and the funding of analytical activities.

THE RELATIONSHIP OF IMPACT EVALUATION TO THE PES

There is a sufficient overlap of concepts and intent between impact evaluation and the existing project evaluation system to suggest that impact evaluation should be a part of this system, rather than a totally new, free-standing system. AID's need for impact evaluation differs from a Mission or project's needs for information in several ways, including a difference in the level of interest in detailed management problems and immediate remedial actions, questions that pertain to the external validity of a project's hypotheses, the secondary effects of projects, etc. In addition, there is a different logical locus for decisionmaking based on evaluation. In regular project evaluations, the project is the level at which most post-evaluation decisions are made. With impact evaluation, this will not be the case. Because of its potential relevance for many projects, the logical home for evaluation reviews of impact evaluations will be Washington and will involve representatives from bureaus that could potentially apply the study's findings. (In addition, it may well be the case that with Mission assistance host countries will want to review the findings of impact evaluations parallel to AID's central review.)

BEST AVAILABLE COPY

The differences are differences in users and to some degree in the substantive empahsis. The two types of evaluation, formative (management) and substantive (impact) share the same basic evaluation framework: The Logical Framework for a project that sets forth objectives, indicators/targets and project assumptions. PCI's examination of AID projects indicated that at present, most Logical Frameworks state the mandate objectives of projects, making it still more likely that AID can continue to use the same conceptual basis for evaluation that has proven useful at the project (management evaluation) level. While PCI found no need for a fundamental change, two activities would help strengthen AID's ability to use the project's Logical Framework as the basis for evaluation:

- A tighter review of the Logical Framework at the time of design. Many are still poorly articulated with only a passing attempt to identify targets, indicators, assumptions and potential means of verification at the Purpose and Goal levels;

- Encouraging projects to articulate all of the key levels of their logic when this activity would:

  - Reduce the tendency to put two logical levels in one matrix cell;

  - Clarify the linkages between Outputs and mandate objectives.

  For most projects the number of additional levels needed would likely be zero to two, with a few projects using more.

At present there are no guidance materials or regulations (published) about evaluation that are likely to conflict with AID plans for ex-post evaluation, or the recommendations herein concerning impact evaluation. AID's published Handbook 3 material on evaluation would, however, need revision if AID modifies the existing system to make room for an impact element managed above the project level.

To the Agency, the most striking difference between project level evaluations and impact evaluations will be the involvement of a different set of actors. As Chapter 6 has suggested, PCI expects that impact evaluation will operate best if managed from AID/W. The level of interest and the availability of

BEST AVAILABLE COPY

time/skills at the project or Mission level suggests that placement of major responsibility for impact evaluation overseas would be dysfunctional. On the other hand, placement of these responsibilities and roles in AID/W will create a new dynamic. AID will need to be clear in its communications about impact evaluation and note for the field that:

- Responsibility is in Washington because the need for information and the primary user interest is there;
- Missions and host countries are expected to benefit; these evaluations are not attempts to look over the Mission's shoulder, but rather to learn from the past for the benefit of all;
- Mission and host involvement is welcomed by AID whenever and wherever either of these actors has the time and interest to participate actively in an impact evaluation.

APPROACHES TO FUNDING IMPACT EVALUATIONS

AID has several options for funding mechanisms, including placing the funds for impact evaluation in project budgets. During the course of this study PCI located examples of project budgets that included funds for evaluation after TTD, thus suggesting that at least a few offices have found ways to work out the details of post-project evaluation via project funds. We do not however see a need to include impact evaluation funds in all project budgets. PCI is recommending that AID use methodologies that can be implemented on an "after-only" basis (cross-sectional studies) wherever possible, and that AID adopt one of a number of partial coverage approaches. Thus, while it may occasionally anticipate that a large or experimental project will require impact evaluation, and include the funds in the project budget, this should not be AID's normal funding mechanism. It would, in the end, we believe, lead to some funding of studies that are never done, and a lack of flexibility to choose, based on future developments, which projects to evaluate.

The basic alternative to project funding is a pool of funds set aside for evaluation, either centrally or allocated by bureau. In the early years, a

BEST AVAILABLE COPY

totally central pool of funds that all bureaus could draw on may make the most sense. Later, when a pattern of impact evaluation practice establishes itself, it may make sense to attach the funds to decentralized units. If AID uses a centralized pool of funds at the start, it may be important, psychologically, to make it an Administrators' fund, rather than a resource pool that seems to be under the control of any single bureau. The procedures and roles with respect to impact evaluation will be assigned to bureaus, and naturally these bureaus will spend the pool of funds AID sets aside. But the funds themselves should belong to the whole community.

STAFFING IMPACT EVALUATIONS

Impact evaluations will compete for scarce resources in the Agency. Money is not the scarcest commodity in this situation; it is people who can devote time and attention to impact evaluations. AID has few personnel with the background and interest required to play leading roles in the Agency's impact evaluation work. Further, the talents of these individuals tend to be wanted in many places, including PPC/E/S's Studies Division, in DSP, for Mission work on project design and evaluation, as trainers, as managers of evaluation task orders, in DSB as technical specialists for subject areas, etc. The same situation applies to AID's pool of RSSA (particularly the Bureau of the Census) and contractor personnel. To some degree it also applies to universities and to PASA arrangements such as those with USDA and DOT. If the RSSA with the Bureau of the Census can be taken as an example, it has been reported to PCI that various offices in AID feel they are "standing in line" to get attention/service on their existing needs. Impact evaluation will create a whole new set of needs for the talents of individuals both in AID and in the organizations with which it works.

Given the fact that talent is likely to be the scarce resource, AID has a number of options. It can try to upgrade the capabilities of its staff who do not now but could play major roles in evaluations. It can try to have more individuals added to the RSSA and PASA arrangements it uses frequently for

BEST AVAILABLE COPY

evaluation assistance. It can hire additional personnel, and it can issue more RFPs. (AID can probably assume that the universities and private evaluation research groups will respond to an increased level of evaluation activity by adding appropriate staff to their organizations.)

In the past AID has tended to set up evaluations so that only one of its human resource groups was involved. Few instances of mixed teams, e.g., AID and university, RSSA and contractor, have beem tried. In beginning its impact work, AID may find that some effort to create mixed teams, each of which involved one or more AID staff members, could work to increase the Agency's capacity to carry out many of the steps in impact evaluation itself. At present there are no easy ways to make this occur. New contracting/funding methods would need to be worked out. A task order approach which allowed AID to pull together individuals from various groups and let them work together on an evaluation is theoretically possible. It would involve some type of "retainer" arrangement with each group, but not one that has the type of tight time frame found in the current IQCs. Impact evaluations should not be force-fit into a 90 or 120 day period where this is not appropriate.

In addition to using the evaluation resources now available to the Agency, AID should seek to involve new faces in the work, particularly those who have undertaken significant efforts to carry out impact evaluations in other contexts. AID should be trying to secure the advice of the "deans" of evaluation, and also seeking from them references to highly qualified former students and associates who would welcome involvement in AID's impact evaluation efforts on a short- or long-term basis. Similarly, in issuing RFPs, AID may want to encourage new groups by placing "points" on impact evaluation experience and temporarily reducing the "point" empahsis on overseas experience.

By whatever combination of approaches AID elects to secure human resources for its impact evaluation work, it should be also trying to deepen the pool of people who can do this work professionally. Inside AID, PCI views the main mechanism for increasing skills as participation on professional evaluation teams. AID's staff will learn more faster from being part of the teams it builds than from any training course that could be devised.

BEST AVAILABLE COPY

BEST AVAILABLE COPY

The Cost of Impact Evaluation

CHAPTER NINE

In this Chapter, PCI provides AID with estimates of the level of effort, stated in terms of direct labor and other direct costs, which the Agency can expect to incur in connection with its program of impact evaluations and related project design improvement efforts. The Chapter provides estimates based on a variety of different assumptions, and it gives AID the basic data and framework for making additional estimates. The specific areas in which we examine cost options are:

- Different assumptions about the number of evaluations AID carries out per year. Estimates are for 10, 20, 30 and 50 impact evaluations. As noted elsewhere, PCI recommends that AID start at or near the low end of this range;

- Differences in the number and type of questions included in the scope of work for impact evaluations. As noted elsewhere, PCI recommends that AID concentrate on the "basic" scope in the first year(s), until it has demonstrated that the evidence needed can be secured with regularity. At the same time, in at least a few evaluations AID should include at least one supplementary scope item, in order to validate our early estimates and examine the difficulties it will face the securing evidence on these additional questions;

- Differences in data collection approaches. As noted elsewhere, PCI expects that evaluation teams will be able to secure evidence of impact that meets AID standards in a number of ways. To the degree possible, AID should try to ensure that all pof the basic approaches are tried early. For estimating purposes, we assumed a mix: Equal frequency of four basic data collection models;

- Differences in the responsibilities AID/W takes on as impact evaluation program "manager". We present a set of minimal tasks, and a set of supplementary tasks. Of the supplementary tasks, we consider the two "quality control" reviews for each project to be of considerable importance. In PCI's judgment they are

BEST AVAILABLE COPY

essential for a high-quality impact evaluation program. Other items in the list of supplementary management tasks are best started the first year as well, but several could be deferred;

- Differences in the mix of personnel used on evaluation teams, including such options as all AID teams, teams made up of non-AID, U.S. personnel (contractors, RSSAs, PASAs, etc.) and combinations of U.S. and LDC personnel. PCI recommends that AID experiment with all of the evaluation team options it has available as it seeks to establish a workable program of impact evaluation.

In the estimates of level of effort, PCI has omitted two considerations that will at some point need to be examined:

- Inflation, which will effect future program years;
- Economies, which we expect may be realized when:
  - AID's experience increases, generally;
  - Designs/instruments are resued;
  - Knowledge concerning specific measurement approaches is consolidated;
  - etc.

  While PCI expects economies of this sort, we do not feel that it it prudent to estimate as if we had a high degree of certainty about which economies AID would in fact realize.

The presentation of costs below is divided into three parts: (a) Costs associated with specific evaluations, (b) costs associated with the management of a program of impact evaluations and (c) additional tasks that should be done at the project design stage (and charged to that function/system).

A. THE COST OF PROJECT IMPACT EVALUATIONS

In this section, the cost of studies that use the "basic" scope are examined first. Supplementary scope items are costed separately at the end of the section.

BEST AVAILABLE COPY

PCI anticipates that AID will find itself involved in four quite different types of situations with respect to the collection and analysis of field data for impact evaluations:

1. Analysis of Existing Data

   There are many situations in which AID can take advantage of records collected for purposes other than impact evaluation. This situation occurs more often than is recognized. Too often new data is collected when, if a little effort was made to identify existing data, the collection of new primary data could have been avoided. For these situations, the costs are the costs of locating existing data, examining the manner in which it was collected to ensure that it meets professional standards, and reanalysis of the data to answer impact evaluation questions.

2. Case Study Data Collection & Analysis

   There will be situations in which the most productive approach involves case studies in locations that did/did not receive project assistance.

3. Sample Survey Data Collection/Analysis Using an Existing Sample Frame

   There will be situations in which a sample frame (area or population) exists, and where the most productive approach will involve a sample survey of populations/areas that did/did not receive project assistance.

4. Sample Surveys where the Sample Frame must be Developed

   There will be situations in which a sample frame (area or population) will need to be created before a sample survey can be carried out.

There is no easy way to determine, *a priori*, how many times each situation will occur. For this set of calculations, PCI has assumed an equal frequency distribution. Table VIII-1 identifies the direct labor and other direct cost elements of each of these types of studies. Only manmonth estimates of labor are provided on this table. Table VIII-2 shows the labor cost together with other direct costs under several assumptions, including:

- Teams composed wholely of AID/W personnel;

BEST AVAILABLE COPY

TABLE IX-1:

DIRECT LABOR & OTHER DIRECT COST COMPONENTS OF EACH TYPE OF STUDY (ESTIMATE APPLIES ONLY TO "BASIC" SCOPE)

	Manmonths & Other Direct Cost Estimates by Study Characteristics
1. Analysis of Existing Data	
Direct Labor:	
Locate and examine existing data	1.0
Analysis of existing data to answer impact evaluation questions	5.0
	6.0
Other Direct Costs (transport, computer, etc.):	$8,000.00
2. Case Studies	
Direct Labor:	
Preparatory work (plan, locate informants)	1.0
Data collection	2.0*
Analysis and presentation	2.0
	5.0
Other Direct Costs (transport, etc.):	$5,000.00
3. Sample Surveys with Existing Frame	
Direct Labor:	
Preparatory work (design, instruments, interviewer training, pretests, analysis plan)	2.0
Data collection (including fieldworker supervision and call-backs)	6.0**
Analysis and presentation	5.0
	13.0
Other Direct Costs (transport, commuter):	$15,000.00
4. Sample Surveys with Frame Development	
Direct Costs:	
Frame development	3.0
Regular sample survey labor	13.0
	16.0
Other Direct Costs:	$20,000.00
Overall Study Cost Averages	
Direct Labor	10.0 (with equal mix)
Other Direct Costs:	$12,000.00 (with equal mix)

* Assumes two communities, or two institutions (one that did receive project assistance and another that did not.

** Assumes either eight short interviews/measurement sessions per day with a study population (treatment and comparison groups) of 640 units (eight strata, with a cell size of 40) *or* four longer interviews per day and a study population of 320 (four strata, cell sizes of 40).

BEST AVAILABLE COPY

TABLE IX-2

Type of Study	Composition of Team: AID/W Staff Only	AID/W & Local	Non-AID U.S. Only	Non-AID U.S. & Local	Average with Equal Mix of Teams
Analysis of Existing Data					
Total MM = 6.0	$ 37,500		$ 49,998		$ 43,749
Local Option = 0 MM	-		-		-
Other Direct Costs = $8,000	8,000		8,000		8,000
	45,500	45,500	57,998	57,998	51,749
Case Studies					
Total MM = 5.0	31,250	18,750	41,665	24,999	29,166
Local Option = 2.0	-	2,500	-	2,500	1,250
Other Direct Costs = $5,000	5,000	5,000	5,000	5,000	5,000
	36,250	26,250	46,665	32,499	35,416
Sample Survey/Existing Frame					
Total MM = 13.0	81,250	53,125	108,329	70,830	78,383
Local Option = 4.5	-	5,625	-	5,625	1,406
Other Direct Costs = 15,000	15,000	15,000	15,000	15,000	15,000
	96,250	73,350	123,329	91,455	94,789
Sample Survey/Develop Frame					
Total MM = 16.0	100,000	65,625	133,328	87,496	96,612
Local Option = 5.5	-	6,875	-	6,685	1,718
Other Direct Costs = 20,000	20,000	20,000	20,000	20,000	20,000
	120,000	92,500	153,328	114,371	118,330
AVERAGE WITH EQUAL MIX OF STUDY TYPES (TOTAL):	$ 74,500	$59,400	$ 95,330	$ 74,080	$ 75,499*

* This final figure averages the row and column totals of $75,827 and $75,071; the error is a function of rounding off liberally, at various points within the table.

BEST AVAILABLE COPY

- Teams composed of AID/W and local hire staff;
- Teams composed of non-AID, U.S. personnel only; and
- Teams of non-AID, U.S. staff with local hire assistance.

Note that Table VIII-2 does not include secretarial time associated with the conduct of evaluations, The costs of clerical assistance would vary widely depending on whether it was procured here or overseas. In reviewing the table, the absence of this figure should be remembered.

Table VIII-3 identifies the cost of the various supplementary items that AID may elect to add to its basic evaluation scope. PCI recommends that in the first years AID experiment with both team composition and scope elements, as well as examine the utility of each of the four study types PCI anticipates will be used.

B. COSTS ASSOCIATED WITH MANAGEMENT OF THE IMPACT EVALUATION PROGRAM

The costs of managing the program are estimated for several levels of effort, including programs involving 10, 20, 30 or 50 evaluations per year. Recommended supplementary tasks are separated from what appears to be a basic minimum set of management activities. Table VIII-4 presents PCI's estimates of the Year One costs; Table VIII-5 presents the Year 2...*n* cost estimate for these AID management activities.

A summary of probable high and low costs for the full effort, at a level of 20 studies per year, is presented as Table VIII-6. It was also presented in Volume I. The high costs identified here are not AID's maximum costs. Below, a maximum cost scenario is described:

> Under the assumptions that (a) virtually all studies turn out to be sample surveys, where a sample frame must be developed, (b) AID uses all non-AID U.S. personnel in impact evaluation work, (c) supplementary scope items for secondary data and benefit continuation use a total of 4.0 manmonths, and (d) the final supplement item turns out

BEST AVAILABLE COPY

TABLE IX-3:

DIRECT LABOR & OTHER DIRECT COSTS FOR "SCOPE SUPPLEMENTS" IN EVALUATIONS

	Manmonths & Other Direct Costs by Scope Supplement Element
1. Examination of Secondary Effects, Including Spread Effects: This estimate is based on a presumption that an evaluation team would identify or establish that a secondary, spread, unplanned or negative effect had occurred. Evidence would be required, however. The evidence gathered by the evaluation team would not be of a type that would document the extent, pattern of dsitribution or provide other detailed measures of such efforts.	
Direct Labor:	2.0 to m
Other Direct Costs:	$2,000, if done while team is in the field.
2. Estimation of Sustainability/Benefit Continuation: This estimate assumes that minimal data would be collected that could be used with models such as PCI's MIM (organizational viability) or DAI's (benefit continuation in cooperatives) approaches to predicting sustainability. As models improve, and AID should try to ensure that such models as it initially uses do improve, the costs will change.	
Direct Labor:	2.0 to m
Other Direct Costs:	$2,000, if done while team is in the field.
3. Followup Studies Based on Impact Evaluations that secure data on second generation effects and (utilization) benefit continuation/sustainability field work that lags impact evaluation by 18-36 months (per recommendations of impact evaluations) to secure data that cannot be secured during the impact evaluation.	
Direct Labor: 3-13 manmonths (while a frame will not again need to be developed, some situations may require two full survey rounds to establish trends)	3.0 - 13.0
Other Direct Costs:	$3,000 - 15,000

BEST AVAILABLE COPY

TABLE IX-4:

IMPACT EVALUATION PROGRAM MANAGEMENT COSTS -- YEAR 1 (LABOR ONLY)

TASK DESCRIPTION	10 Studies		20 Studies		30 Studies		50 Studies	
	MM	$	MM	$	MM	$	MM	$
A. MINIMUM SET OF TASKS								
A.1 Selection of Topics	1.0	6,250	1.0	6,250	1.5	9,375	2.0	12,500
A.2 Selection of Specific Projects	1.0	6,250	1.0	6,250	1.5	9,375	2.0	12,500
A.3 Preparation of Budget, etc.	1.0	6,250	1.0	6,250	1.0	6,250	1.0	6,250
A.4 Arrangements for Teams (including procurement for 1/2 the studies)	3.0	18,750	6.0	37,500	9.0	56,250	15.0	93,750
A.5 Briefings/Startup with Team	1.0	6,250	2.0	12,500	3.0	18,750	5.0	31,250
A.6 Monitoring, Logistics, Troubleshooting during the Study Period	1.5	9,375	3.0	18,750	4.5	28,125	7.5	46,875
A.7 Report Dissemination, Evaluation Reviews, Utilization Monitoring	1.5	9,375	3.0	18,750	4.5	28,125	7.5	46,875
SUBTOTAL (for tasks done by PPC, regional bureaus, contracting) i.e., *professional time**	10.0	62,500	17.0	106,250	25.0	156,250	40.0	250,000
Support Personnel Time/Dollars**	5.0	10,415	8.5	17,705	12.5	26,037	20.0	41,660
	15.0	72,915	25.5	123,955	37.5	182,287	60.0	291,660
B. RECOMMENDED SUPPLEMENTARY TASKS								
B.1 Set Standards of Evidence	1.0	6,250	1.0	6,250	1.0	6,250	1.0	6,250
B.2 Develop "Standardized" Scope	1.0	6,250	1.0	6,250	1.0	6,250	1.0	6,250
B.3 Identify Basic Evaluation Designs that will Yield Adequate Evidence	1.0	6,250	1.0	6,250	1.0	6,250	1.0	6,250
B.4 Identify Measures/Field Data Collection/Analysis Techniques that have Proven Record for Producing Evidence of Specific Impacts, e.g., Income, Health Status Changes, etc.	3.0	18,750	3.0	18,750	3.0	18,750	3.0	18,750
B.5 Redraft/Improve New Handbook Regulations on Evaluation	1.0	6,250	1.0	6,250	1.0	6,250	1.0	6,250
B.6 Draft/Disseminate Impact Evaluation Guidance Based on B.1 - B.4	3.0	18,750	3.0	18,750	3.0	18,750	3.0	18,750
B.7 Outside "Quality Control" Review of Evaluation Plans for each Project Evaluation (Detailed Material)	1.5	9,375	3.0	18,750	4.5	28,125	7.5	46,875
B.8 Outside "Quality Control" Review of Draft Evaluation Reports	1.5	9,375	3.0	18,750	4.5	28,125	7.5	46,875
B.9 Secondary Analysis of Sets of Evaluation Reports & of Methodologies/Measures--Culling Additional Lessons	1.5	9,375	3.0	18,750	4.5	28,125	7.5	46,875
SUBTOTAL (for tasks done by PPC, regional Bureaus, DSB, FHA and other central Offices), *professional time*	14.0	90,265	19.0	118,750	23.5	146,875	32.5	203,125
Support Personnel Time/Dollars	7.0	14,581	9.5	19,783	12.0	24,995	16.0	33,328
	21.0	105,206	28.5	138,538	35.5	171,871	48.5	236,453
TOTAL OF MINIMUM AND RECOMMENDED SUPPLEMENTARY IMPACT EVALUATION MANAGEMENT TASKS (LABOR ONLY): YEAR ONE	36.0	175,121	54.0	262,493	73.0	354,158	108.5	528,113

* Professional time for AID/W is calculated at $75,000 based on figures provided to PCI in July on the "Average Worldwide Overseas Cost Per U.S. FTEPP Employee: FY80 Operating Expenses". The overseas average was given at $87,500, of which $16,600 was for allowances given to personnel overseas.

** Support time is estimated to be one-third of the cost of AID/W professional time and expected to include basic materials and services (Items) required to perform tasks.

BEST AVAILABLE COPY

TABLE IX-5:

IMPACT EVALUATION PROGRAM MANAGEMENT COSTS -- YEAR 2...*N* (LABOR ONLY)*

TASK DESCRIPTION	10 Studies		20 Studies		30 Studies		50 Studies	
	MM	$	MM	$	MM	$	MM	$
A. MINIMUM SET OF TASKS (per Year One List)	15.0	72,915	25.5	123,955	37.5	182,287	60.0	291,660
B. RECOMMENDED SUPPLEMENTAL TASKS								
B.4 Annual update of information available to teams on measures/field data collection/analysis techniques, i.e. a digest of the methodology lessons learned from prior years' work	1.0	6,250	1.0	6,250	1.5	9,375	1.5	9,375
B.7 Outside "quality control" review of evaluation plans for each evaluation	1.5	9,375	3.0	18,750	4.5	28,125	7.5	46,875
B.8 Outside "quality control" review of draft evaluation reports	1.5	9,375	3.0	18,750	4.5	28,125	7.5	46,875
B.9 Secondary analysis	1.5	9,375	3.0	18,750	4.5	28,125	7.5	46,875
SUBTOTAL: Professional time	5.5	34,375	10.0	62,500	15.0	93,750	24.0	150,000
Support Personnel time/dollars	3.0	6,249	5.0	10,415	7.5	15,622	12.0	25,000
	8.5	40,624	15.0	72,915	22.5	109,372	36.0	175,000
TOTAL OF MINIMUM & RECOMMENDED SUPPLEMENTAL IMPACT EVALUATION MANAGEMENT TASKS (Year 2...n)	23.5	113,539	40.5	196,870	60.0	291,659	96.0	466,660

* Estimates of cost as per Year I; no margin for inflation computed.

BEST AVAILABLE COPY

TABLE IX-10:

IMPACT EVALUATION COST ESTIMATE SUMMARY*

(This summary assumes 20 studies/"basic" scope)

KEY TASKS	YEAR ONE		YEAR TWO..."N"	
	Probable Low	Probable High**	Probable Low	Probable High**
1. Conduct 20 project impact evaluations per year. (This estimate assumes use of the "basic" scope, a mix of four types of studies, and a mix of approaches to team composition. Backup sheets in Volume II show other options.)	$1,188,000	$1,906,600	$1,188,000	$1,906,600
2. Manage AID's program of impact evaluations. (The high estimates involve all of the recommended supplementary tasks. AID can calculate intermediate options from backup data in Volume II.)	123,995	262,493	123,995	196,870
TOTAL:	$1,311,995	$2,169,093	$1,311,995	$2,103,470
UNIT COST:	$ 65,599	$ 108,454	$ 65,599	$ 105,173

* No estimate is made for inflation.

** High does not mean "most expensive" approach AID could take. A description of the "most expensive" scenario is provided with the backup sheets in Volume II.

BEST AVAILABLE COPY

to involve a rerun of the initial survey using the same sample frame, to establish trends as well as get second-generation, the cost for one study would be:

$166,452	(single study with management burden: AIDs)
37,322	(secondary data supplement)
123,329	(second run of survey for sustainability data)
$327,113	

For 20 studies at this total cost, AID would need $6,542,260. p.a.

C. LABOR ESTIMATES FOR ADDITIONAL PROJECT DESIGN PHASE WORK THAT WILL PROVIDE THE AGENCY WITH AN IMPROVED BASIS FOR EVALUATION

Below, PCI estimates the labor required to undertake two design phase tasks that would improve the basis for impact evaluation. We do not recommend that these costs be considered part of the impact evaluation program costs. Rather, the cost of the first item, improved project logic, should be borne by AID as part of the cost of the management of the project design system. The second item, baseline data collection, should be borne by each project. A requirement for this type of information already exists and some projects now collect this type of information.

1. Improved Project Logic

Each Logical Framework deserves a more thorough review than it appears to receive at present. Many are not in compliance with the guidance on their preparation; some appear to be last-minute exercises. Two aspects of the design logic, in the text and matrix, warrant attention as AID moves toward impact evaluation:

- The extended logic of the project should be discernable from the project paper. At present, "levels" of the hierarchy of objectives are often meshed in order to present, as the highest objective of a project, a statement that aims at mandated or country (CDSS) goals. When this meshing is done, evaluation teams are forced to reconstruct (with the possibility of error) intermediate steps in the logic of the designers;

BEST AVAILABLE COPY

- Other aspects of the design, including indicators and their targets, means of verification and assumptions, are not always well thought out or fully specified. Again, attention to these items, i.e., a critical review during design would strenghten the basis for evaluation.

Estimated Level of Effort Required to Strengthen Project Designs as a Basis for Impact Evaluation

a. One day per project (X 200 projects per year) for PPC/E examination and critique of text/Logical Frameworks with respect to full statement of extended logic of project and other elements of design from evaluation perspective. This review should include an effort to identify whether an "after-only" design will be appropriate if the project is selected for impact evaluation. Where this proves not to be the case, a recommendation for investments in baseline data, etc., should be made. The review should take place *as early as possible* in order to allow time for Mission/office revisions before final PP preparation.

Total Manmonths: 20 per year

b. Baseline data collection, including the collection of baseline data for a comparison group that will not receive project assistance, *if*:

- A determination is made at the point of design that an impact evaluation will be carried out later, *or*

- AID's design stage review of the project suggests that there is important information that can be gained only by having longitudinal as well as cross-sectional data on the project, e.g., migration effects.

Total Manmonths: 11-12 per project*

* Assuming costs for baseline data collection to be equal to the costs of later impact data collection, with the possibility for savings later if the instruments and procedures used at the design stage can be reused during impact evaluation.

BEST AVAILABLE COPY