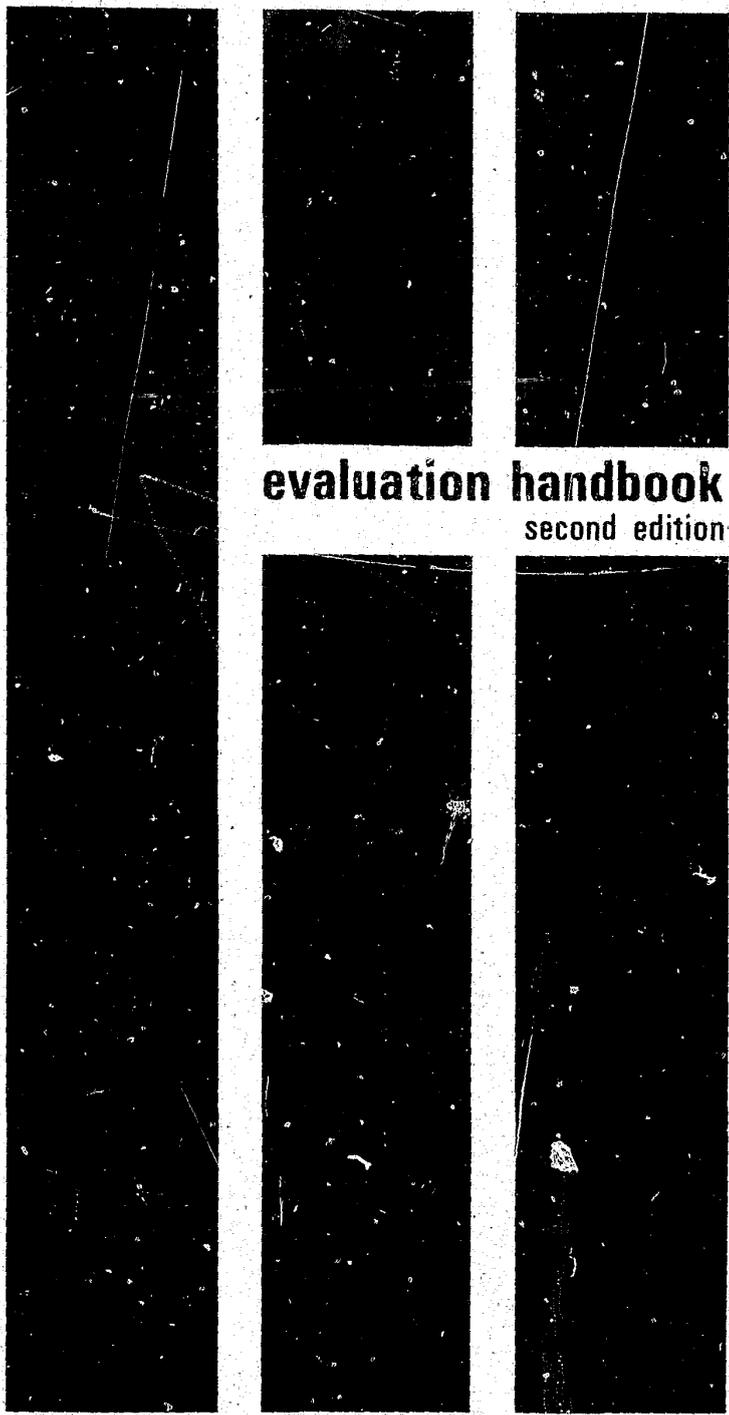


X



evaluation handbook
second edition

156472
EVALUATION

M.C. 1026.1
Supplement II

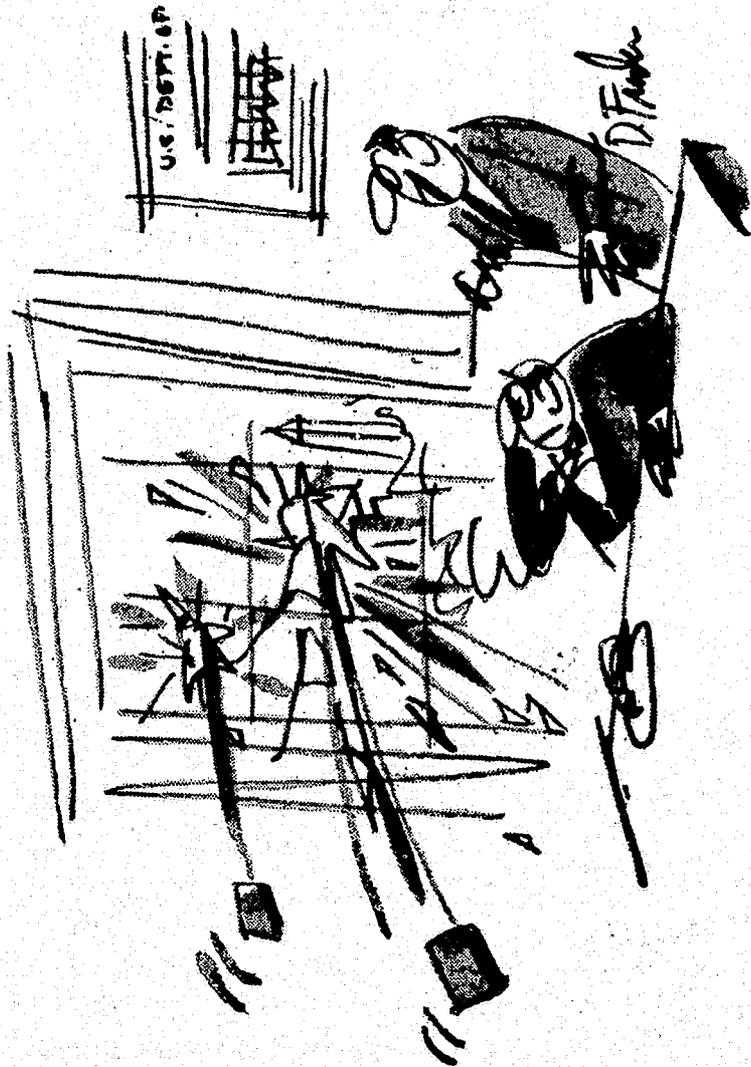
evaluation handbook 2D ED
second edition



Office of Program Evaluation
United States
Agency for International Development
Washington, D.C. 20523

M.C. 1026.1
Supplement II

third printing
MAY 1974



"Ah! We're beginning to get some feedback!"

Copyright 1971 Saturday Review, Inc.

TABLE OF CONTENTS

	<u>PAGE</u>
FOREWORD.	vii
INTRODUCTION.	ix
CHAPTER	
I. THE <u>WHAT</u> AND <u>WHY</u> OF PROGRAM EVALUATION.	1
II. THE AID EVALUATION SYSTEM	5
AID Evaluation Process.	5
AID Evaluation Organization and Responsibilities.	6
Director of Program Evaluation, AID/W	6
Regional Bureau Evaluation Officers	6
The Program Evaluation Officer.	7
AID Reference Center (Memory Bank).	8
III. THE ANNUAL EVALUATION PROCESS	11
Annual Program Evaluation Plan.	11
Project Proposals	12
Implementation Plans.	13
Annual Evaluation of Technical Assistance and Other Noncapital Projects	14
- The Process.	14
- The Concept.	16
- Relation of Project Purpose and Program Goals	18
Evaluation of Capital Assistance.	20
IV. EVALUATION STUDIES	21
Indepth Project Level Evaluations	22
- Technical Assistance and Food Assistance.	22
- Development Loans - Capital Projects.	24
Sector Level Evaluations.	26
Program Level Evaluations	29
- Country Program Evaluations	29
- Program Loan Evaluations.	29
Multi-Country Evaluation Studies, Spring Reviews.	29
Special Evaluations of Assistance Techniques and Policies.	31

CHAPTER	<u>PAGE</u>
V. DESIGN OF EVALUATION STUDIES	33
Criteria for Designing the Study	34
A Basic Study Design	35
Suggested Checklist for Planning An Evaluation Study	40
The Selection of Evaluators	41
- Basis for Selection	43
- Combinations of In-House and Outside Experts	43
- Sources of Evaluators	43
Consultants, the Care and Feeding of	44
- Briefing of Consultant	44
- Mission Participation and Liaison with Consultants	45
- Timing and submission of the Report from Consultant	45
Analysis of Data	45
Preparation of the Final Report	46
VI. MEASUREMENT, DATA COLLECTION AND ANALYSIS	47
Data Collection	49
- Direct Methods	50
- Indirect Methods	51
Dimensions of Progress	53
- Baseline Data	53
- Indicators	54
Non-economic Indicators	56
Performance Standards	58
Advantages and Disadvantages of Using Indicators and Standards	58
Quantitative vs. Qualitative Measures	59
VII. ISSUES IN PROGRAM EVALUATION	60
Candor and Objectivity	60
Joint Evaluations with Cooperating Countries . . .	61
- Types of Participation	62
- Pros and Cons of Joint Participation in Evaluation	64

CHAPTER	<u>PAGE</u>
APPENDICES	65
APPENDIX A Glossary of Terms	66
APPENDIX B Illustrative Baseline Measures	70
APPENDIX C Selected Output	76
APPENDIX D Suggested "Social Indicators"	79
APPENDIX E Indicators - Alliance for Progress . . .	87
APPENDIX F Performance Standards	93

ANNOTATED BIBLIOGRAPHY	103
----------------------------------	-----

LIST OF CHARTS

Figure 1 - The Logical Framework of a Technical Assistance Project	19
Figure 2 - "Ideal" Study Design for Making Comparisons	36

FOREWORD

Over the last few years the Agency for International Development has created and put into use a program evaluation system which has helped significantly to improve both our assistance programs and our understanding of the development problems which those programs aim to solve.

We cannot rest on past accomplishments.

In a 1972 memorandum to heads of Executive Departments and Agencies, President Nixon stated that:

"Program evaluation is one of your most important responsibilities...As the President's Advisory Committee on Executive Organization has emphasized, each Agency must continually evaluate its own programs."

In AID's highly decentralized organization, Missions and individual project officers play an important role in program evaluation activities. This edition of the Evaluation Handbook is designed to stimulate and assist AID staff abroad and in Washington to do an even better job of evaluating in the future.

INTRODUCTION

The U.S. Agency for International Development and the foreign governments it assists are faced with three basic issues: to identify the more important goals which need to be addressed, to design activities which are most likely to bring about the desired changes, and to administer the activities as efficiently as possible.

Each of these three issues can be met more successfully with the use of findings from evaluation of experience. As Sir Winston Churchill once said, "I pass with relief from the tossing sea of Cause and Theory to the firm ground of Result and Fact."

The material contained in the following pages represents a compilation and a condensation of the information on the Agency's evaluation system. It is presented in handbook form to assist evaluation officers, program and project officers, contract team chiefs, and anyone else concerned with evaluation. We hope that it will help them in the performance of their duties, and provide a ready reference work for all those interested in learning more about this subject.

- This second edition of the Evaluation Handbook was edited by Gerald Schwab, U.S. Operations Mission to Thailand, who together with Philip Sperling, AID/W, prepared the first edition. Significant contributions were made by Robert L. Hubbell, Ronald W. Jones, and Herbert D. Tunner, as well as the other members of the Program Evaluation Committee and many Mission Personnel. Special appreciation is expressed to Lea Knott of the U.S. AID Mission to Laos for her editorial assistance, to Joan Silver for managing production arrangements, and to Marilyn Steenburgh for her patience and skill in typing both the draft and the final copy.

The second edition of the Evaluation Handbook was originally published in February of 1972. This second printing of that edition reflects demand for copies both from AID and its intermediaries, and from other organizations.

Chapter I

THE WHAT AND WHY OF PROGRAM EVALUATION

One prematurely gray colleague characterized A.I.D. as having a

20-year job with a
10-year plan, a
2-year tour, and a
1-year appropriation.

While the frustrations inherent in such a situation are obvious, it is clearly incumbent on A.I.D. to make the best possible use of its resources at all times. It is our contention that evaluation can play a great part in this effort, provided the findings are applied to planning or replanning. If used properly, evaluation findings should permit A.I.D. to materially improve the quality of performance; if not so used, evaluation is not worth the effort, despite its historical interest.

The classic dramatic character, Lothario, when queried about the secret of his success, explained that over a long period of time he had found it most helpful to break each conquest down into three distinct parts: planning it, doing it, and then analyzing it to determine why it had (or occasionally had not) worked as planned.

A.I.D.'s analysis of its program management procedures also has identified three similar factors which look -- but are not always -- as easy as PIE.

- P - Planning - Deciding what (and how much) to do and how to do it;
- I - Implementing - Doing it;
- E - Evaluating - Appraising the actual results in order to determine effectiveness, significance, and efficiency.

Evaluation provides the factual information about what happened, and thus becomes a key management tool for improving planning and implementation of new and ongoing activities.

There appears to be relatively little disagreement in defining planning and implementation. However, a discussion among interested parties resulted in a variety of definitions of the term evaluation.

- Some said it meant measuring progress toward a target.
- Others said it was analyzing reasons for the outcome.
- Still others said that there is no evaluation unless we look at the significance of a project, at linkages, at relationships to sectors, to economic development, to civic participation, to something bigger than the project.
- Some said evaluation is a Project Appraisal Report -- a PAR.
- And others said, that an evaluation which produces only a PAR is PARalysis.

A possible conclusion: Evaluation can be many things. It can be ascertaining whether we are meeting the targets. And, if not, why not? Should we do more of the same? Should we change? Should we quit? Do the targets make sense? Or, to use a somewhat more formal definition, program evaluation can be described as a systematic assessment of actions in order to improve planning or implementation of current and future activities. It is one aspect of the intertwined program management cycle consisting of planning, implementation, and evaluation.

Evaluation seeks to answer three basic questions which should be asked of all kinds of assistance at all levels -- project, sector, country program:

Effectiveness - Are the targets for outputs and purposes being achieved? What are the reasons for success or failure?

Significance - Will the achievement of the targets contribute to economic development or other higher goals beyond the project purpose? To what extent? What are the activity's advantages over possible alternatives? What about side effects?

Efficiency - Do the benefits justify the cost? Are there more efficient means of achieving the same targets?

The primary purpose of evaluation is to assist planners and managers in making decisions about programs and projects by:

- Verifying the activity's appropriateness and effectiveness in order to permit an informed decision about continuing the activity;
- Providing a basis for selecting alternative courses of action; and by
- Making lessons learned available for current or future planning.

In brief, evaluation is designed to assist management to obtain reasonably objective information about projects and programs in a regular fashion so that lessons learned can be applied to current planning decisions or to future operations.

Evaluation, as used in the context of this Handbook, differs materially from monitoring or from regular audits and inspections. The latter are generally designed to appraise operations in order to determine compliance with management controls and regulations. As such, they do not as a rule challenge the choice of targets. Evaluation, on the other hand, questions the relevance of the project, challenges all aspects of the project design, examines performance of inputs and implementing agents, measures progress toward targets and may well result in redesign and replanning actions. Audits may uncover inefficiencies in implementation or lack of clarity in targets which concern the planner and manager. Hence, evaluators must keep informed of audit findings and avoid duplication of work in looking at project effectiveness and efficiency. Finally, evaluation also differs

from project monitoring which is concerned with the day-to-day supervision of procurement, delivery, and installation of inputs, and the production of outputs to assure that progress is on schedule. A good monitoring system will, of course, make periodic evaluations much easier.

Aside from the primary purposes of systematic evaluations, there are likely to be derived from the process certain benefits which may be of equal or perhaps even greater value. These include:

- Sharper definition of purposes and goals. Evaluations have a way of exposing high-sounding projects which have not been reduced to measurable or verifiable targets. How does one evaluate a project which has as its purpose, "to help improve the quality", "to expand and improve", or "to increase the effectiveness" of an institution (not to speak of making it "viable"), when specific targets are not provided? At times, the evaluation process will result in a more clearly defined purpose, thus providing a better basis for measuring progress and planning actions.
- Improved understanding and internal communication. As a result of analyzing and discussing a project, vertical and horizontal intra-office communications are greatly facilitated. Technicians and contractors learn more precisely what is expected of them. Supervisors acquire a better understanding of the problems encountered by staff members, and vexing problems may for the first time be brought to the attention of top management.
- API (Anti-procrastination Incentive). Without going into the question of whether any component of A.I.D. could ever be accused of procrastination, it has been observed that an evaluation, or the mere scheduling of an evaluation, frequently causes offices to address themselves posthaste to elements known to be behind schedule or of poor quality, and to place these on their action agenda.

Chapter II

THE A.I.D. EVALUATION SYSTEM

I'd like to know
what this whole show
is all about
before it's out.

Piet Hein

The Foreign Assistance Act of 1961, as Amended, makes explicit the expectation that the Agency will conduct evaluations. Part I, Chapter 2, Title V, Section 241 reads:

(a) The President is authorized to use funds made available for this part to carry out programs of research into, and evaluation of, the process of economic development in less developed friendly countries and areas, into the factors affecting the relative successes and costs of development activities, and into the means, techniques, and such other aspects of development assistance as he may determine, in order to render such assistance of increasing value and benefit.

A.I.D. Evaluation Process

A.I.D. assigns primary responsibility for program evaluation to the action units of the Agency. Missions and appropriate AID/W offices are expected to appraise progress toward targets and also to consider the validity of the targets themselves. Responsibility is so placed because only the action units can effectively make changes indicated by evaluation findings. This requires a regular evaluation process which calls for the systematic collection and analysis of objective data, which periodically brings a variety of viewpoints to bear on activities and problems, and which relates evaluation findings to action

decisions. This process goes far beyond the preparation of reports, although its conclusions may be recorded in reports. The process is described in detail in Chapter III.

A.I.D. Evaluation Organization and Responsibilities

Specific evaluation activities are largely the responsibility of individual Missions and those AID/W offices charged with direct supervision of specific programs. Coordination and supporting functions are provided by the Director of Program Evaluation in cooperation with AID/W offices and the Regional Bureaus. Internal coordination among these offices is facilitated by their membership on the Program Evaluation Committee (PEC)^{1/}, which meets regularly to discuss procedures and to exchange information.

Director of Program Evaluation, AID/W

The Director of Program Evaluation, located in the Bureau for Program and Policy Coordination, develops evaluation methodology and coordinates the evaluation activities of the various bureaus and staff offices. He arranges for the exchange of information pertaining to techniques and results of evaluation within A.I.D. and with other donors; provides general guidance and training in evaluation; and conducts or supports evaluation studies of Agency-wide policy and program issues and problems. He carries out these functions in cooperation with the members of the Program Evaluation Committee, which he chairs.

Regional Bureau Evaluation Officers

Regional Bureau evaluation officers backstop the overseas evaluation activities in their respective geographic areas, serve as advisors on evaluation matters within the Bureau, and represent the Bureau on the A.I.D. Program Evaluation Committee.

^{1/} PEC members include representatives of each of the Regional Bureaus, the staff bureaus, and of the Office of Food for Peace, and the Auditor General.

Although their specific tasks differ somewhat from region to region, Regional Bureau evaluation officers are generally responsible for:

- facilitating AID/W review and use of annual evaluation plans, Project Appraisal Reports, and special evaluations, and for coordinating ensuing comments and support to the Missions;
- serving as the focal point in the Bureau for the collection and dissemination of evaluation experience, methodology, and findings;
- participating in the selection and training of Mission evaluation officers and of special evaluation teams;
- assisting in the introduction and supervision of the regional evaluation activities as well as participating in the conduct of these as need arises.

The Program Evaluation Officer

The primary responsibility for assuring adequate program evaluation rests with each Mission Director and AID/W Assistant Administrator. His attitude towards evaluation shapes that of his organization, and it is up to him how he specifically decides to organize for this purpose. To assist him, he should have an officer responsible for the staff functions needed to make the evaluation system work effectively. Each Mission and AID/W office responsible for project activities has been asked to designate an evaluation officer for this purpose.

The core assignment of the evaluation officer is to coordinate and facilitate the planning and carrying out of evaluation activities of the various office elements, in order to assure a unified and orderly annual evaluation program. For this core assignment, he is the systems manager, and not the evaluator.

The evaluation officer plans the organization's evaluation activities and participates in their execution to the extent considered appropriate under local circumstances. Since the reason for involving action officers in evaluation is to have them participate in the development of changes in plans so that

they will execute these changes, the evaluation officer loses effectiveness if he completely takes over the evaluation. The action officers would then defend themselves against the evaluation officer rather than working with him.

The evaluation officer also directs the analysis and dissemination of evaluation data -- both those data developed internally and those received from other sources -- to insure maximum utility of the findings for program planning and improvement, and to facilitate the transfer of insights gained to other potential users.

A.I.D. Reference Center (Memory Bank)

Program evaluation assumes that we can learn from our experience. For the most part, lessons learned are used in the offices where the evaluation occurred in order to improve ongoing activities or to plan similar future activities. However, some conclusions based on experience in one country may be applicable elsewhere. The conclusions may apply not only to the substance of projects and programs, but also to techniques for studying feasibility or for conducting evaluations.

In the past, A.I.D. has been characterized as an Agency without a memory. If a project manager sought reports on experience elsewhere, his technical backstop or desk officer had to undertake a search to discover where similar activities had been tried, and to locate reports from scattered files. Regular retirement of records made it unlikely that reports over three years old could be easily located. Within the recent past, however, significant progress has been made in overcoming this amnesia through the establishment of the A.I.D. Reference Center.

Contents of the Memory Bank

The A.I.D. Reference Center (ARC), located in Room 1656 New State Building, is popularly known as the Memory Bank. It consists of a central, permanent collection of selected "AID memory" materials -- e.g., reports and documents which help in the transfer of A.I.D. experience. Highest priority is given to the collection of the following kinds of materials:

- Evaluation documents and case studies: Materials that analyze A.I.D. experience in development assistance situations. These documents describe the experience, assess accomplishments, and discuss possible alternatives for future similar situations.
- Special Studies: Various A.I.D.-generated special studies or issues papers which analyze development assistance problems.
- Program documents: These include formal documents (project budget submissions, country field submissions, program memoranda) and informal documents (sector analyses, country programs, interregional programs, and others).
- Project documents: Substantive documentation such as Noncapital Project Papers (PROPs) and Project Appraisal Reports (PARs) which will enable users to draw on A.I.D. project experience.
- Reports: Feasibility studies, A.I.D. research reports, various kinds of progress and terminal reports on A.I.D. projects, and end-of-tour reports by A.I.D. technicians and contractors.

Mail rooms, contractors, etc., systematically send formal recurring documents, such as PROPs and PARs to ARC. However, many other valuable documents, such as special evaluations, termination reports, issues papers, etc., may be missed unless originating officers remember to direct them to ARC. Documents of interest should be addressed as follows: Attention: PPC/ARC, Room 1656, New State. 2/ If possible, two copies should be sent.

Use of Memory Bank

Overseas personnel should send requests for information to be obtained from Memory Bank materials through their Regional Bureau. This has the advantage that an informed backstop

2/ Detailed instructions for sending documents to the ARC are covered in the Annual Evaluation Plan messages, the Project Management Handbook, the Disposition Handbook, and the A.I.D. Procurement Regulations.

person may help the reference librarian select useful documents from ARC. Another way to assure getting the right information is to describe the problem precisely. For example, an individual who requested documents on artificial insemination received in response a veterinarian's technical explanation obtained from the Department of Agriculture. What the requester had really wanted was an account of someone else's experience with the kind of government set-up and farm organizations required to ensure success in a better breeding program.

ARC also assists in the completion of annotated bibliographies in the A.I.D. Bibliography Series, which are issued as guides to materials on development assistance in various fields. Each bibliography is compiled by an expert in a subject matter field of development. The bibliographies contain hundreds of references on subjects such as land reform, urban development, civic participation, malaria eradication, book and library development programs, nonformal education, community water supply, etc. Each bibliography contains instructions on how to order copies.

Chapter III

THE ANNUAL EVALUATION PROCESS

There is measure in all things.

Horace

The evaluation system is an integral part of the overall planning and management process. This section describes, in abbreviated form, some of the required background documentation and procedures. ^{1/}

Annual Program Evaluation Plan

Each year, usually at or near the end of the fiscal year, appropriate A.I.D. Bureaus and Offices are requested to submit their evaluation plans for the coming year. Although the specific information to be provided will differ from year to year, three basic elements will usually be required; a review of evaluation activities carried out during the previous year; a schedule of evaluations planned for the coming year; problems encountered and lessons learned in the course of the previous year's activities.

In order to relate the evaluation plan to the basic issues, key officers must be involved in the formulation of the plan. Field Missions which have some type of evaluation review panel will find it a useful forum for this purpose.

^{1/} In view of the changing nature of these procedures, and the fact that the Evaluation Handbook will not be reissued with every change, current Agency regulations should be consulted for specific guidance and instruction.

Project Proposals

Planning for all types of assistance -- capital, technical, food or a combination of these -- should be based on a sector analysis and strategy statement. For each project, a proposal is required for AID/W authorization which relates it to the sector plan and describes its purpose, implementation, and inputs.

Although the preparation of project proposals is relevant to this Handbook only insofar as the project proposal contains the targets and criteria against which later evaluations can be made, the importance of planning for evaluation at the beginning of an activity within the context of the project proposal cannot be overemphasized.

By using essentially the same structure for both the Non-capital Project Paper (PROP) and the Project Appraisal Report (PAR), -- A.I.D.'s basic evaluation document for technical assistance projects -- a significant step has been taken toward integrating the key elements of the evaluation process into project design at the outset. The definition of specific targets, of the purposes they are to serve and of the means by which they are to be achieved will greatly facilitate subsequent evaluation of performance.

It must be kept in mind in designing a project that it is important not only to define the changes which are to result, but also to establish a baseline reflecting the original situation in which changes are to be made. Thus, it will not be enough in the long run to have PROPs which identify exactly what is to be achieved by the end of the project (i.e., End-of-Project Status or EOPS) and how one verifies that these targets have indeed been achieved. It is necessary also to record the status at the beginning of the project (i.e., Beginning-of-Project Status or BOPS) in such a fashion that subsequent measures can be made against it.

The final step in planning evaluation as part of a project is to determine the indicators or other data that will be needed to ascertain progress. If possible, the planners will use existing sources of data, but they may need to arrange for regular collection of selected information as part of project implementation. A special aspect of data collection may be the use of a comparable control group which will permit better interpretation of the causative relationships between project

activities and observed changes. If a control group seems practical, project planning should include means to select control units and to collect baseline and change data from them. ^{2/}

The amount of data needed for evaluation purposes will, of course, vary with the nature of the proposal. For some types of loans, particularly those which involve tranches where the second phase depends on meeting certain specified conditions in the first phase, inclusion of a satisfactory scheme for evaluation may be required. For certain non-capital projects, particularly those of an experimental nature or those for institutional development, the details of conducting special evaluations may be specified as these go beyond the minimum AID guidelines and instructions providing for annual Project Appraisal Reports.

Implementation Plans

As life-of-project documents, PROPs deal more with general project design than with detailed tactics and schedules. The same is generally true of loan papers, although some may contain considerable detail. In either case, specific plans of action are needed.

For noncapital projects, the Joint Project Implementation Plan (PIP) is prepared in the early stages of the project, usually in conjunction with preparation of the bilateral Project Agreement. It sets out the work schedule and certain output indicators, as well as such key inputs as personnel, participants, and commodity requirements. The progress of a project toward its established targets can be measured against these output indicators in quantitative terms. Some projects, such as those of an advisory or institution-building nature, do not readily lend themselves to quantitative measures. However, even in these cases, it should be possible to provide some defined steps or forms of behavior which can be objectively verified as evidence of achievement.

The documentation for implementation of loans is more complex than for noncapital projects. In part, this difference reflects the fact that the cooperating government is more directly

^{2/} For a detailed treatment of baseline data collection and comparisons, see Chapter VI.

responsible for implementation, and a loan may involve various conditions precedent, each with its own specified reports. A loan may also depend heavily on implementation plans prepared by engineering or management consultant firms.

Whatever the formats and whoever the authors, the totality of the implementation plans should make clear the interim and final objectives so that progress and completion can be observed and evaluated.

Annual Evaluation of Technical Assistance and Other Noncapital Projects

Missions and AID/W offices responsible for the administration of technical assistance and certain other noncapital projects are required to evaluate them annually. The self-evaluation approach should enlist the judgments and suggestions of all knowledgeable personnel, including members of contract and PASA teams, and insofar as practical, of the cooperating country and other donors. This approach goes against past notions that evaluation should be conducted by outsiders -- inspectors, auditors, or other headquarters staff (although they have important roles to play) -- because outsiders cannot achieve complete coverage and are not responsible for putting recommendations into effect. To achieve objectivity in self-evaluation, there is an established process.

The Process

The elements of the noncapital project evaluation process are:

1. A logical framework in which the Mission or AID/W office
 - (1) Defines project inputs, outputs, purpose and goal in measurable or objectively verifiable terms;
 - (2) Hypothesizes the causative linkage between outputs, purpose, and goals; and
 - (3) Establishes the indicators that will permit subsequent measurements or verification of achievements of the defined outputs, purpose, and goal.

The logical framework is not itself an evaluation device; rather, it sets the stage for the evaluation. Evaluation consists of determining and validating whether or not the project outputs were produced, whether these outputs in fact

achieved the project purpose, and finally whether this achievement made a significant contribution, as planned, to the higher goal. By focusing on the causative linkages between inputs, outputs, purpose, and goal, evaluation avoids extraneous and irrelevant questions and looks for possible improvements.

The logical framework requires reexamination of the original design of the project as an integral part of the evaluation. It permits a clear separation between manageable interests (managing inputs to produce outputs) and those factors that appear to be beyond the project team's managerial control. Beyond this stage, it is necessary to act as a social scientist in testing the hypotheses that (1) producing the planned project outputs will result in achieving the project purpose, and (2) achievement of this purpose will result in a significant contribution to a sector or program goal. The review of project design is then followed by an examination of (1) the performance of input factors (personnel, training, commodities) and action agents (USAID, contractors, other donors, cooperating country), and (2) actual progress toward outputs, purpose, and goal.

2. A group review, an interactive process among interested parties, is essential for reaching the best evaluative conclusions and determining future actions. Therefore, formal reviews represent an integral part of the process. The desired approach is a collaborative effort rather than a judicial inquiry. The attendance at these reviews depends on the project. Some Missions have a regular evaluation panel consisting of such officers as the Director, or Deputy Director, Program Officer, Evaluation Officer and Controller, supplemented by people concerned with the particular project. The review might include representatives from the cooperating country government, other donors, or representatives from AID/W (in the case of Mission-managed projects) or the Mission (in case of AID/W-managed projects).

3. A process manager who is responsible for helping project personnel analyze their projects in accordance with the logical framework, and for managing group reviews. Missions and AID/W bureaus and offices have designated an evaluation officer for this purpose.

4. A simplified Project Appraisal Report (PAR), which conceived as a low-cost by-product of the evaluation process, and which

is designed to provide a permanent record of the findings and decisions arrived at during the evaluation review.

At first glance this approach to evaluation may appear too elementary, too pat, to provide a tool for the serious examination of the more profound aspects of economic development. However, closer examination will show that, in fact, the format allows the widest possible latitude for examining the project and its implications. Depending on the size of the project, or other considerations, the evaluator can apply the requisite degree of sophistication and analysis to the collection of data, the examination of causative linkages, or other aspects.

The Concept

Underlying the concept of evaluation is the recognition that much of what A.I.D. is doing is experimental in nature and as such cannot be expected to be successful in all cases. In fact, the development assistance process, like a scientific experiment, may be described as a series of hypotheses. We anticipate that if donor and recipient countries provide certain inputs, a predicted output will occur. This is presumed to be manageable. We then hypothesize that, if this output occurs, certain economic or social changes will follow. We hypothesize further that, if these changes take place, then higher living standards or national income or political stability or other broad goals will be achieved.

The evaluator first confirms that inputs indeed produced intended outputs. If not, he ascertains the changes needed to produce the outputs. He then becomes the social scientist who tests the hypotheses. Were they valid? If not, what explicit or implicit presumptions proved incorrect? Such testing of presumptions moves evaluation beyond monitoring and auditing.

To recapitulate, the process of analysis should follow the logical progression of a development project:

- (1) If adequate inputs are provided, then planned outputs will be produced.^{3/}

^{3/} See Appendix A - Glossary of Terms.

- (2) If these outputs are produced, then purpose will be achieved.
- (3) If purpose is achieved, then the planned degree of progress toward a higher goal will occur.

The first stage of the progression -- inputs to outputs -- is manageable. The next two stages -- outputs to purpose and purpose to goal -- are hypotheses which can be tested. Evaluation assesses progress at all stages and checks linkages. If one stage does not lead to the next, evaluation reexamines the implicit presumptions and considers alternatives to the mixture of inputs or to the nature of the purpose and goal.

Note that the word manageable is used here in its twentieth-century sense. A manager promotes the cooperation of equals to achieve results; he does not act as a czar who issues orders. Especially in A.I.D., which operates in an "open system" with a cooperating country government and other donors, project teams need to use tact and persuasive means. When A.I.D. provides inputs to supplement cooperating country and other donor inputs, it assumes a degree of responsibility for outputs in a complicated joint situation. Its power consists of knowledge, attention, and persuasion, and this is what modern management is about. A comparable situation is the project officer for the launching of an Apollo shot to the moon, who cannot order the U. S. Navy to deploy ships in the South Pacific to recover the astronauts, but had better be sure such arrangements are made before the launching.

Use of this logical framework in evaluating projects demands that project progress be measured in two stages: First, inputs to outputs must be measured because it is necessary to measure that which management is expected to produce. Secondly, the evaluation process must then independently measure progress toward the project purpose. (The measurement of progress toward purpose must be independent of measuring outputs, otherwise a logical fallacy results. It would not prove or test the hypothesis that if the output, then the purpose.)

By focusing on independent measures of outputs and progress toward project purpose, the use of the logical framework should help reduce management's preoccupation with inputs. Adopting the experimental viewpoint of a scientist, as opposed to that of a manager does not lessen management accountability and the distinction between the subjective and the objective. Production of outputs and achievement of purpose are objectively

verifiable; the subjective element is the judgment that producing the outputs will achieve the purpose. To adopt the experimental viewpoint does not imply that there can be little confidence in judgments regarding achievement of purpose. The scientist states premise and process from which he deduces certain probable results. An equally salient aspect of the scientific method is a painstaking review when results are not as expected. The careful and objective sorting of evidence is what assistance managers must strive for, and the logical framework was designed to support such a careful and objective process. The logical framework is shown in Figure 1.

For the evaluation process to be useful, it must be carried out with the utmost candor and objectivity. Proposals to change or adjust shortcomings in strategy are the mark of alert and flexible officers who take advantage of experience. Adjustments may also be regarded as a necessary facet of the difficult process of trying to generate economic and social changes.

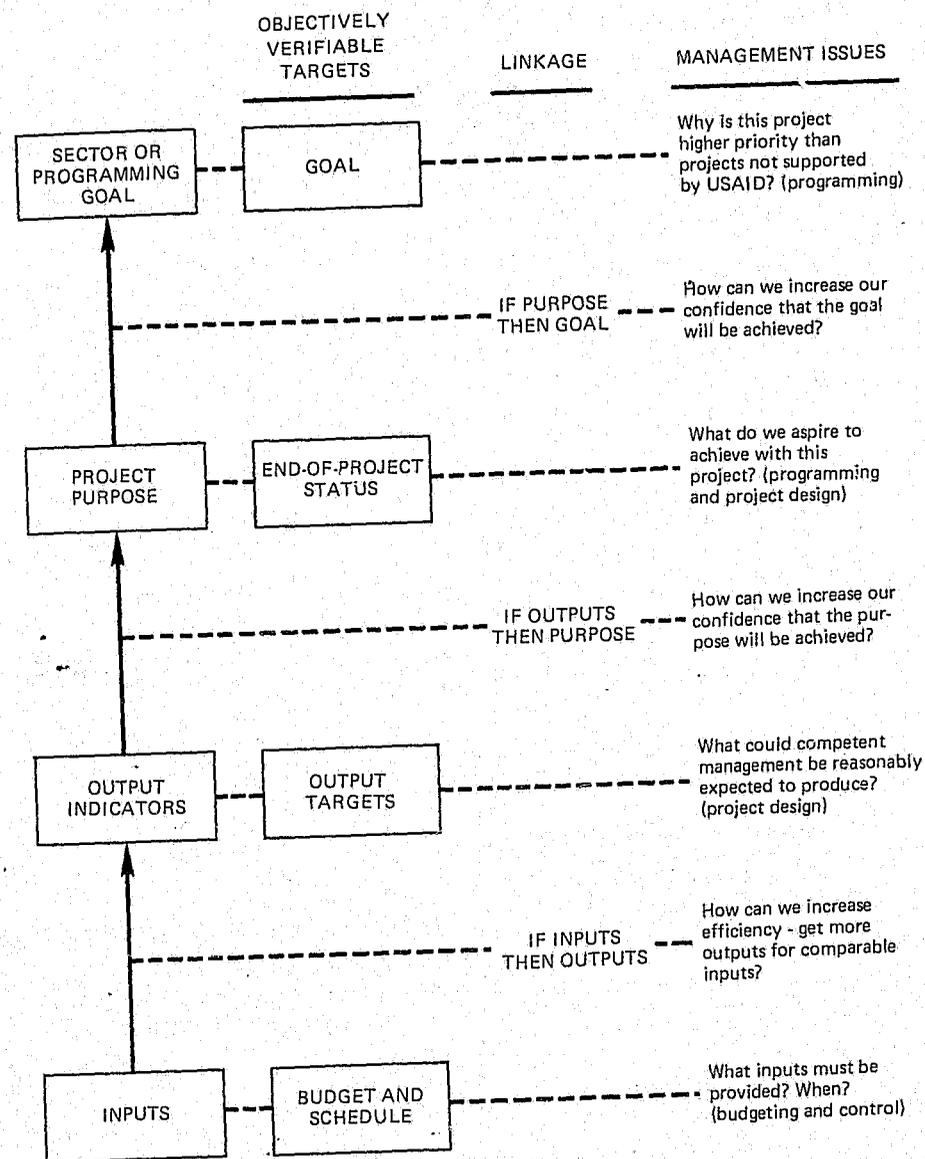
Relation of Project Purpose and Program Goals

A.I.D.'s present evaluation system is project oriented. Although the evaluation instructions provide for scrutiny of major objectives, the causative link between the project purpose and the broader sector objectives or program goals for the particular country may be difficult to see. The linkages between project outputs and purposes, between purposes and country program goals or objectives are considered to be a series of interconnected hypotheses about economic, social, and political development.

In actuality, the impact of a small project such as a pilot agricultural school upon a broad objective, such as "self-sufficiency in agriculture", is not going to be great and would be exceedingly difficult to trace. Such is the case when a country strategy includes such broad objectives as "reducing the balance of payments gap" or "making the distribution of income in the rural areas more equitable." It could be useful then to approach a project from a different perspective; for example, to analyze it in relation to the sector goal.

FIGURE 1

THE LOGICAL FRAMEWORK OF A TECHNICAL ASSISTANCE PROJECT



Approaching a project, particularly a small one, from the narrower confine of the sector goal may provide a project manager with a better framework within which to judge the relationship between project purpose and higher goal.

Evaluation of Capital Assistance

Annual evaluation according to the aforementioned PAR process is required for technical assistance components over \$100,000 which are part of capital projects. Other required evaluation for development finance -- which includes not only capital projects but sector and program loans -- is quite widespread, however, it is concentrated in the area of evaluation studies, rather than encompassed by the systematic, annual approach applied to technical assistance. The nature of evaluation studies as they apply to capital assistance projects and other forms of development finance is discussed in the chapter immediately following, Chapter IV; the methodology for carrying them out is described in Chapter V.

Chapter IV

EVALUATION STUDIES

Problems worthy
of attack
prove their worth
by hitting back.

Piet Hein

The other key element in the overall A.I.D. Evaluation System, in addition to the non-capital evaluation process described in Chapter III, consists of Evaluation Studies. These are defined as studies which encompass a deeper analysis than that involved in the annual project evaluation process (although the problem being studied, in the case of technical assistance, may well have been flagged during that process and recorded in the PAR); require technical or analytical skills which may not be available in kind or quantity in the Mission; or endeavor to answer questions beyond the project level. Evaluation studies, in addition to being an instrument for conducting in-depth evaluation of on-going projects satisfy several other needs which the annual non-capital project evaluation process was not designed to serve. These include evaluation of terminated projects; evaluation of activities which cut across project lines, such as third-country training or multi-project or sector activities; and analysis of multi-country experiences -- a component of the evaluation system for which the Agency, through "Spring Reviews" and other devices has been building a capability over the past few years. Finally, evaluation studies are the area in which evaluation of development finance is concentrated.

There are three basic types of A.I.D. assistance to which evaluation studies are applied, and which need to be distinguished from each other because the evaluative approach may differ somewhat between them. These aid categories are:

- Technical Assistance -- which also includes the special category of Participant Training;
- Development Finance -- which includes loan-funded capital assistance projects, sector loans, and program loans; and
- Food Assistance projects.

Technical assistance is generally grant-funded, but occasionally is loan-financed, particularly under a comprehensive sector loan. Capital projects are generally loan-funded, but a few are grant-funded, mainly from supporting assistance.

There are four basic types of special evaluations which may be applied to any of these aid categories, depending on the kind of information required. These types of evaluation are defined by the level on which the analysis focuses, i.e., project level, sector level, country program level, or multi-country level.

In addition, special evaluations may be done of assistance techniques and policies. These do not concern specific projects or programs.

The various types of evaluation studies are described below. More detailed discussion of the methodology which can be applied to these is contained in the following chapter, Chapter V.

In-Depth Project Level Evaluations

In-depth project level evaluations can be and are regularly applied to each of the three types of A.I.D. assistance.

Technical Assistance and Food Assistance: Despite the value of the non-capital project evaluation system as a tool for evaluating individual projects and replanning activities, there remain instances in which in-depth evaluations of specific projects will be both appropriate and desirable. In some instances, the annual evaluation process -- including regional projects -- may be instrumental in calling attention to the

needed for an in-depth study. 1/

The reasons for an in-depth study vary greatly, but they are likely to fall into one of the following categories:

- To reappraise a project's rationale or direction and to examine planned or alternative courses of action with the assistance of outside observers or persons of specific technical or analytical skills;
- To examine in depth some key linkage(s), perhaps identified in the course of the annual evaluation process;
- To carry out extensive field studies in connection with the examination of a project's performance;
- To establish a historical record and analysis covering the life of the project; and
- To study completed or terminating projects, putting special stress on recording the significant techniques or lessons learned which might be transferable or applicable to other activities.

Ways of designing an in-depth evaluation study are numerous, depending on the reason for which it is being undertaken and the information sought. An idea of the variety that is possible is evident in the following examples of studies which have been carried out within the past few years:

- The evaluation of the institutional maturity of a country's agricultural university, under an A.I.D. contract, was carried out over a six-week period by two visiting consultants. Their recommendations were considered in developing plans for an agricultural research project.

1/ In the design of an in-depth study, it will be helpful under most circumstances to keep the logical framework and the technical assistance project evaluation system in mind as a way of assuring that the important issues are addressed, and that the study and its findings are related to subsequent annual PARs.

- A joint Mission-cooperating country team examined an institute of business administration to ascertain the current effectiveness of the institution (formerly assisted by A.I.D.), and to assess the relationship of the institution to the cooperating country's basic educational needs at the time of the study.
- A team of experts from the National Communicable Disease Center reviewed the Mission's malaria eradication program to identify reasons for failure to interrupt malaria transmission and to evaluate the adequacy of methods being taken to cope with the problem.
- A full-scale evaluation of a PL 480, Title II, Food-for-Work program was carried out by a Task Force of Participating Agency team, contract and Mission direct-hire employees, representing a wide range of professional disciplines, and a representative of the cooperating country's Ministry of Planning. The work of the Task Force was coordinated by and the final report prepared by the Mission's Evaluation Officer.
- With the assistance of a consultant from the U. S. Department of Labor, Bureau of Apprenticeship and Training, a two-stage evaluation was conducted of a terminating central training institute project. The study was designed to assess the success of A.I.D.'s institution-building effort -- the ability of the project to carry on without U.S. assistance -- and the relevance and value of the project to the cooperating country's development. The first part involved a three-month assessment to review the history of the project and the quality of technical assistance supplied; the second stage, conducted twelve to eighteen months after the completion of the first, was to determine if U.S. assistance had had a sustained impact.

Development Loans - Capital Projects: Project loans finance the foreign exchange costs of constructing infrastructure such as roads, airports, power plants, or irrigation systems. They are preceded by economic feasibility and engineering design studies. There are often conditions precedent and implementation papers. A supervisory

engineering firm deals with the construction contractor(s). During the project, there are inspections and monitoring reports. There may also be evaluative analyses of non-physical aspects of the project such as management, training, or rate-setting policies. Some project loans finance intermediate credit institutions which lend to industry, farmers, cooperatives, or housing. For these, the evaluation studies the types of loans issued, repayment experience, development impact, management practices, etc.

On occasion, for selected completed projects, Missions and AID/W have carried out special evaluations with a view toward lessons for similar future projects. These evaluations put considerable emphasis on whether the initial feasibility studies were well done, but also look at operating and construction questions. Examples of post-project questions in different problem areas are:

- Engineering - architecture to examine such questions as:
 - (1) What is the use experience -- traffic patterns, power plant loads, acre-feet of irrigation water, classroom hours, number of out-patients and types of in-patients? etc.
 - (2) What is the maintenance experience -- Amount of machine downtime? Do culverts carry floods? Does reservoir silt too rapidly? Does road surface hold up? Does building heat? etc.
- Accounting to compare actual costs and income for income-producing projects with those in the feasibility studies; to analyze cost elements for ways to reduce operating burdens; to provide data for rate-setting, etc.
- Economics to assess actual cost/benefit ratios and compare them to predicted ones; to study correlations between various types of projects and general economic growth; to examine the effects of various types of transport systems, or power generation or skill training; to compile data on aspects which are ancillary to projects, etc.
- Political science and public administration to look at the effective methods of internal organization and training; the ways of gaining political support; the procedures to avoid graft; the advantages and disadvantages

of independent regulatory agencies, or regional or planning agencies; the techniques for obtaining, using or controlling local participation, etc.

- Timing. A problem which can pervade all the various problem areas noted above is timing. For example, was the project conceived at the right stage of development? Was its capacity usable immediately upon construction? Was there a reasonable period allowed for growth (without too long a period for servicing debt on unproductive capacity)?

Sector Level Evaluations

In recent years, A.I.D. has endeavored to relate its assistance more to the development of a sector than to total national growth or to disconnected projects. The sector approach offers possibilities for concentrating technical assistance in order to exert a noticeable influence toward change. It also facilitates the transfer of resources to make a significant impact, either through infrastructure projects, development banks, or commodity imports.

The sectoral viewpoint often affects the approach of an A.I.D. Mission to evaluation and analysis. When a sector goal and program have been articulated, the evaluation of an individual project is facilitated since the connection between that project's purpose and a broader goal is known and may be measurable. On the other hand, Missions often decide after project evaluations as recorded in PARs that an especially organized, in-depth analysis of total sector progress and problems is advisable. The combined results of several PARs on related projects probably will not cover all activities in a sector: grants and loans, capital and non-capital projects, all sources of support (cooperating country, international, private). Nor is a PAR evaluation likely to give enough attention to the relationships or dynamics of a sector that it will reveal bottleneck areas not being touched by existing activities.

Sector analysis, then, is an effort to understand what makes a sector tick. First, the analysis sets out to specify and measure the inputs, outputs and relationships within a sector and between the sector and the rest of the economy. Second, it tries to estimate the direct, indirect and induced effects of alternate policies on output objectives.

Such an analysis may indirectly result in evaluative findings about projects even when the analysis pays little attention to the particular progress of inputs, outputs or purpose of projects. This evaluative spillover occurs because the relative importance (or unimportance) of the problem being addressed by the project becomes more evident in the longer perspective. Indeed, the wide-angle lens of a sector analysis may be the only practicable way to inspire program managers to ask seriously questions which should be a part of every project evaluation; namely, "Have I selected the right targets?" or "Would it matter if this project ceased?"

Agricultural sector analysis has a ten year history in the Agency. Until recently, the standard procedure involved a short-term team, composed of subject matter experts, whose recommendations were based on intuition and a broad familiarity with the country situation. Starting in the late 1960s, an effort has been made to introduce computerized mathematical procedures to agricultural sector analysis, to reduce its dependence on subjective judgments and mental arithmetic. At least three different types of models are being developed. One, by A.I.D. staff, is based on the input-output method with linear programming components. It has been used for sector loans in Colombia. Another, by a team at Michigan State University under contract with A.I.D., uses simulation techniques. The first efforts here used Nigerian data. The third, by the IBRD, is primarily a linear programming exercise. It is being tried in Mexico. No one method can claim absolute superiority, though there is agreement among the analysts that the end product will offer decisionmakers a much more flexible and reliable instrument for planning sector programs. The introduction of mathematical rigor into sector analysis will proceed slowly, however, since it demands a data base which some countries cannot supply and since it is expensive in terms of time and money. One might argue that neither the time nor expense should be constraining elements if the strategies made possible by the computerized analysis of many variables and of tertiary effects facilitate more rapid progress with less investment. However, at this point, the new techniques are not entirely proved or accepted. One difficulty is that both Missions and cooperating countries may lack absorptive capacity for using sophisticated techniques. This is not unique to LDCs. In several American cities or firms, decisionmakers have refused to adopt a course of action which runs counter to their intuition or which they cannot explain to their constituents.

The sector analysis techniques used for agriculture in Colombia have related investment for alternative crops or processing activities to total impact on employment, income distribution and foreign exchange -- three national goals adopted by the Colombian Government. Thus, subsequent evaluation of sector loans and project results should be able to use the same technique and baseline to measure progress toward these national goals.

For other sectors, the introduction of mathematical rigor has proceeded less slowly, partly because of the lack of satisfactory production functions comparable to the one for agriculture. However, the manipulation of massive data which may be available even in less developed countries can provide guidance for program planning. For example, use of various kinds of operating reports of school systems can give clues about problem areas in curriculum or costs. Similarly, studies of demographic and vital statistics indicate target audiences for family planning, education and services.

Whatever the design of the analysis during program planning, evaluations of ongoing programs in the sector must grow apace. This is because policy prescriptions must be related to A.I.D. and cooperating country government programs already underway. The sector evaluation is called for to get a reliable description of present programs and show how much needs to be done to bring them in line with the preferred strategy. Further expansion of sector analysis and improvements in techniques will facilitate subsequent sector evaluations, just as the adoption of the GPOI discipline in project planning simplifies the job of project evaluation for the PAR. However, Missions will undoubtedly rely heavily on temporary duty teams for sector evaluations because such evaluations usually need an interdisciplinary approach and several man-months.

When teams are used, the role of the Mission is to help define the scope of work, to collect data and records in advance of the team arrival, to suggest and arrange appointments and field trips, to react to tentative conclusions, and to follow-up on recommendations. This role is discussed further in the next chapter, especially in the section on the care and feeding of consultants.

Sector loans are the most recent form of development finance; they have been used primarily in Latin America. The criteria for decisions and the methods for programming them are still evolving. These loans start with an agreed upon strategy of policies, investments and technical assistance for a sector or

partial sector such as education or higher education, agriculture or small farmer food crops. The loans may fund imports, local costs, and technical assistance. Often they are disbursed in annual installments or tranches related to progress stages. The loan agreements usually specify that periodic joint evaluations be made, sometimes tying the next disbursement to the extent of progress. The methodologies for such evaluations vary with the aspect of the performance examined -- be it the overall policies, the capital component, or the technical assistance.

Program Level Evaluations

Country Program Evaluations: A country program evaluation consists of reviewing the significance and success of all A.I.D. developmental activities within a particular country. Such evaluations are undertaken when an in-depth and comprehensive view of the A.I.D. program is required, particularly with a view towards replanning strategy and/or levels of assistance. Country program evaluations take place relatively infrequently, and in a variety of circumstances as regard local situation, kind and level of program, specific problems addressed, etc. Evaluative design and approach to these evaluations, therefore, tends to be developed on an individual basis rather than to follow any prescribed pattern.

Program Loan Evaluations: Program loans finance imports into less developed countries. When a second-year program loan is under consideration, an evaluation of the first year's experience is required. This usually consists of ascertaining the extent to which agreed-upon policy changes on the part of the borrower were implemented, and an analysis of the impact of the imports. For example, the imports might have been designed to keep industry working at or near capacity; this is a target which can be measured.

Multi-Country Evaluation Studies/Spring Reviews

Comparative evaluations can reveal important causes or effects which are obscured by conditions peculiar to individual countries. They can cover comparisons within a single geographic region, or around the world. Although there is danger in assuming that what has happened in one country will necessarily happen in another, presumably more confidence can be placed in findings based on experiences drawn from five different countries than from knowledge of a single country. There is a certain safety in numbers (a simplistic way of expressing faith

in statistical analysis), and good reason for A.I.D. to draw upon and intelligently apply lessons from its worldwide activities. An intensive investigation of the A.I.D. experience should provide answers to many of the crucial questions concerning the process of accelerating development.

A characteristic of these evaluations is that they cover the record of a number of years. Comparative evaluations are usually not undertaken until results attributable to the project can be expected to appear. In fact, the longer the historical perspective the better, although the problem of trade-off between additional years and record quality presents itself.

Finally, these evaluations offer a mechanism for bringing lessons of the past to bear on questions of efficiency. By studying several projects which used different means to accomplish similar purposes, it is possible to arrive at conclusions about relative costs and the effectiveness of the different methods that were used.

Within the three-phased evaluation approach -- effectiveness, significance, and efficiency -- comparative studies can play a particularly important role in evaluation of effectiveness and significance.

Multi-country evaluation studies which have been carried out in the past include a study of A.I.D.'s use of program loans to influence the economic policies of developing countries; an analysis of building extension services in Latin America; and a worldwide evaluation of malaria programs. A number of important issues are amenable to this type of analysis.

A special kind of multi-country evaluation is the Administrator's Program Evaluation Reviews (popularly known as Spring Reviews). These began in 1967. They were designed to coordinate the resources of AID/W offices and the Missions for evaluating program areas of high priority. They concentrate on the historical record, with a view to applying the lessons of the past to improve A.I.D. programs in the future. Sometimes the Reviews look at development experience beyond that of A.I.D. For example, the land reform review examined experience in thirty countries, about half of which had not received A.I.D. help on the problem. These reviews have ranged from comprehensive studies involving many months of intensive

preparation, and outside expertise, to studies of a more narrow scope conducted by a small group of in-house staff. The former were each culminated by a three-day conference involving several hundred people for both A.I.D. and the public; the latter were culminated with half-day, in-house review sessions. Most of these sessions have been chaired by the Administrator. The findings of the reviews are widely circulated, and program policy makers are encouraged to apply the results and findings to A.I.D. programming decisions.

All the conferences to date have been conducted in Washington. There will undoubtedly be experimentation with the design of reviews in future years. Meetings may be shifted to the field; they may be divided by geographic region and further split into working sessions that are aimed at practitioners and informative sessions that are aimed at decisionmakers.

Special Evaluations of Assistance Techniques and Policies

Some important evaluation studies look at problems and issues which are related to A.I.D. projects and programs, but which do not focus on these as the unit of analysis. They include such questions as those concerning the effectiveness of certain techniques of administering or delivering development assistance, e.g., use of Participating Agency teams versus direct-hire personnel, the effectiveness of loan-financed technical assistance, the upward mobility of returned participants, or principles and doctrines of aid. The latter could cover for example, historical analysis of the advantages and disadvantages of coordinating with other donors, of multi-lateral aid, or of the benefits that can be attributed to making aid contingent on self-help.^{2/} Many of these issues are perhaps best suited for scrutiny at the AID/W level; where they have worldwide applicability they could in fact be evaluated as one of the multi-country Spring Reviews described above. Individual Missions, especially larger ones, may however, find it profitable to engage in such analyses of assistance techniques and policies.

Participant training activities are usually carried out as an integral part of a technical assistance project in a

^{2/} Some of these questions may result from entries in the Assumpt s column of the logical framework.

functional area, and therefore regularly evaluated under project inputs and outputs in PARs. Similarly, special evaluations cover participant training whenever they are done for technical assistance projects of which training is a component.

The Office of International Training in AID/W has pioneered a systematized form of worldwide evaluation covering the overall participant training process. Structured questionnaires provide the basic data that are then analyzed by the statistical techniques used in survey research. An entry interview shortly after the participant arrives in the United States supplies information on such points as his selection, his predeparture orientation and other preparation, his language capability, and understanding of his training program. At mid-point in his training, he completes a questionnaire which is designed to call attention to any difficulties he may be encountering. After his training has been completed, he is given an exit interview. Special reports on the exit interviews are issued from time to time, in addition to periodic reports. Evaluation studies are also done at various training facilities to determine the facilities' effectiveness. In addition, a Returned Participant Follow-up Activities Report is submitted annually by the Missions, which provides a source of data on utilization of training. Almost all the follow-up activities are behavioral indicators which lend themselves to quantification. (For example: How many requests for technical literature were made? How many returnees requested and/or took supplementary training? How many returnees trained others in the new technology they had learned?)

The most comprehensive evaluation of participant training as a technique of development assistance included interview data compiled for participants from thirty-four countries. The findings were published as country reports, four regional reports, and a global combination issued in 1966, entitled, AID Participant Training Program -- An Evaluation Study.

Chapter V

DESIGN OF EVALUATION STUDIES

Find out the cause of this effect
Or rather say, the cause of this defect,
For this effect defective comes
By cause.

Hamlet,
William Shakespeare

Probably the most difficult portion of any evaluation study is the initial phrasing of the question to be asked. If the wrong questions are raised, or the problems are not adequately identified in the first place, time and effort may be wasted in coming up with irrelevant answers. When a decision is made to undertake a study, the following questions must be asked:

Why is the study to be done?

What is to be learned?

Who wants to know?

How is the study to be done?

Where is the study to be done?

When is the study to be done?

The answers to why, what, who, how, where, and when will help shape the phrasing of questions, and will help ensure that whatever study plan is devised, it will reflect realities.

The kind of question raised may sometimes run into conflict with the program policies of management. The potential for conflict is greatest when questions concerning the why of things are asked. This kind of question challenges the most fundamental premises, while the how questions pertain only to methods

or techniques used within existing premises or policies. Decisions frequently must be made in the context of administrative or political pressures which are important to consider in the design of the evaluation study.

There is an underlying philosophy of "operationism" in most social sciences which requires a problem or question to be stated in such a way that one has to specify the operations or measures to be taken to define the concept and to provide an answer. For example, the typical example of meaningless scholasticism was the question "How many angels can dance on the head of a pin?" But a more modern question such as "Are we getting any Title IX effects out of the 'such-and-such' project?" is also non-operational. It should be rephrased into a question such as "Was there popular participation in the decisionmaking, the carrying out, and the sharing of benefits in the 'such-and-such' project?" This question in itself leads to other specific questions: "How is popular participation measured? How is decisionmaking determined? How are the dimensions of carrying out a project fixed? How does one quantify the sharing of benefits?" If a question cannot be so stated -- forget it. Restate it so that it is realistic and meaningful. State it so that the operations required to measure it are clear.

Criteria for Designing the Study

Evaluation's primary purpose is to assist management to fulfill its decisionmaking responsibilities. Evaluation studies should be designed to meet the following criteria:

- **Objectivity:** Evaluation activities must minimize subjectivity and must be as candid and factual as possible.
- **Timeliness:** Evaluation studies must become available to management on a timely basis, whether designed to provide feed-back to an ongoing project or information in connection with other activities.
- **Applicability:** The study must produce operationally useful conclusions or recommendations.
- **Communicability:** Findings should be amenable to "translation" from academic language or techniques, into a form readily understood by those who will use the study's results.

- **Validity:** The design of an evaluation study must adhere to principles that assure the reliability of the data being gathered. Collection and processing of the data should be appropriate to the design of the study and to the conditions under which the study was conducted.
- **Scope or Depth:** Evaluation should measure not only progress or quality of performance in a project, but should also seriously question the premises on which the entire project is based. (This point, often overlooked, appeared in connection with a recent study of a malaria program. In the past, rigorous evaluations had been carried out by epidemiologists and other specialists, but only late in the process was the strategy questioned: Was the conventional strategy of attack, consolidation, and maintenance practical in a country with a rudimentary public health infrastructure? In another instance, evaluators found an agricultural institution project effective in meeting its purposes, but the project purpose had become outdated in terms of national needs.)

A Basic Study Design

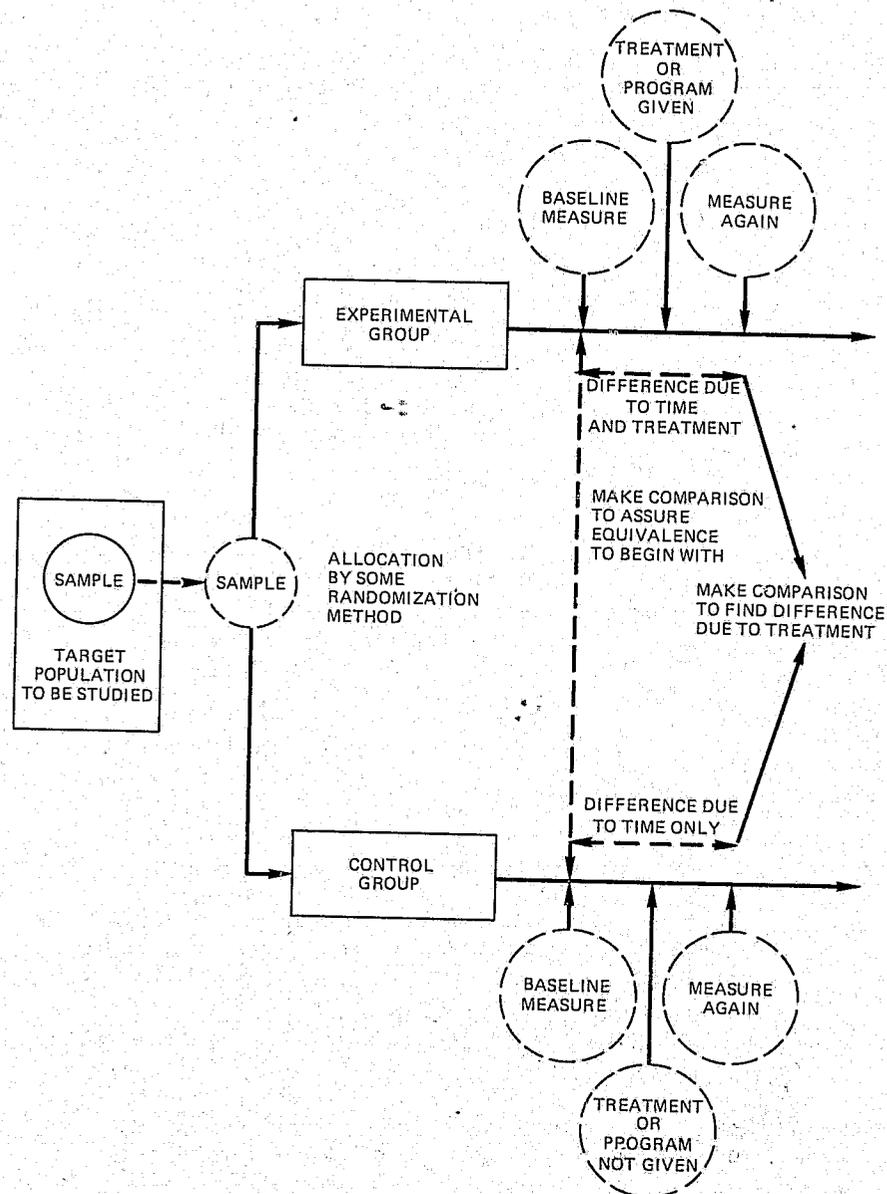
In the design of a study, care must be taken to show comparisons clearly -- i.e., not to confound or confuse the elements with extraneous matter. To accomplish this, a study should be so designed that when comparisons are made, the results are clearly attributable to one or the other of the factors involved. This cannot always be done. Real-life situations tend to be complex and to be made up of interacting factors. If this is the case, conclusions should honestly reflect what is happening -- including the confusion. The best method is to try to control as many of the factors as possible and to let only one or more factors vary except in instances in which multiple correlations are possible.

Figure 2 shows a basic research design to which almost all other study designs are traceable. There may be all sorts of variants to the logic which this diagram pictorializes, but the logic remains fundamentally the same. It is a means of contrasting one variable with another while all other factors are considered equal -- or at least kept under a form of control.

The design of the study should indicate the approach to data gathering to be used -- e.g., use of regular operating reports, field reports, field surveys, interviewing, administering of tests; the type of experimental design -- e.g., control group,

FIGURE 3

"IDEAL" STUDY DESIGN FOR MAKING COMPARISONS



before and after, and whether "treatment" with some kind of program is involved. It should also define the group to be studied and how a sample is chosen. These factors influence the kind of statement that is made at the end of the study -- how general it can be or how specific it may have to be.

The diagram shows a particular target population selected for study and a sample taken from that population. The sample next is divided into two groups by a scheme which assumes that the factors in the groups which might influence the results have, if not an equal, at least a probable chance of occurring in both groups. Tests are given, or baseline measures are taken, in both the experimental and control groups. This comparison is made to assure that the two groups are similar at the beginning. If there are differences, at least the differences are known. Then one group receives "treatment" or program input, and the other does not. The same measurements applied at the baseline are applied again after the "treatment" has had time to take effect. Then three more comparisons are made:

- (1) The experimental group is compared with itself before and after "treatment";
- (2) The control group is compared with itself before and after the "nontreatment" period;
- (3) The main comparison is really a comparison of the comparisons ($3 = 2 - 1$).

Following are the basic steps in designing and carrying out an evaluation study:

- State the problem.
- Select the standards or criteria against which judgments are to be made. What do you hope to accomplish by the end of the project (or have accomplished at the time of the evaluation)?
- Identify the indicators which will permit measurement of the changes to be brought about. (The criteria and indicators should be found in the second and third columns of the logical framework matrix if the activity being evaluated has earlier been analyzed in accordance with the matrix.)
- Collect data on indicators, including baseline data if not already available.

- Analyze data for (1) rates of change,
 - (2) direction of change,
 - (3) nature of change,
 - (4) amount of change.
- Interpret the data analysis:
 - (1) Was the planned purpose (or intermediate target) accomplished?
 - (2) Did it make a significant impact on broader development goals?
 - (3) Was it worth the cost and effort?
 - (4) What lessons are there to be learned?
 - (5) What were the critical factors that determined the outcome?

This basic study design is admittedly just that; regrettably, it cannot always be duplicated.

The basic design for comparative study is similar to the logical framework used in appraising projects. The line showing the experimental group can be read as "approved A.I.D. project" and the baseline measure is essentially the Beginning-of-Project-Status (BOPS). The "treatment" or the program given for comparative study is essentially the same as the input/output phase. The point at which measures are again taken is essentially the same point at which the End-of-Project-Status (EOPS) is measured.

There are a great many reasons why it may be necessary to modify this basic study design. Economic assistance programs are developmental in nature rather than controlled laboratory experiments. Furthermore, factors independent of the "treatment" may act as agents of change during the reform period, and the very fact that a test is under way may influence the outcome. Political and administrative circumstances may inhibit setting up control units for programs of a social or economic nature, and it is obviously impossible for social action programs to achieve experimental isolation comparable to the conditions in a laboratory or even to the conditions in agricultural test plots. Even when the ideal cannot be reached,

however, judicious planning will allow the evaluator to obtain the maximum possible benefits from evaluation activities, provided the pitfalls are recognized.

An example of a comparative study with controls in the education sector is that carried out by USAID/Guatemala. To test alternative strategies, two schools were provided with special classroom equipment and with the services of technicians. One was in an Indian-speaking area and the other in a Spanish-speaking area. These two schools were compared with two established control schools where the same languages were spoken but in which no innovations were introduced. In order to be sure that the students of the four schools were essentially equal educationally, baseline measures were taken of such factors as teacher training, pupil-teacher ratios, and level of pupil achievement. After that, any differences found in attendance, drop-outs, promotions or achievement levels might be traced to the innovations. But which innovation? The special facilities? Or the technicians' services? To clarify this point, two more experimental schools were planned with the same baseline measures and technician services, but without specially constructed facilities. At the end of the study, comparisons will be made of the attendance records, drop-outs, promotions, and educational achievement to determine the schools with the best records.

This method can help to determine the effectiveness of our inputs or treatment. Conversely, if the same changes occur in the control group, we must assume that the changes are due to some unrecognized factor and an attempt should be made to identify these.

Other design examples of special evaluation studies are available on "Institution Building" and "Population and Family Planning Programs." In addition, a series entitled Manuals for Evaluation of Family Planning and Population Programs are being prepared by the International Institute for the Study of Human Reproduction, Columbia University, with the support of Ford Foundation and A.I.D.

Suggested Checklist for Planning an Evaluation Study

Objectives

- (1) What is the study (not project) objective?
- (2) Does the study have a potential for providing new (and needed) information? A new method? Technique? Procedure? Policy?
- (3) Will the final results be important or significant for the project or program? Might they change some policy or way of doing things? Would confirmation of validity of earlier expectations warrant the cost of the study?

Methods

- (1) Are the techniques, instruments, or modes of inquiry appropriate to the study design? To the foreign context?
- (2) Will the methods require adaptation to a local condition? Will this adaptation do violence to the design?
- (3) Are there sampling problems?
- (4) If interviewing or opinion-survey techniques are to be used, have the questions been reviewed for meaningfulness in the local language and culture? Good taste? Political sensitivity? Religious connotation? Language problems?
- (5) Will the methods gather more data than are required? Less? That is, are they efficient, economical, and effective in terms of the goals of the study?

Data Processing

- (1) Are the procedures for the statistical manipulation of the data stated clearly? Is there a clearly conceived plan for the analysis that will be done once the data have been collected?
- (2) Have statisticians or ADP systems experts been consulted regarding the program to be used?

- (3) Are the analytical procedures likely to produce meaningful statements?

Analysis and Interpretation

- (1) Have a wide variety of potential findings been considered?
- (2) Does the logic or design of the study permit clearly stated generalizations?

Costs

- (1) Are the dollar costs for the evaluation study reasonable for the various categories (personnel, travel, supplies, overhead, etc.)?
- (2) Are local currencies being used to the maximum extent possible?
- (3) Are there luxury or unnecessary items in the budget?
- (4) Has the budget estimate omitted consideration of some item (services by foreign personnel, differences in living costs from one place to another, etc.)?
- (5) Are the total costs proportional to the scope or importance of the study? Is the study worth the investment? Will the study cost more than its results might save?

General

- (1) Will the study answer the questions it set out to answer?
- (2) Will it produce explicit and usable results?
- (3) If it is not completed, will there be salvage value?
- (4) If the study is completed -- THEN WHAT?

The Selection of Evaluators

The selection of the evaluator(s) is of paramount importance to the success of the endeavor. Should the work be done by in-house or outside personnel? Once this decision has been made, where can the appropriate evaluator(s) be located?

The value of the program evaluation process is in direct proportion to its use by management in planning and implementing projected and ongoing programs. Evaluations carried out by, or under, the direction of action offices are most relevant to their needs and the findings are more likely to be accepted and applied. This placement of responsibility, however, poses several problems. Action office personnel may find it difficult to be objective, they often lack time, and they may not be acquainted with data gathering and analytical techniques. Various approaches can help overcome such difficulties. Consultants (outside individuals, headquarters, officers or contractors) help provide objectivity, time, and expertise. Missions can organize special task forces which take advantage of skills available in university or Participating Agency teams or in AID/W, and joint evaluation with cooperating governments can provide additional manpower for data gathering.

Some of the pros and cons involved in using consultants are:

- One of the primary problems is to minimize subjectivity. Consultants in specific functional fields may have a strong bias one way or the other; however, disinterested consultants should be able to offer greater objectivity in the evaluation of a project.
- In most cases, the consultant will be handicapped by his lack of familiarity with the project or program and the country or Mission perspective. Unless familiar with prevailing local conditions and customs, the consultant-evaluator is likely to encounter difficulties and unexpected delays in the design and conduct of an evaluation study.
- The consultant may be able to bring into play specialized knowledge and familiarity with different techniques and fresh viewpoints which are not otherwise available.
- Consultants may also be able to assemble a staff of varied and cross-disciplinary expertise which cannot readily be matched within the organization.
- The effect on the host government of recommendations by a recognized non-U.S. Government source may be greater than the effect of those coming from U.S. Government sources. A consultant may be able to prepare and present a more frank and candid report than an agency of the U.S. Government.

Basis for Selection

The selection thus comes down to the type of study desired and the information or data to be derived. Problems likely to be encountered and basic qualifications expected from the evaluator(s) (such as language, knowledge of local conditions, technical expertise) should be spelled out in detail. On the basis of this information, an intelligent selection can be made, not only between possible groups of evaluators, but also of the individual(s) from within the group. In addition, this information will help provide potential candidates with an understanding of what is expected.

In choosing a consultant for an evaluation study of narrow scope, or one encompassing limited technical aspects, a perceptive and inquisitive observer from outside the discipline may be able to make a valuable contribution by challenging basic assumptions and bringing a new perspective to the task. This consideration increases substantially the sources of evaluators, especially in the case of in-house or locally available personnel.

Combinations of In-House and Outside Experts

These considerations should not be construed as forcing a choice between in-house and outside experts. In fact, a team consisting of A.I.D. personnel and outside consultants provides many advantages, e.g., the fresh outlook and objectivity of the outsider and the familiarity with the project and/or area, as well as the A.I.D. perspective of the direct-hire employee.

Sources of Evaluators

In-house evaluators can be drawn from the office responsible for the project, another Mission, or AID/W; Participating Agency personnel; U.S. university or contract personnel in the area; a task force of experts formed from a combination of the above groups, with the Evaluation Officer serving as an advisor and ex-officio member. The AID/W geographic bureaus provide assistance in recruiting outside evaluators. Potential sources include the group of consulting firms under contract with the AID/W Program Evaluation Office, other past and present A.I.D. consultant and contractors, professional organizations, international organizations, U.S. Government agencies, roster of retired U.S. Government employees, U.S. university personnel independently in the area, third-country experts, etc.

Consultants, The Care and Feeding of

If the services of an outside consultant are retained, the action office should undertake the following steps to maximize his contribution:

Briefing of Consultant -- As a means of bringing into focus the evaluation study specified and to make the maximum use of the consultant's time, he should be given a detailed briefing document prior to his beginning his task. This document should contain the following categories of data:

- Project background and history;
- Project and sector goals;
- Operating strategy of the project to date and anticipated strategy, including the assumptions about conditions or actions of other interested parties;
- Project operations;
- Reasons for making an evaluation;
- Scope of evaluation to be carried out;
- Extent of cooperating government participation and contracts.

In addition to this briefing document, the consultant should also be given a document, prepared in cooperation with the action officer, executive office, and other interested offices, which outlines in detail the logistic support that can be provided and the facilities available to him (e.g., housing, transportation, PX and commissary privileges, etc.).

- Finally, special care should be taken to acquaint the consultant with the concept and methodology of A.I.D.'s annual noncapital evaluation process. While the consultant's specific assignment may not cover all aspects of the project, an acquaintance with the system and the total project design will help him to formulate his recommendations in such a manner that they can be integrated into future, regular in-house evaluation efforts.

Mission Participation and Liaison with Consultants

The Mission should designate a counterpart (e.g., the project manager) as liaison officer responsible for keeping abreast of the consultant's work, and assuring that all relevant data are made available. In addition, there should be periodic review sessions between the consultant and appropriate A.I.D. personnel to check the consultant's progress and to discuss the direction of his efforts. It is the responsibility of the liaison officer to follow through on proposed changes after the departure of the consultant, as well as to facilitate his work, to assist him in overcoming local problems and to prevent any duplication of efforts. A substantial input of Mission or AID/W skills in the course of the evaluation is desirable.

Timing and Submission of the Report from Consultant

The consultant should be held to a mutually agreed-upon, realistic schedule. Except when clearly not possible (as in the case of collected data being analyzed by computers at the consultant's home institution), he should be required to submit his report (or at least a good draft) prior to his departure from the Mission or AID/W office.

Analysis of Data

If data are to be analyzed by statistical techniques which may also involve use of a computer, a statistician or ADP systems expert should be consulted early in the evaluation. He may want the data to be collected or to be expressed in a particular form; he can frequently suggest shortcuts in data collection, provided that the information desired on completion of the analysis can be delineated. This may save much effort because people frequently collect far more data than is needed. It may also be necessary to describe in detail the methods by which the data were collected and the procedures used in obtaining the sample. In both cases, errors may have occurred. The statistician may be able to correct for some of these; however, he should be aware of what happened in the data collection stage so that if errors are present to begin with, they will not be compounded during the analysis. In this era of the information explosion, there are many spurious reports because data were collected and analyzed without a validity and reliability check.

Preparation of the Final Report

It is expected that when a special study has been completed, a report telling what was done, how it was done, and containing conclusions and recommendations will be written. It is often helpful to draft a preliminary outline before the study begins. Drafting such an outline will help to clarify the thinking of the evaluator as to what should be done, how it should be done, and the kinds of problems involved. Care must be taken that the outline is used only as a device to help plan the study.

When the initial proposal for a special study is made, the proposal is questioned from the standpoint of why, what, who, how, where, and when. When the study has been completed, the final report should cover similar points. It should state clearly and succinctly

- Why the study was undertaken. Every effort should be made to be explicit in the rationale so that others may understand the reasons for inclusions or omissions in the study.
- What the problem was.
- Who performed the study.
- How the problem was studied. What procedures were used. What information was collected. How were the data analyzed. How were the data interpreted.
- Where the study was carried out.
- When the study was carried out.
- The final question to be answered in the report is, SO WHAT? State the conclusions clearly and concisely, and recommend the next steps to be taken.

Chapter VI

MEASUREMENT, DATA COLLECTION, AND ANALYSIS

When you cannot measure what you are speaking about, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science, whatever the matter may be.

Lord Kelvin

Measurement provides a means of replacing qualitative distinctions with quantitative distinctions. It introduces precision into judgments. Of course, the mere act of assigning numbers can lead to all sorts of errors. The most serious of these is the common belief that the differing degrees of a particular quality always bear the same ratio as the numbers assigned to them. (For example, is a day when the temperature is 100° twice as hot as a day when the temperature is 50° ?)

Another kind of error is the belief that certain kinds of A.I.D. operations cannot be quantified at all. At present, for many of our non-economic programs, this may be so. Institutional growth and maturity, expansion of human skills and knowledge, the adaptation and transfer of technology, are exceedingly difficult to pin down. However, they provide a challenge to creativity in a problem area where much innovation is needed.

Another common error is the belief that direct measurements can be made of the phenomena observed. This is not always so. Usually, manifestations or indices of these phenomena are observed and measured. For this reason, the selection of indicators become critical. Indicators are selected because they are the manifestations of output or change per se, or

because they are considered equivalents or representations of the output. When they are the latter, they serve as proxy or surrogate indicators which stand for the real thing. To know whether the indicators have accurately measured what they are supposed to measure, validity must be considered. To know whether the measures are dependable measures, reliability must be considered.

- Validity refers to the degree with which a measure or indicator actually does what it purports to do.
- Reliability refers to the degree of consistency or dependability with which results will be obtained upon successive applications of the measure.

Both concepts are necessary to provide an estimate of the degree of error in our measures. Without them, there will be errors anyway, but their existence or magnitude will not be recognized.

The threats to validity and reliability are many, and great care must be taken to spot them because they may occur when and where least expected. An example of a test influencing the outcome is found in the famed "Hawthorne" effect, named after a Western Electric plant of that name. In the course of a study of environmental factors affecting productivity, it was found that productivity improved not only when lighting was increased, but again when lighting was decreased; the workers were pleased by the attention of the management. Such threats to validity can be mitigated by the use of control units, which are included in the test, but receive no actual input to produce change. Well-known instances of this approach are medical experiments requiring a placebo.

The Land Tenure Center of the University of Wisconsin has pointed out that the first conclusion about the effect of land reform on production in Bolivia was that production decreased for a few years and then increased. Now scholars are not so sure. The apparent early decrease in some regions may have occurred because the newly independent farmers avoided the use of middlemen in marketing. The observers were not gathering data on the independent farmers; they were looking for the traditional proxy indicators of production by collection of sales data from established wholesalers. Some interviews with

representative farmers might have revealed the realities.

Measurement methods may vary between the two units compared. For example, the safety records of two similar factories differed. The factory with fewer reported accidents had first-aid kits throughout the plant. Hence, the only accidents reported were the more serious ones that required a visit to the nurse. The factory with more reported accidents, prohibited first-aid kits in the plant and thus forced all injured people to visit the nurse.

Similar threats to validity occur when there are changes in the means of measuring the effects of the program. For example, law enforcement, accident prevention, disease prevention or other "drives" are often accompanied by improved record-keeping. There may then appear to be an increase -- more crimes or accidents -- simply because the new reporting system does not miss as many cases as the old reporting system. This threat should not be used as an excuse to defer improved records; rather, the inability to make comparisons should be recognized.

Data Collection

Project planning and evaluation both require data before either function can be performed. If project planning and evaluation are to be improved, objective data must be substituted for intuition. Data can be as varied as the number of farmers who planted the new high-yielding variety of rice; the amount of fertilizer, pesticide, and water used; or how much was paid to the landlord for rent, to the bank for credit, to the merchant for seed, or to others for storing, milling, and marketing the harvest. All these are data, whether expressed in hectares, pounds of fertilizer, piasters, baht, or pesos. The first problem in data collection is to specify the data that are required.

If evaluation is to be built into the project, the best data to be gathered are the kinds of information needed by the project manager for project operations. But with a view to their being used as evaluative data, they should be couched in terms of output indicators.

Direct Methods

Even in less-developed countries where statistical services are not very well developed, there are likely to be substantial sources of data which are often ignored. One problem with their use, however, may be that the method by which they were collected or the scope of problems they cover, was determined on the basis of purposes different from those now to be served. On occasion, it may be possible to modify the data collected. It must further be recognized that LDC statistics are often of questionable reliability and must be used with caution. This, of course, is equally true of statistics developed solely in connection with a particular project, although the method of collection may provide an indication of the degree of trust the data merits. Thus, an effort to obtain one-time baseline data may require combing through source materials. This method of collection is likely to improve reliability. On the other hand, to obtain regular progress data, it will usually be necessary to rely on the routine data collection of others. These data may be less reliable as a result of efforts to "look good", overwork on the part of statistical personnel, etc.

Available Data. The following brief list will illustrate the kinds of information recorded by government agencies or private organizations. It is not exhaustive. See Appendix C for selected output indicators which have been used for various subjects.

- Public records: Vital statistics on births, deaths, marriages, divorces; school attendance; arrests; court convictions; prison records; taxes and customs collected; welfare payments; bridge and highway toll receipts; automobile registrations; etc.
- Private Organizations: Union records; farm co-op records; business payrolls; factory production records; shipping records; warehouse inventories; bank deposits; credit institution loan applications and approvals; truck company records; railroad passenger load; freight car loadings; hospital and insurance company data; import licenses; store sales; market prices, etc.

In addition, U.S. Embassy attachés collect and report data to Washington. USAIDs can probably also arrange to obtain data collected by other donors of foreign assistance, the UN family of specialized agencies, multilateral banks, regional councils, Ford, Rockefeller, and other foundations, and voluntary agencies.

Direct Observation. This can be costly and time consuming. It has the advantage of not being dependent on the availability of persons willing to cooperate or capable of reporting the desired information. It also may permit the observer to stay out of what is being observed, although there are techniques for becoming a participant observer.

Questionnaires and Interviews. These usually require highly skilled specialists in order to collect valid and reliable data, and to avoid collecting a good deal of spurious information. There are ample reference works. USAIDs should rely on these and on specialists wherever surveys, opinion polls, or attitudinal studies are needed.

Indirect Methods

In less-developed countries where it may be difficult to obtain a population census, an interviewer who queries a farmer about his last year's income or rice harvest might immediately encounter cultural or other problems. The farmer may not be willing to report these data accurately. He may suspect the interviewer of being a government agent who will eventually raise his taxes. Whether meeting willingness or suspicion, these attitudes too constitute data which have to be taken into account; they not only influence the kind of information the farmer gives, if any, but may determine whether he responds to a technical assistance effort at all. When obstacles of this sort arise and data cannot be obtained directly, it is sometimes possible to do so indirectly or by proxy.

Estimates. These are personal judgments. They are sometimes, but not always, reasoned judgments and it is not possible to place the same degree of confidence in them as in objective facts. Nevertheless, decisions may have to rely on the best estimate which can be made.

Guesses, Conjectures, or Surmises. These are opinions or personal judgments based on insufficient evidence; confidence placed in them is still lower. Decisions made on the basis of guesses may be entirely random. If statements have little evidence to back them up, it is best not to try to quantify them.

Other Indirect Methods. When farmers cannot be counted directly, it may be possible to substitute a method in which something else is counted, and from logical deduction and inference, to gain an estimate of the number of farmers. For example, aerial photos of hectareage under cultivation are taken; the average number of hectares per farmer is assumed; and the number of farmers is deduced. The average number of hectares per farmer is assumed on the basis of what is known about the number of hectares per farmer from another part of the country; this would be a reasonable but not necessarily accurate assumption.

Examples of other substitute methods of counting farmers include the following: Compile from agricultural bank records the number of farmers who requested loans. (Some may not have asked for credit and thus will be missed.) Land title records will give owners but not tenants. (Then, names of tenants will have to be requested from the owners.) The miller, the fertilizer salesman, the storage warehouse, the farmers' cooperative, and other groups dealing with farmers will have slightly different numbers of farmers with whom they deal. All taken together will permit the best estimate with the minimum of error.

Other problems in the field hamper collection of data directly. Illiterate persons cannot complete questionnaires themselves. Different languages or dialects in the same country compound interviewing problems. USAIDs are understaffed and trained counterparts cannot be found. There may be travel restrictions. Aerial photographs are too expensive. The invasion of privacy of the family is forbidden, etc.

One Mission which had protested to AID/W that the data collection problem was practically insurmountable in the cooperating country later realized that an impressive amount of data could be gathered by exercising ingenuity. The food and agriculture officer hired local moonlighters to gather information on market retail prices in the bazaars. The field extension advisors obtained samples of crops produced in different parts of the country and noted the prices farmers received for their harvest. A Participating Agency economist interviewed farmers on farm costs and income. A scholar on a university contract team collected data on a rural family budget on his own time, and made this available to the Mission. An ILO advisor arranged for a sample survey of the labor force using

as interviewers, local high school girls who returned with good answers on the number of such people in households. An engineering team promoted the establishment of an advisory committee from industry. A highway engineer arranged for traffic counts on major market roads. A visiting graduate student had done research on land tenure. In some less-developed countries there may be more data gatherers than are suspected; e.g., local libraries and universities, research firms, professional societies, and public and private educational agencies. The point is that in many cases the data are already there; it's a matter of pulling these data together.

Dimensions of Progress

The evaluator is faced with the need to establish tangible indicators of the changes that are occurring over the life history of the project. While the changes can be observed, there is no way of sampling the dynamic process itself. It is therefore necessary to fall back on the next best substitute, namely taking two static measures -- the before and after situations -- and inferring the in-between situation as a changing one. A combination of baseline data and indicators will in most instances provide the evaluator with the necessary information:

Baseline Data. These data provide information about the status of things at the start of the project or BOPS (Beginning-of-Project Status). These data become the "fix", zero point, anchor point, or benchmark against which later measures will be taken.

The establishment of baseline data can be simple or complex, depending on the circumstances and the project purpose. Thus, for example, if the project seeks only to increase numerical output of a given kind, and provided that adequate statistical data are available or can be procured, the establishment of suitable baseline data will be relatively simple. On the other hand if, as is frequently the case, a project seeks to effect certain qualitative changes, the establishment of suitable measurement data becomes more difficult. One way of dealing with this problem is to establish rating scales as a means of determining baseline measurement. (See Appendix B for rating scales for housing development and community development. These are intended as suggestions only.)

Indicators. These are variables in the cooperating country situation which indicate change in the areas treated (either directly or indirectly), and which lend themselves to simple quantification to indicate a magnitude. These variables can be used to measure performance.

The selection of baseline data and indicators is of course governed by the changes that are sought or anticipated. In planning for evaluation, the project planner or evaluator must ask himself the following questions:

- What changes are anticipated?
- What will the end-results of these changes be?
- How are these end-results to be indicated in the future?
- What data are available at present which resemble the indicators? (And which can increase, improve, grow or change into the future indicator?)

Appendix C shows a list of selected output indicators which have been used in various A.I.D. projects. The elements of variables in the cooperating country situation considered changeable have been identified, and a simple quantification of each element is issued to indicate a magnitude; e.g., graduates per year. There is a tendency to confuse progress in marshaling inputs, with progress towards output targets. There may be an output target of doubling the enrollment of a vocational school. This increased enrollment will require new buildings. Counting the number of additional classrooms built is an input measurement; counting the numbers of students is an output measurement.

Within the context of the noncapital project evaluation system, separate measures of indicators are required on the output and purpose levels. However, the latter may prove considerably more difficult to quantify and thus require other methods of verification. For example:

Output Level

Houses sprayed

Skill training provided

Business loans made

Family planning clinics established

Textbooks printed

Examiners trained

Fertilizer distributed

Purpose Level

Malaria reduced

Employment obtained

Exports increased

Birthrate reduced

Education improved

Increased taxes collected

Crops increased

Indicators may be used to measure significance if they are used to compare what happened with a goal other than the project target. For example, to determine whether 100 graduates per year in an education project has any significance for the cooperating country economy, one must compare that output indicator with a goal pertaining to the entire education and human resources sector in that country, or to other sectors. Such a goal might be found in the national manpower survey. For Nepal, 100 graduates per year may be significant; for India, it may not be. Inter-country comparisons may also help in judging significance. For example, if 100 graduates per year in India only adds to the ranks of the unemployed intelligentsia, the first conclusion may be that India is educating too many people. But international comparison will show that Korea and Taiwan have a higher proportion of educated people and a lower rate of unemployment. The problem in India may be the type of education or the nature of the labor market.

The amount of change or progress is measured by examining the indicator in relation to the life span of the project. The simple indicator "number of graduates per year" becomes meaningful only when the number of graduates this year is compared with the number of graduates last year.

Indicators may be used to measure effectiveness if they are used in such a way as to compare what actually happened with what was expected to happen (project targets). They may also be used to measure efficiency if they are used in such a way as to show the cost per unit in relation to the benefit

accrued. Suppose a project goal was to turn out 100 graduates per year and that actually only 92 persons graduated. Suppose also, that the annual project costs, not including initial capital expenditures, could be expected to amount to \$560,000. To oversimplify, the effectiveness was 92 percent, and the cost can be stated most simply as \$560,000 divided by 92, or \$5,097 per student. Is that efficient? To answer this question, information is needed on the usual cost per student for this type of school (medical, law, or teacher training, etc). If experience factors show it should cost only \$3,000 per student, the school is expensive and is thus inefficient. Either the cost has to be reduced, an increasing number of graduates have to be turned out at the same overall expenditure, or some other vehicle for the training of the required number of students must be developed.

Non-economic Indicators

The emphasis on development by A.I.D. and its predecessor agencies has been preponderantly on economic growth and development. This is evident in the A.I.D. staffing patterns, in the way A.I.D. is organized to provide capital and program assistance, and in the procedures whereby program decisions are made and priorities determined. These latter are largely in terms of the impact that projects may have on increasing the Gross National Product (GNP) of a particular country.

However, the Foreign Assistance Act, as Amended, in 1969 clearly gives political and social development a comparable priority with economic development. Efforts are now being made to develop indicators which will permit measuring the effectiveness, efficiency, or significance of projects in terms of impact on the social or political aspects of a country's development. Part of the problem encountered lies in the state-of-the-art of the social sciences. Theory and doctrine involving socio-political phenomena generally are described in qualitative terms. We are only beginning to quantify such matters as social concerns or political affairs.

Considering the time taken by economists to devise methods of accurately measuring GNP as an index of economic growth, a similar approach should be attempted for the social and political aspects of growth, e.g., an equivalent of GNP such as Net National Welfare (NNW). A.I.D. has devised social

indicators which are shown in Appendix D. These are designed to be incorporated in a country analysis to evaluate civic development activities. They permit a systematic consideration of social development and popular participation, and can be used in developing program priorities and objectives.

These macro and sectoral indicators focus on the population's access to resources (land, credit, education, etc.) and change in this access over time, rather than on the more conventional aggregate measures which assess levels of living or welfare (health, nutrition, literacy, per capita GNP). Some of the latter are, however, included. Level-of-living averages can conceal gross inequalities. The primary purpose in selecting these indicators is to obtain a better picture of the extent to which different groups in the society have opportunities to participate. Income distribution would be one of the best indicators for this purpose, but because data on this subject are scarce, this has not been included. If income distribution data can be obtained, this indicator should be added.

In this section, an attempt has been made to show the relevance of data for social development and popular participation. Overall, the data should help in the Missions' analyses of four factors essential to determining the need and priorities for increasing popular participation as an objective of the A.I.D. program:

- The pattern of modernization and its effects; i.e., what sectors are most affected (either positively or negatively) by the spread of modernization and in what ways?
- Which groups seem likely to be affected adversely by present trends (e.g., small farmers, wage earners, professional people)? Over what lengths of time?
- What opportunities are open to these adversely affected groups to redress the balance (e.g., increased access to credit, effective unions, more jobs in the cities, labor-intensive rural public works programs, etc.)?
- What changes in cooperating country development plans and/or programs are necessary to promote broader access to resources and opportunities? How feasible are such changes?

Knowledge of these four factors will allow specific A.I.D. strategy and program recommendations to follow.

Performance Standards

The question arises, Is the degree of change that has been brought about significant? Other ways of asking this are, How much of a difference makes a difference? Or, how much change must take place before it is considered to have an impact on development?

The degree of progress achieved can be labeled minimal or maximal or optimal, in which case the range of progress expected has to be known in advance. Further, to know whether the minimal or maximal change observed should be labeled unsatisfactory, adequate, or satisfactory, other things have to be known. The meaning of unsatisfactory would have to be given in terms of a standard. (For example, an infant mortality rate of 75 per 100 live births might be considered unsatisfactory until it reaches a more tolerable or adequate rate of less than 30 per 100.) Such a standard can be obtained only by collecting the historical experience in various countries and (1) determining the current status of development by using indicators, and (2) making intra-country and inter-country comparisons of these indicators to see where on the scale of comparison a particular country lies. These measures often go beyond the evaluation of A.I.D. activities; they are a step in the direction of assessing a country's total development program. If A.I.D. is only one of several donors, its contribution to development may be difficult to discern.

Once the particular status of a sector's growth in a country is known, the rate of progress in the less-developed country may be seen to be very low or slow as compared to the same sector in developed countries. Once the range of indicators or the rates of growth for a number of countries have been ascertained, they can be used as standards of progress against which to describe a particular less-developed country's growth.

Advantages and Disadvantages of Using Indicators and Standards

If properly formulated and applied, progress indicators and performance standards can:

- Establish that change has occurred and indicate the character, direction, and rate of change;
- Permit comparison of the actual change against that which was planned;
- Permit assessment of the impact of this change on higher goals;
- Compare a project's performance with that of similar projects;
- Allow the examination of the relation of input to output and of cost to benefit.

Indicators and standards have a tendency to cause apprehension and can indeed be harmful if wrongly applied because they may:

- Force the setting of targets more precisely than perhaps they should be set, given the uncertainties of the cooperating country situation;
- Require quantitative measurements when much of the project's concern is with qualitative improvements in human knowledge and skill, institutional capacity, etc.;
- Subject the project's efforts to comparison with other projects and programs which are not comparable because of differences in cultural, economic, political, or other characteristics.

Quantitative vs. Qualitative Measures

Much of what has been said supports Lord Kelvin's contention that when it is practical, quantitative measures are preferable to qualitative measures, and it therefore behooves the evaluator to strive for quantification. However, the central issue in evaluation is not so much one of quantitative vs. qualitative measures, but rather that indicators of change be objectively verifiable, whether they be quantitative or qualitative.

Chapter VII

ISSUES IN PROGRAM EVALUATION

On this very ground with small flags flying, and tinny blasts on tiny trumpets, we shall meet the enemy. And he may not only be ours, he may be us.

"PGGO," Walt Kelly

Of the issues in program evaluation discussed in this section, the one of candor and objectivity is fundamental to all evaluations; the other, that of joint evaluations with cooperating country personnel, is one that offers an opportunity to broaden the scope and depth of project evaluation.

Candor and Objectivity

Candor means forthrightness with the additional sense of freedom from bias, prejudice, or malice. Objectivity means to operate independently and to be capable of making observation or verification by scientific methods.

The current program evaluation system is a somewhat biased one in that project managers take an active role in the evaluation of the projects that they themselves are managing. The important issue here then is to minimize the subjective element. The project must be given as honest an appraisal as possible. Stating facts, with all the "warts and pimples," can be a tremendous advantage. Conversely, there are great disadvantages in not being candid and objective. The facts become blurred with emotional or personality overtones. Decisions cannot be made readily when the facts are fuzzy.

Opinions, beliefs, and values are blended in people's mental processes after long exposure to life experience and

education within a particular culture. Americans tend to view the world through "red, white, and blue colored" glasses. Sometimes there is an awareness of these attitudes, inclinations, ideals, and interests, but not always. As a result, predispositions and values are not visible and cannot be fully controlled. Subjectivity can be reduced by recognizing their existence, and by stating as explicitly as possible what the value premises are.

We do not need to rely entirely on exhortation to obtain objectivity, even with self-evaluation. There are a number of tools at the disposal of the evaluator to assist him in minimizing subjectivity. These include:

- Statistical data to replace conjectures and opinions held by the evaluator;
- Judgments of individuals and groups not directly involved in carrying out the project, such as
 - (1) The local academic community, graduate students, etc.,
 - (2) Persons directly affected by the measures,
 - (3) Consultants,
 - (4) Other A.I.D. offices not directly involved in the project;
- Joint evaluations with the cooperating country government;
- Comparisons with
 - (1) Control groups,
 - (2) Inter-country and intra-country standards.

Joint Evaluations with Cooperating Countries

Development assistance involves working with cooperating countries to add to their own resources a critical margin of additional resources or technical knowledge, so that their development programs will succeed. More and more, A.I.D.'s emphasis is on the cooperating country taking the initiative in planning and in executing plans involving A.I.D. assistance. In conjunction with this, the United States is lowering its donor profile and is thus moving toward greater use of non-government intermediaries in administering assistance.

Consistent with these approaches to development assistance is the thorough-going participation of cooperating country officials in the evaluation of U.S.-assisted activities.

The policy of the Agency for International Development (as it is with other donors), is to encourage joint evaluations. Such evaluations are not required by A.I.D. on the basis that circumstances vary with different types of countries, projects, and personalities. Partly because of these variations in circumstances, Missions have used many different arrangements for involving cooperating countries in evaluations.

As this edition of the Handbook is written, more than half of the Missions have engaged in some form of joint evaluation exercise. Their reports indicate that the effort is generally useful and that most of their original reservations proved to have been unfounded. Conversely, some Missions which decided not to undertake joint evaluations regretted their decisions because the evaluation findings often pointed to the need for action changes by cooperating countries. To convince the governments in later negotiating sessions of the need to undertake such actions proved more awkward than might have been the case in joint evaluation proceedings.

One caveat to the above conclusions needs to be noted. Evaluations can serve several purposes. The most common one of assessing progress and considering how to progress further might often be pursued jointly. But the purpose of planning strategy vis-a-vis the cooperating country should obviously be private. Some Missions have two evaluation review sessions -- one internal and one joint to accommodate these circumstances.

Types of Participation -- The least inclusive form of joint participation is to have informal discussions with responsible cooperating country officers to get their opinions about the activity being evaluated. This should occur frequently. These informal soundings should reach beyond cooperating project personnel to higher officers, including those in planning and budget offices, and to persons and/or organizations whom the activity is ultimately designed to serve.

Another and more comprehensive form of joint participation is joint preparation or review of the project design. The project adviser and his counterpart may meet together with the Mission Evaluation Officer to work out the logical framework. In one such case, the two key project officers spent several

hours actively sorting out the project purpose while the Evaluation Officer sat by. The clarification helped both the USAID and cooperating country project officers. In another country, a group of cooperating country officials and their American adviser went through all the worksheets and redesigned their project in the process. A variation on involving counterparts in preparation or review of project design is where the Americans take a draft logical framework to their counterparts for comment.

Some Missions limit joint evaluation to the design stage; others, as described below, extend such participation to Mission review sessions; still others commence joint evaluation with such sessions. Missions which feel issues raised at the Director's review are too sensitive to involve cooperating country personnel may choose, as mentioned, to hold a separate review session with them. Or, this may be a reason for holding joint participation to the design and progress measurement stage.

One Mission which invites cooperating country persons to sit in the Director's review sessions, sometimes invites them on a personal basis, and other times issues an invitation to a Minister to send an official representative.

A more comprehensive joint review has occurred annually in Uganda for five years, even bridging a change in governments. Leading Ugandan and USAID officials go on a retreat for several days, away from interruptions. The Deputy Minister of Planning presides. The Uganda project directors report on actions concerning recommendations from the previous review, on progress achieved during the year, and on problems outstanding. The respective USAID advisers comment. Officials of both governments question and offer comments. The conclusion is a joint communique listing actions for each party.

Another approach to joint reviews is to work through review sessions sponsored by the cooperating country government. For years, some Planning Ministries have taken the initiative in holding semi-annual meetings to review the status of projects. Often these sessions, however, have not been structured nor have they looked systematically at facts; rather, they have simply been a forum for asking whether there were any problems. Sometimes their usefulness has been limited by the absence of knowledgeable, low-level personnel. To take the approach of working through cooperating country reviews provides an approach to improving the government's own capability for evaluation.

Yet another kind of joint participation has occurred in connection with evaluation studies. Here the evaluation has been planned and conducted jointly, with the evaluation task force comprised of persons from both the cooperating country government and from the USAID.

Finally, some types of activities have continuous evaluation built-in as a part of the activity. Data are regularly collected and analyzed. Such evaluations are usually conducted for programs of mass participation such as those to provide family planning services, to deliver seeds and fertilizers for agricultural production campaigns, or to eradicate malaria. They are also used for educational experiments in which achievement tests are administered to groups of students, etc. Such mass evaluative efforts cannot be conducted without much responsibility being shouldered by the cooperating country, particularly insofar as data collection and tabulation is concerned.

Pros and Cons of Joint Participation in Evaluation --

The possible advantages of some form of joint evaluation are (1) more complete development of a factual base, including cooperating country attitudes, so that the evaluation findings and recommendations are more realistic, and (2) more effective communication. Joint participation in evaluation can educate top officials and arouse their interest. And, when Americans are observed looking at their own shortfalls, cooperating country people will find it easier to do likewise without losing face.

On the other hand, joint participation in evaluation may be cumbersome; time is required to plan the scope of work or the review agenda; overworked officials, both from the USAID and the cooperating government are subjected to another burden on their time; language differences may complicate sessions. Also, Missions sometimes feel that to surface minor issues in another forum would complicate major negotiations. Cooperating countries may have internal jurisdictional problems which make it difficult to establish which is the responsible operating Ministry or Office. Such potential drawbacks to participation may be obviated by the form of participation selected and by careful planning.

Another way of looking at joint evaluations is that they are themselves a form of technical assistance. When less-developed countries reach the point of self-analysis of their own operations, they will have passed an important milestone on the road toward ability to plan and manage their own development.

APPENDICES

GLOSSARY OF TERMS

ASSUMPTION

A situation or a condition which must be assumed to exist if the project is to succeed, but over which AID/W or the Mission has little or no control.

DEVELOPMENT HYPOTHESES

"If outputs, then purpose" is called the project development hypothesis. The hypothesis that purpose will lead to goal is called the program hypothesis. These are hypotheses because we are not certain of the causative relationship between the if statement and the then statement.

END-OF-PROJECT STATUS (EOPS)

The objectively verifiable targets that signal the successful completion of the project purpose. Also referred to as "Conditions expected at end of the project."

EVALUATION

Analysis and comparison of actual progress vs. prior plans, oriented toward improving plans for future implementation. It is part of a continuing management process consisting of planning, implementation, and evaluation; ideally with each following the other in a continuous cycle until successful completion of the activity.

EVALUATION OFFICER

The person responsible for managing the evaluation process.

EVALUATION REVIEW

The process whereby evidence from a project evaluation is reviewed to confirm actions requested and proposed for the coming year:

GOAL

The term designating the programming level beyond the project purpose. It provides the reason for the project and articulates the end toward which the efforts of A.I.D. (and the cooperating government) are directed.

HYPOTHESIS

A statement in the form "if A, then B" where there is uncertainty about the causative relationship between achieving A and achieving B.

INPUTS

Inputs are the goods and services (personnel, commodities, participant training, etc.) provided by the Mission, AID/W, other donors, and/or the cooperating country, with the expectation of producing specific outputs.

LOGICAL FRAMEWORK

A summary of project design, emphasizing the results expected when a project is successfully completed. Results are expressed as objectively verifiable indicators.

MEASURES OF GOAL ACHIEVEMENT

The means of verifying through indicators the achievement (in either quantitative or qualitative terms) of the goals.

OBJECTIVELY VERIFIABLE INDICATORS

Good project design must include prior definition of what will be measured to demonstrate progress (indicators) and how much (targets). Ways of verifying progress should be objectively stated so that both a proponent of a project and an informed skeptic would agree that progress has or has not been as planned. Preestablishing objective verifiable indicators and targets helps focus discussion on evidence rather than opinions.

OUTPUTS

The specifically intended kind of results (as opposed to their magnitude) that can be expected from good management of the inputs provided. A project manager might be considered responsible for producing specific outputs; the Mission or AID/W action office shares responsibility for the judgment that producing these outputs will result in achieving purpose.

PROJECT

A planned undertaking that clearly specifies what will be accomplished, over what period of time, and at what cost.

PROJECT APPRAISAL REPORT (PAR)

The by-product of the project evaluation process that reports the results of evaluations.

PROJECT DESIGN

A summary of what the project is expected to achieve (purpose), and how it will be achieved with the inputs and time available. The key elements of project design may be summarized in the logical framework format.

PROJECT MANAGER

The individual responsible for a project. More specifically, the individual who is charged with protecting A.I.D.'s manageable interests, producing the agreed-upon outputs within the specified time and cost constraints.

PURPOSE

That which is expected to be achieved if the project is completed successfully and on time. It expresses in quantitative or qualitative terms (within parameters capable of verification) that which we hope to create, accomplish, or change with a view toward influencing the solution of a country or sector problem.

TARGET

An indicator with a magnitude to be realized at a specific date; an explicit and objectively verifiable measure of results expected.

APPENDIX B

ILLUSTRATIVE BASELINE MEASURESI. Housing Quality

(This* has been used as a rating scale by a housing officer to get a quantified measure of housing quality in different cities or different sections of the same city.)

	SCORE		
	Yes	No	
1. Inadequate original construction or conversion: dirt floors	1	3	
2. Considerable wear on inside steps or floors	2	3	
3. Are the rooms in good order?	3	2	
4. Is the furniture in good repair?	3	2	
5. Substantial sagging or bulging of outside walls or roof	1	3	
6. Shaky or unsafe porch, steps or railing	2	3	
7. Broken or missing window panes	2	3	
8. Rotted or loose window frames	2	3	
9. Deep wear on doorstll, door frames or outside steps	2	3	
10. Badly rusted or partially missing gutters and downspouts	2	3	
11. Is the lot clear and in good order?	3	2	
12. Inadequate original construction or conversion: makeshift interior walls	1	3	
13. Inadequate original construction or conversion: makeshift exterior walls or roof	1	3	
	over large area	over small area	none
14. Holes, open cracks, rotted, loose, or missing materials on inside walls	1	2	3
15. Holes, open cracks, rotted, loose, or missing materials on floors	1	2	3
16. Holes, open cracks, rotted, loose, or missing materials on ceilings	1	2	3
17. Substantial sagging of floors or walls	1	2	3
18. Holes, open cracks, rotted, loose or missing materials on foundation	1	2	3
19. Holes, open cracks, rotted, loose or missing materials on outside walls	1	2	3
20. Holes, open cracks, rotted, loose or missing materials on roof	1	2	3
21. Where is water obtained?			
Other			(Score 1)
Pipes or wells outside			(Score 2)
Piped into house			(Score 3)

* Adapted from Cornell University Index of Housing Quality (Contract AID/csd-817).

22. What type of lighting does unit have?
 Other (Score 1)
 Electric (Score 3)
23. What kind of fuel is used for cooking?
 Other (Score 1)
 Electric or gas (Score 3)
24. What kind of refrigeration is used?
 Other or none (Score 1)
 Electric (Score 3)
25. What toilet facilities are available for this household?
 Other (Score 1)
 Flush toilet inside (shared) or outside (Score 2)
 Flush toilet inside, exclusive use (Score 3)
26. What kind of bathing facilities are available for household?
 Other (Score 1)
 Installed tub or shower inside (shared) or outside (exclusive use) (Score 2)
 Installed tub or shower inside, exclusive use (Score 3)

TOTAL score possible = 3 x 26 = 78

II. Measuring Community Development*

This is a draft of an instrument for comparing the level of development of communities and urban barrios. Its purpose is to provide a systematic way of selecting communities which are most ready to take advantage of development programs or outside help such as Peace Corps Volunteers. It is designed to be completed by one person in about half a day in small communities or, at most, one full day in large communities or barrios in cities. It is not an instrument for thorough, in-depth study of the community. Rather, it represents the first step in choosing high potential communities for development. The baseline measures will be obtained:

(1) by walking up and down each street of the community, counting and classifying houses, and counting stores, public buildings, restaurants, theaters, etc.

(2) by talking to four or five knowledgeable community members, such as the local priest, teniente politico, school teachers, coop leaders, and others to find out such factors as existing active organizations, outside entities represented in the community, community projects, social problems or health problems.

* For illustrative purposes only. By courtesy of Richard J. Greene, USAID/Ecuador.

These baseline measures of community achievements and activity should reflect the will and energy of community leaders and members. In other words, communities that are well organized and have many improvements and services are likely to have more dynamic populations than do less developed communities. These active communities are the ones which, hypothetically, should benefit most from development resources, whether Volunteers, technical assistance, organization efforts for coops, education programs, and the like.

COMMUNITY SURVEY

I. IDENTIFICATION

- A. Name of community _____
- B. Location (approximate time by car and direction from major town or landmark) _____
- C. Is community capital of canton _____ or parish? _____
- D. Region: Coast _____ Sierra _____ Oriente _____
- E. Date of founding _____
 Predominant first language: Spanish _____
 Quechua _____ Use both Quechua and Spanish _____

II. House types and population estimates (tabulate number in each category) TOTALS

- A. Chozas (houses markedly poor; shacks compared to rest) _____
- B. Paja, palm, wood roof _____
- C. Zinc, ardex, cement roof _____
- D. Tile (clay or cement) roof _____
- E. Cement roof _____

Total houses in community - _____

- F. Houses under construction (foundation begun or more) _____
- G. Give estimate of number of people per house _____
- H. Estimate of total population _____

_____ =
 (Total houses) x (People)

III. COMMUNITY SERVICES (Indicate type or number in each category) A. Water System (check which are used)

- Wells _____
- Community Faucets _____
- Water in Houses _____
- No improved water system - river, irrigation ditches lake, etc. _____

B. Community Electric System
 Present _____ No. customers _____ (ask company or coop) None _____

C. Communications (check every mode that is in community) Telephone _____ Telegraph _____ Radio transmitter _____ Newspapers delivered daily _____ Number per day (ask agent) _____

D. Street System - No streets, only trails _____ Only one street _____ Number blocks dirt streets _____; gravel _____; cobblestone _____ paved _____

E. Transportation System - Number roads to community _____ Number hours by foot to road _____ On main road _____ Distance (time) by car to main road _____ Taxi service in community _____ Number buses per week _____ Train service _____ Plane service _____

F. Public Services (indicate number)

Plaza _____	Military Buildings _____
Chapels _____	Municipal Government Bldgs. _____
Catholic Churches _____	Agency offices _____
Protestant " _____	Community Center Bldgs. _____
Post Office _____	Primary schools _____
Police Station _____	Colegios _____
Fire Department _____	Parques Infantiles _____
Municipal Bathrooms _____	Canchas _____
Open Markets _____	Health Posts _____
Covered Market Buildings _____	Hospitals _____

G. Private Services (indicate number)

Banks _____	Hotels or Pensiones _____
Restaurants _____	Drugstores _____
Movie Theaters _____	Barbershops _____
Billiard Halls _____	Shoe Repair _____
Gasoline Station _____	Tailor/Seamstress _____
Mechanic Shop _____	Carpenter Shop _____
Print Shop _____	Other (specify) _____

IV. COMMUNITY SPECIALISTS (indicate number)

Priests (full time) _____	Doctor _____
Teniente Politico _____	Nurse _____
Jefe de Registro Civil _____	Dentist _____
Police _____	Teachers _____

V. COMMUNITY ORGANIZATIONS (padres de familia, Recreation, social, religious, cooperatives, political, agricultural)

Name	Type (Purpose)	Frequency of Meeting (Formal or Informal)	Number of Socios
1. _____	_____	_____	_____
2. _____	_____	_____	_____
3. _____	_____	_____	_____
4. _____	_____	_____	_____
5. _____	_____	_____	_____

VI. COMMUNITY PROJECTS (Physical improvements planned or in process)

A. Project description _____

B. Community Organization Sponsor _____

C. Work stage: Only Planning Underway (explain progress, e.g., start of organizing, talk to agency, etc.) _____

When actual work started _____ Date scheduled completion _____

D. Agency Participation
 No agency help _____ Community initiated, agency help with execution _____
 Agency initiated, community execution _____
 Agency initiated and execution _____
 Agency(s) which are participating _____

VII. COMMUNITY ECONOMICS

A. Land tenure of surrounding community
 Mainly commercial haciendas _____
 Mainly small property owners _____ Estimated plot size _____
 Mainly haciendas which are subdivided: arrendatarios, desmonteros, arimados, partidarios (circle which is the dominant arrangement); estimated plot size _____

B. Production (List major crops or products shipped for sale outside of community)
 1. _____ 4. _____
 2. _____ 5. _____
 3. _____ 6. _____

C. If city barrios, list major occupations of inhabitants

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
| 2. _____ | 5. _____ |
| 3. _____ | 6. _____ |

D. Industries (list all types, include artisan industries)

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
| 2. _____ | 5. _____ |
| 3. _____ | 6. _____ |

VIII. COMMENTS (Explain if any of the following are present)

- A. Fundamental social or economic change movements (e.g., plans for land acquisition, obtaining water rights, etc.)
- B. Community Problems (e.g. serious health problems, delinquency, alcoholism)
- C. Special economic circumstances (e.g., artisan economy, presence of important industry, etc.)

SELECTED OUTPUT INDICATORS
(For illustrative purposes only)

MARKETING AND DISTRIBUTION

- *Number firms participating in sales training program
- *Number national sales training seminars held
- *Number product-use pamphlets produced
- *Number training films produced
- Number warehouses erected
- *Number trainers trained
- *Number training meetings conducted (in sales techniques, technical use of product, and management procedures)
- Number trained farm organization supervisors on duty
- *Number education meetings (for fertilizers, pesticide)
- Number of farm organizations

CREDIT

- Increase in field staff
- Number rural banks established
- Number bank branch offices opened
- Number of import and distribution loans
- Value of import and distribution loans
- Number of loan applications received
- Number of loan applications processed
- Number of loan applications approved
- Proportion of cultivators receiving loans (number recipients of loans divided by number of cultivators)

CROP PRODUCTION

- *Hectares improved variety planted
- Seed standards developed
- Seed growers' association established
- *Number farmers trained in new techniques
- *Tons seed grain imported
- Tons seed grain produced locally
- *Seed storage facilities constructed and equipped
- Private sector seed importation system developed (number of importers)
- Number tons of yield harvested (milled)

ANIMAL PRODUCTION

- Number breeder hatcheries (broiler and egg producers) established

* These are input measures showing progress in a course of action towards a target but are not the target outputs themselves.

- Number day old chickens produced per year
- Number market eggs produced per year
- Number swine farms established (or improved)
- Increase in brood sows
- Increase in market hogs
- Number vaccine production and testing centers established
- Number quarantine stations existing
- Number animal disease diagnostic centers established
- Amount vaccine produced
- Number hogs (chickens, dogs, etc.) vaccinated
- Number feed mills established
- Amount produced per year of balanced formulated feeds
- Number abattoirs established
- National livestock center established
- Number pigs for sale

LAND REFORM

- Number hectares aerial photographed (or surveyed)
- Number of titles registered or distributed
- Necessary legislation passed
- Percent farmers on own land

MANPOWER DEVELOPMENT

- Number occupational employment surveys completed
- Number on-the-job training systems in operation

TAX COLLECTION

- Increase in revenue over last year

FAMILY PLANNING

- Number of home visits by F.P. personnel
- Number of pills distributed
- Number of training courses given
- Number of trainees graduated
- Number of research projects completed
- Number of new acceptors

COMMUNICATIONS

- Newspaper circulation per 1000
- Number pieces mail per 1000
- Radio - TV per 1000
- Cinema attendance per 1000
- Total number telephones in country
- Number telephones in major cities
- Number telephones outside major cities

INSTITUTIONAL MATURITY

Political viability demonstrated
 Professional status recognized
 Technical competence proved
 Survival capacity demonstrated
 Ability to attract financial resources shown
 Capacity to innovate demonstrated
 Services being used in community

LABOR

Number collective bargaining contracts
 Number members in unions divided by number of wage earners
 Changes in real wages and benefits

EDUCATION

Number classrooms built
 Number graduates of teacher training colleges
 Number prototype libraries established
 Number returned participants assigned to appropriate positions
 Percent literate adults in population
 Percent children able to pass UN reading test
 School enrollees; ratio to school-age population
 Number of drop outs; % drop outs by grade and age
 Access to education - number of members of minority group
 - girls, numbers and percent of total
 Student-teacher ratios
 Number of teachers in position
 Literacy rates - changes for total population and percent over 15 years old
 Number textbooks written, printed, revised, distributed
 Percent vocational education graduates placed
 Earnings of vocational education graduates vs. untrained
 Budget support from local or central government

SUGGESTED "SOCIAL INDICATORS"I. GeneralA. Population Distribution

Knowledge of the spatial distribution of population is useful for many types of social, political and economic analysis. The reason for requesting a division of the population into rural vs. various size urban categories instead of the more conventional urban-rural classification is to obtain some picture of the relative significance of urban communities of different size with different socio-economic functions: 1) market-towns(5,000 - 20,000) which can serve as centers of agro-industrial activity, 2) medium sized cities (more than 20,000) which serve as regional centers and can absorb much of the rural-urban migration, and 3) vast urban agglomerations to which villagers flock after leaving intermediate cities in which their integration is probably difficult.

	<u>1960</u>	<u>1965</u>	<u>1970</u>	<u>1975</u>	<u>1980</u>
a) Rural Population					
b) Towns of 5,000 - 20,000					
c) Intermediate Cities					
d) Major cities					

B. Access to Education - Primary School Scholarization Rate

School attendance in relation to school-age population indicates how much of the population has access to education. Differential urban and rural rates are especially significant since the rural population generally has inferior access to education and similar services. Because education is so important a factor in social mobility, school attendance ratios (scholarization rates) may also serve as an indicator of social mobility.

If school enrollment and population data are broken down by urban and rural, as it is for some countries, differential urban and rural scholarization rates can be calculated. In the absence of such data it may be possible to make an estimate based on general knowledge of the availability of primary schools in rural areas.

Primary School Scholarization
Number of grades
Age at entrance to first grade

		1960	1965	1970	1975	1980
<u>NATIONAL</u>	1 Enrollment					
	2 School-Age Population (Age to)					
	3 Scholarization Rate (1÷2)					
<u>URBAN</u>	1 Enrollment					
	2 School-Age Population					
	3 Scholarization Rate (1÷2)					
<u>RURAL</u>	1 Enrollment					
	2 School-Age Population					
	3 Scholarization Rate (1÷2)					

C. Distribution of Service Activities: Telephones

The number of telephones in the major cities should be stated along with the total number in the country. The number of actual instruments is preferable to the number of telephone numbers listed in directories since it gives a better indication of telephone use, but if the former is not available the latter can be used. These data are presumably available at the telephone bureau (PTT) or company. The number of telephones per 100,000 of population is useful as a measure of the development of communications, but the purpose of this indicator is as a measure of the extent to which service activities (businesses, government offices, commercial agriculture, etc.) are geographically dispersed throughout the country or narrowly concentrated in one or two centers. The distribution of telephones is thus a proxy for the distribution of economic activity other than traditional agriculture and handicrafts.

	1960	1965	1970	1975	1980
1. Number of Telephones (Total)					
2. Number in Major City (Cities)					
3. Number outside Major City (1-2)					
4. Percentage Outside Major City (3÷1)					

D. Communications: Newspaper Circulation

The circulation of newspapers expressed as the daily sales of newspapers per 1,000 of population gives an indication of what proportions of the population is participating in the national economic, social, political and cultural life. All newspapers, including local weeklies, can be included but it is presumed that the total circulation is preponderantly accounted for by metropolitan dailies and that this figure is relatively easy to get.

	1960	1965	1970	1975	1980
1 Daily Newspaper Circulation					
2 Population (1,000)					
3 Circulation per 1,000 people (1÷2)					

II. Agricultural

The following are combinations of economic and social data and various indicators useable for evaluations in the agricultural field. National accounts information is assumed to be already available, both in the countries and in AID/W.

A. Distribution of Land Ownership

The pattern of land ownership is closely tied to social structure and the distribution of power as well as to production. It is therefore important to know the existing situation and to have some understanding of the way it is evolving, i.e., toward greater concentration or greater equality. The pattern of land holdings may be described by size and by type of holding. Missions should use some recent year for which information is available. Repeating these data for five year intervals will show trends. The entries under column (1) "Hectares," may need to be revised depending on how the country groups farms by size. (One hectare = 2.47 acres.)

Land Holdings Pattern, 19__

Hectares	Land in Farms (000 hectares)	Number of Farms (000)	Average Size of Farms (2÷3)
(1)	(2)	(3)	(4)
0 - 2.4			
2.5 - 4.9			
5.0 - 9.9			
10.0 - 19.9			
20.0 - 49.9			
50.0 - 99.9			
100.0 & over			

Farmer - Land Relationship, 19__

Hectares	Owner	Tenant	Share-cropper	Landless Laborer	Other	Total
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0.0 - 2.4						
2.5 - 4.9						
5.0 - 9.9						
10.0 - 19.9						
20.0 - 49.9						
50.0 - 99.9						
100.0 & over						

B. Access to Modern Farm Technology

The extent to which farmers are participating in the use of improved inputs is an important determinant of the rate at which the agricultural sector is able to modernize. Use of chemical fertilizers, on which data are relatively good, may be taken as a proxy for the whole range of improved inputs and practices. For this purpose the most useful indicator of fertilizer consumption is the proportion of cultivators (excluding farm laborers) using chemical fertilizers. If this is not available, annual consumption of chemical fertilizers (expressed as kilograms of plant nutrient, not bulk fertilizer) per hectare of cultivated land would be an acceptable alternative.

	1960	1965	1970	1975	1980
1. Number of Cultivators (excluding farm laborers)					
2. Cultivators using chemical fertilizers					
3. Proportions using fertilizers (2 ÷ 1)					

or:

	1960	1965	1970	1975	1980
1. Annual Consumption of Chemical Fertilizers (M.T. of nutrient value)					
2. Cultivated area (1,000 hectares)					
3. Use of fertilizer per hectare (kg) (1 ÷ 2)					

C. Access to Agricultural Credit

Access to credit on reasonable terms is a major factor affecting the adoption by farmers of improved practices and purchased inputs. It is therefore important to know what

proportion of the agricultural population (cultivators, not farm laborers) has access to such credit.

Distribution of Credit by Farm Size, 19__

Hectares	Number of Loans	Total Value of Credit	Average Value of Loans 3:2
(1)	(2)	(3)	(4)
0 - 2.4			
2.5 - 4.9			
5.0 - 9.9			
10.0 - 19.9			
20.0 - 49.9			
50.0 - 99.9			
100. & over			

Distribution of Loans by Source, 19__

Total, All Sources	Number of Loans	Total Value of Credit	Average Value of Loans 3:2
(1)	(2)	(3)	(4)
Government Agr. Bank			
Private Banks			
Farmers Cooperatives (incl. Credit Unions)			

Separate tables on this sort of information may be gathered for short, medium and long-term loans - the latter being those lasting more than twelve months.

D. Access of Farm Population to Markets

Farm-to-market roads make it possible for farmers to produce for an off-farm market and thus constitute a major determinant of whether they adopt improved practices. The possibility open to farmers of participating in the market can be gauged by the extent of the feeder or farm-to-market road system. Kilometers of farm-to-market roads usable throughout the year by motor vehicles (and kilometers of canals, if relevant) per square kilometer of cultivated land give a good measure of the extent of the transport system. The national highway system should be excluded, but if it is impossible to separate it out, use total road mileage.

	1960	1965	1970	1975	1980
1. Kilometers of feeder roads					
2. Area cultivated (1,000 ha.)					
3. Roads/cultivated area (km/ha) (1÷2)					

E. Monetization of Agriculture

The relative sizes of the subsistence (or non-monetized) and the commercial (or monetized) sectors are an important indication of the extent to which farmers are participating in the national economic system and in the national life generally. This can be measured in terms of the share of total agricultural output produced in the subsistence sector or in terms of the proportion of cultivators working in the subsistence sector. (The two ratios will differ since productivity in the subsistence sector is lower than in the commercial one.)

1960 1965 1970 1975 1980

1. Gross value of agricultural output
2. Gross value of subsistence output
3. Share of subsistence sector (2÷1)
4. Number of cultivators
5. Number of subsistence cultivators
6. Share of subsistence cultivators (5÷4)

III. Employment and Wages

A. Structure of Employment: Wage and Salary Earners

The size of the wage and salary earning component in the total economically active population reflects rationalization and institutionalization of economic activity. It can be used as an indicator of modernization. This group consists of those paid regularly by the week, month or year, such as the employees of government agencies, public or private business enterprises, commercial agriculture, and organizations dispensing professional and personal services. It does not include the self-employed (e.g., in agriculture, handicrafts, small shops or street-vending) or casual labor employed for short periods (e.g., migratory agricultural workers).

1960 1965 1970 1975 1980

1. Economically Active Population
2. Wage and Salary Earners
3. Ratio (2÷1)

B. Unemployment

Unemployment is a structural problem of modernization that may have economic, social, and political consequences if it rises steadily or is not alleviated over long periods of time. The number of unemployed is, of course, more meaningful if related to the total labor force as provided for in the table below. Since urban unemployment presents special

CONTINUED

1 OF 2

problems, provision is made in the table for presenting it separately in relation to the urban labor force.

	<u>1960</u>	<u>1965</u>	<u>1970</u>	<u>1975</u>	<u>1980</u>
1. Unemployed					
(a) Urban unemployed					
2. Labor Force					
(a) Urban Labor Force					
3. Unemployed as proportion of Labor Force (1:2)					
(a) Urban unemployed as proportion of urban labor force (1a:2a)					

C. Trend in Real Wages

The purpose of this measure is to ascertain whether the economic position of wage earners has improved or deteriorated, and how much. The average daily wage (for that portion of the labor force on which wage statistics are available) should be deflated by the index of the cost of living (or other appropriate deflator).

	<u>1960</u>	<u>1965</u>	<u>1970</u>	<u>1975</u>	<u>1980</u>
1. Money Wages					
2. Cost of living index (1960=100)					
3. Real Wages $100 \times (1:2)$					

D. Unionization

The extent of unionization, as measured by the percentage of the wage earning population which belongs to a union, when taken with the activeness of the trade union movement, as measured by the number of workers engaged in strikes during a 12-month period, gives an indication of the degree of organized expression available to the wage-earning population. The data are more relevant when compared with real wage trends in III.C. above.

The membership data are presumably available from the trade unions. The wage earning population used as the denominator should (like the numerator) exclude agricultural workers and civil servants, but include employees of state enterprises.

The data on strike participation are simply an estimate of the number of workers who participated in strikes, not of man days (or years).

1960 1965 1970 1975 1980

1. Number of Wage Earners
2. Union Membership
3. Union members as % of
Wage Earners (2:1)
4. Number of Workers
Participating in
Strikes

INDICATORS - ALLIANCE FOR PROGRESS

PER CAPITA GROWTH

Goal - 2.5% growth per capita per year.

Indicators - GNP, total and per capita
GNP, Growth rates total and per
capita
GNP, indexes total and per capita.

Advantages of Indicators - Combines effect of production and
population growth
Best single overall measure.

Shortcomings of Indicators - Intercountry comparisons need
adjustment for constant dollar
exchange rates
Masks or omits other significant
variables such as income
distribution or rural-urban
disparities.

INCOME DISTRIBUTION

Goal - More equitable distribution to
economic and social groups, with
larger shares of benefits of
progress going to needier
sectors and investment

Indicators - Index of investment
Income distribution
Average earnings by sector (where
available)
Social progress - life expectancy
- access to education
- agricultural productivity

Advantages of Indicators - Income distribution is best
available quantitative indicator
of general welfare
Relate to some of necessary policy
measures for social progress

Shortcomings of Indicators - Standards of living affected by
prices and social services, so

that inter-country comparisons less meaningful than intra-comparisons over time.

TRADE DIVERSIFICATION

Goals - Make national income structures increasingly free from dependence on export of a few primary products and on import of capital goods
Stabilize export prices or income

Indicators - Composition of exports
Trends of GNP sectors
Indexes - production manufactured exports

Advantages of Indicators - Like the income distribution, supplement GNP as an indicator of general development

Shortcomings of Indicators - Do not relate to price stability.

INDUSTRIALIZATION

Goal - Accelerate rational industrialization to utilize natural resources and provide employment, taking full advantage of both public and private sectors

Indicators - Value added by manufacturing
Power production
Output of specific manufactures
Export of manufactures

Advantages of Indicators - Value added measures actual contribution of processing, while output figures may be better for inter-country comparisons by eliminating comparative price problems
Export of manufactures gives a clue to their competitiveness
Power consumption is recognized as a good general indicator of industrial sophistication

Shortcomings of Indicators - Should be used in conjunction with other indicators for agriculture and education, since LDC's have often been tempted to over-emphasize investment in the visible aspects of modernity at the expense of general development.

AGRICULTURE

Goals - Raise the level of agricultural output and productivity greatly
Improve related storage, transportation, and marketing services

Indicators - Central government agriculture expenditure -
 . index
 . % of GNP
 . % of total government expenditure
 Total agriculture production -
 . aggregate value
 . index
 . per capita index
 Total crop production -
 . aggregate value
 . index
 Total food production -
 . aggregate value
 . index
 . per capita index
 Agricultural schools - enrollment and graduates
 Agricultural coops - numbers and members

Advantages of Indicators - Production was considered best general comparable indicator because it tends to average out variations in individual crops, soils, weather, etc.
Per capita indexes relate production growth to population growth
Expenditures show level of government interest

Shortcomings of Indicators - Production does not necessarily indicate progress in technology as do F.A.O. reports on yields per acre for many crops (although these figures must be compared over an extended time series to average out weather variations).
Production and needs do not always relate directly, since countries can or should import and export widely different proportions of their consumption and output.

AGRARIAN REFORM

Goal - Comprehensive reform leading to effective transformation of unjust systems of land tenure and use so that, with timely and adequate credit, technical assistance and facilities for marketing and distribution, land becomes a basis of economic stability, welfare and dignity of man who works it.

Indicators - No uniform indicators possible.

Shortcomings of Possible Indicators - Uniform figures not available. Reform consists of more than tenure. Credit and other supporting measures.

EDUCATION

Goals - Eliminate adult illiteracy.
Assure access to 6 years of primary education for each school age child by 1970.
Modernize and expand vocational, technical, secondary and higher educational and training facilities.
Strengthen capacity for basic and applied research.
Provide the competent personnel required in rapidly growing societies.

Indicators - Central government education expenditures -
. index
. % of GNP
. % of total government expenditures
Primary schools -
. enrollment
. student-teacher ratios
. teachers
. graduates
. classrooms constructed
Secondary schools -
. student-teacher ratios
. teachers
. graduates
General secondary and higher schools - enrollment
Teacher training institutions - enrollment
Teacher training institutions - teachers
Teacher training institutions - graduates
Higher schools - graduates
Illiteracy

Advantages of Indicators - Generally relate directly to targets

Shortcomings of Indicators - Do not report on qualitative goals such as "modernize," "strengthen research capacity."

HEALTH

Goals - Increase life expectancy at birth by a minimum of 5 years and
Increase ability to learn and produce by:
. Providing public water and sewage disposal to 70% of urban and 50% of rural population
. Reducing mortality of children less than 5 years of age by one-half
. Controlling more serious communicable diseases
. Improving nutrition
. Improve basic health services

- . Train medical and health personnel
- . Intensify health research

Indicators - Practicing physicians
 Practicing nurses
 Hospital beds
 Life expectancy
 Potable water availability
 % of population provided with sewage facilities
 Death rates for major epidemic diseases
 Food calorie availabilities

Comment - General goal of increased ability to learn and produce was generally translated into countable actions.

GOVERNMENT REVENUES

Goals - Improve ability to collect revenues needed to support other goals
 Improve equity of tax systems
 Improve effectiveness of tax systems in promoting development

Indicators - Domestic revenues - index
 Domestic revenues - % of GNP
 Tax revenues index
 Central government tax revenues - % of GNP
 Central government tax revenues - % of domestic revenues

Advantages of Indicators - Total revenue as a % of GNP is probably the best single indicator of country self-help, although some non-tax revenue may reflect entrepreneurial activities of governments

Shortcomings of Indicators - Data on regional and local revenues likely to be incomplete
 Central government revenues may not be useful for inter-country comparisons because of variations in reliance on local governments.

PERFORMANCE STANDARDS

APPENDIX F

These are Edgar L. Owens' "working" standards of progress. There is nothing "official" about them. But they are among the few rule-of-thumb standards that are available and useable to make comparisons. They are summarized here in the interest of generating further discussion and research on them.

A. General Economic Indicators

1. Per Capita Income

A good rate indicates rapid progress in both industry and agriculture. A poor rate suggests some major problems which, historically, we know are probably found in agriculture and agro-industries, since rapid industrial progress usually follows farm progress. For a good rate, a norm seems to be 5% or more, while a poor rate is something substantially less than 5%.

Per Capita Domestic Product Percent Annual Growth, 1960-69

Japan	10.0
Korea	6.4
Taiwan	6.3
Puerto Rico	6.0
Israel	5.3
Thailand	4.7
Ivory Coast	4.7
Yugoslavia	4.6
Malaysia	3.8
Mexico, Turkey and Morocco	3.4
Argentina	2.6
Venezuela	2.5
Tunisia	2.1
Philippines	1.9
Chile & Uganda	1.7
Tanzania	1.6
Colombia & Kenya	1.5
Brazil & Peru	1.4
India	1.1

cont'd

Senegal	0.1
Ghana	0.0
Nigeria	-0.3
Uruguay	-0.8

SOURCE: World Bank

2. Exports

Increases of \$2 to \$5 (current prices) per capita per year have been recorded. It ought to be possible to increase exports at a rate of \$1.50 per capita annually at a minimum. Very low rates, such as 20¢ or 30¢ indicate major problems.

Equally important, the proportion of exports that are processed in some fashion should rise by several percent a year.

The first table shows exports per capita for a number of countries, 1950 and 1969. The variation in performance is very considerable and is essentially a reflection of a country's capacity to diversify its production base and to meet international standards in quality, delivery dates, spare parts, and so forth.

The second table, the comparison of Taiwan and Mexico, is an example of how export data can be analyzed to get some notion of how well a country is developing its capacity to pay its own way in the international community. The capacity to compete is essentially a processing and manufacturing capacity. As the table shows, Taiwan has been developing this capacity much more rapidly than Mexico. And, as shown on the first table, Taiwan's exports per capita have multiplied six times faster than Mexico's.

Two qualifications should be added to the above. First, the oil-mineral rich countries, such as Venezuela, Iran, and Malaysia are obviously in a special category. The question for these countries is how they use their ample export earnings.

Second, the entries in the left-hand column of the second table can be made more or less detailed than shown here, and they should be adjusted somewhat to suit the composition of exports of a country. The table is included here simply to illustrate how export statistics can be used as an analytical tool.

Exports Per Capita: Early 1950s-1969

Country	Early 1950s	1969	Change
Israel	\$27.93	\$242.27	\$214.34
Taiwan	10.66	76.05	65.39
Yugoslavia	10.91	72.46	61.55
Korea	.71	19.83	19.12
Mexico	17.84	29.23	11.99
Morocco	21.19	32.23	11.04
Egypt	17.73	22.92	5.19
India	3.39	3.51	.12
Argentina	69.01	67.21	-1.80
Brazil	26.90	25.04	-1.86
Indonesia	10.42	6.80	-3.62
Colombia	39.89	29.69	-10.20

SOURCE: UN Yearbook of International Trade Statistics.

Percentage Distribution of Exports, Taiwan and Mexico, 1951 and 1968

	Taiwan		Mexico	
	1951	1968	1951	1968
Traditional Agricultural Exports				
Taiwan: Sugar, Tea, Rice	73.9%	8.9%		
Mexico: Cotton, Coffee, Fish			29.5%	20.0%
Other Unprocessed Agricultural Products	9.7	14.4 ^{a/}	10.6	27.4
Processed Agricultural Products	4.1	19.0 ^{a/}	18.9	13.7
Sub-Total, Agricultural	87.7	42.3	69.0	61.1
Mineral Ores and Oil Manufactures	3.2	1.4	37.8	24.0
Sub-Total, Non-Agricultural	4.1	57.5	41.0	34.9
Miscellaneous Exports, Errors and Omissions	8.2	.2	.0	4.0
Total	100%	100%	100%	100%
Total Dollar Value of Exports	\$98.3	\$802.5	\$468.7	\$1,257.6
Proportion of Products Exported as Processed Agricultural Commodities or Manufactures	5.0%	75.1%	22.1%	24.6%

SOURCE: UN Commodity Trade Statistics Bulletins
^{a/} Excludes wood products made from imported logs.

3. Birth Rate

Once a secular decline in the birth rate sets in, as in Taiwan and Puerto Rico, then the rate should decline by around 1/2 per 1,000 per year for 2 or 3 decades until it is down to 20 per 1,000 or lower.

Birth Rates Per 1,000 Population

	1948	1967	Change
Puerto Rico	40.2	25.4	-14.8
Taiwan	39.7	28.1	-11.6
Israel	28.6	26.8	-1.8
Mexico	44.6	41.3	-3.3

Late 1960s

Indonesia	48.3	Brazil	37.8
Philippines	44.7	Colombia	44.6
Iran	45.4	Peru	41.5
Morocco	49.5	Turkey	39.6
Tunisia	46.3	India	42.8
Thailand	42.8	Egypt	44.1

SOURCE: UN Demographic Yearbook

B. Agriculture1. Agricultural Productivity

Yields per acre of the basic food grains of a country are a general indicator of the extent to which small farmers are going modern since the only countries with high yields and a high rate of increase are those in which small farmers have been brought into a modern agricultural system. As one person has expressed the point, "Food shortages are not due to a lack of technology, but to the inability to apply existing technology." The following table shows the enormous variation in capacities to apply technology.

Foodgrain Yields: 1948-50 to 1968-70 (pounds per acre)

	1948-50	1968-70	Increase
Taiwan	1800	3510	1710
Egypt	2120	3370	1250

Korea	1640	2850	1210
Yugoslavia	1145	2185	1040
Ceylon	1265	2060	795
Mexico	700	1265	565
Colombia	915	1480	565
Chile	1125	1630	505
Thailand	1190	1670	480
India	640	945	305
Turkey	835	1105	270
Peru	1225	1495	270
Philippines	930	1145	215
Brazil	1170	1225	55
Iran	900	950	50
Tunisia	440	395	-45
Japan	2920	4285	1665
USA	1495	2895	1400
Denmark	2670	3860	1190
Gt. Britain	2155	3170	1015

2. Fertilizer Consumption

When fertilizer usage is virtually nothing to start with, consumption ought to rise very rapidly. How much fertilizer a country ought to use varies very much according to demand, the type of farming system, and physical conditions. However, it is clear from the following table that in many countries fertilizer usage is much less than it should be. The principal reason is the low usage rate among small farmers.

Fertilizer Consumption, 1969/70
(Pounds fertilizer nutrient per acre)

Japan	415
Taiwan	266
Korea	206
Egypt	103
Yugoslavia	76
USA	74
Ceylon,	54
Mexico	21
Philippines and Turkey	16
Thailand	12
India	10
Morocco	7

3. Agricultural Credit

Preliminary research on production credit suggests that the annual requirement is somewhere in the neighborhood of a quarter or more of gross annual agricultural product. The proportion of farmers receiving institutional credit should be 60-80%, which can be taken to mean that such credit is available to all farmers. There are always some who do not need it or use it.

4. Extension and Research

More work needs to be done on quantitative measures of qualitative inputs. For example, looking at countries where agricultural extension works and where agricultural research is, first, good, and second, communicated to farmers, might give a clue to desirable ratios. Tentative suggestions are:

- a. One extension worker for every 1,000 agricultural workers
- b. Perhaps almost as many researchers as extension workers
- c. Expenditures for agricultural research should be around 1% of the value of annual agricultural output.

C. Rural Development

1. Rural Capital Formation

Capital formation is a necessary component of an agricultural revolution as well as of other development. Moreover, part of this capital should come from rural areas. Generally speaking, if statistics are available, the deposits in rural banks, cooperatives and other institutions are close to zero because local financial institutions that farmers are willing to use do not exist. In Taiwan, in 1970, such deposits amounted to \$125 per acre, and are the principal source of funds for agricultural production credit. Taiwan has one savings institution for each 2,500 farms. How well these ratios would fit other countries would need to be determined.

2. Farm-to-Market Roads

If, general, geographically dispersed development is to occur, a country must move from an acute shortage of farm-to-market roads (including canals where feasible) to adequacy in some reasonably short period, say one

decade. A possible standard of adequacy may be 2 1/2 to 3 miles of road for each square mile of cultivated land. To reach this ratio in a decade would require construction of about 1/4 mile of road per cultivated square mile per year, if the country starts with 1/2 mile of road per cultivated square mile.

Farm-to-Market Roads - Ratio of Miles to Cultivated Sq. Miles

U.S.A.	3.28	Philippines	1.14
Taiwan	2.67	India	.79
East Pakistan	2.45	West Pakistan	.71
Chile	1.91	Tunisia	.58
Colombia	1.59	Iran	.47

SOURCE: Statesmen's Yearbook and FAO Production Yearbook

Note: The metric equivalent of 2 1/2 - 3 miles of road to one square mile of cultivated area is approximately 1 1/2 - 1 3/4 km. of road to one square km.,

3. Location of Facilities

A good deal can be told about the quality of economic development by statistics on the distribution of various physical facilities between the capital or the largest city and the rest of the country. For example, 3/4 of the telephones in Thailand are in Bangkok. In Taiwan, the proportion in Taipei is much lower. The same kind of unequal distribution is true of post offices, schools, clinics, factories, financial institutions, warehouses, etc. Such simple statistics tell a good deal about the ability of a government to get development underway outside of urban complexes, which, again, tells something about the state of agriculture. Work is needed before standards of performance can be developed.

D. Industry and Power

1. Manufacturing Output

In countries with little industry, an increase of output of 10% or more per year ought to be possible for at least a decade, and possibly several.

Percent Increase in Manufacturing Output 1960-69

	<u>Av. An. Rate of Increase</u>
*Korea	20.4
Taiwan	18.0
India	11.2
Pakistan	11.1
Mexico	8.6
Philippines	8.1
Brazil and Colombia	5.8
Chile	4.8

SOURCE: UN Statistical Yearbook

2. Electricity

If electric power production is more than 100 kwh per capita per year, an annual increase of close to 10% is acceptable. If production is less than 100 kwh per capita per year, percentage increases are misleading because the starting base is so low. Below 100 kwh an increase of 10 kwh per capita per year appears to be a reasonable target.

Increase in Electric Power (kwh per capita) 1948-1969

	<u>1948</u>	<u>1969</u>	<u>Increase</u>	<u>Av. An. Percent Increase</u>	<u>Kwh per Capita per year Increase</u>
Puerto Rico	218	2582	2364	12.5	
Yugoslavia	129 ^{a/}	1128	999	10.9	
Taiwan	116 ^{a/}	824	708	10.3	
Israel	364	2156	1792	8.8	
Brazil	138	458	420	5.8	
Mexico	162	522	360	5.7	
Argentina	281	829	548	5.2	
Chile	484	746	262	2.0	
Korea	65 ^{b/}	256	191		17.4
Egypt	55 ^{c/}	225	170		10.6
Turkey	34	228	194		9.2
India	16	105	89		4.2
Morocco	44	127	83		3.9

^{a/}1949 ^{b/}1958 ^{c/}1953
SOURCE: UN World Energy Supplies

E. Education

1. Since UNESCO recently changed the statistical basis for calculating primary and secondary enrollment, international comparisons for the 1950s and 1960s are not possible. On the other hand, the new series represent a considerable improvement so that international comparisons in the future will be more reliable than they have been.

2. Third Level School Enrollment

Universities, technical schools, normal schools and others beyond the secondary level should have 500 students per 100,000 total population. Because of the enormous variations among countries in the starting point, it is hard to suggest an optimum rate of increase toward this goal.

Increase in Third Level Students Per 100,000 People
(1950-1967)

	<u>1950</u>	<u>1967</u>	<u>Change</u>
Brazil	98	251	153
Taiwan	87	1054	967
Egypt	167	565	398
Chile	160	625	465
Turkey	118	384	266
India	113	225	112
Pakistan	93	278	185
Colombia	94	268	174
Iran	34	149	115
Tunisia	50	161	111
Morocco	15	64	49
Malaysia	5	184	179
Thailand	141	102	-39
Mexico	135	338	202
Korea	126	574	448

F. Health1. Infant Mortality

If infant mortality is high to start with, say 75 per 1,000 or more, then a reduction of around 3 per 1,000 per year would be a reasonable standard until the rate is down to less than 30 per 1,000. Such a decline can be taken as evidence of a reasonably effective rural health service.

Infant Mortality Per 1,000 Live Births (1948-1969)

	<u>1948</u>	<u>1969</u>	<u>Change</u>
U.S.S.R.	81.0	25.8	-58.2
Taiwan	56.6	17.5	-39.1
Puerto Rico	78.3	28.2	-50.1
Philippines	114.4	67.2	-47.2
Colombia	136.1	70.4 '68	-65.7
Chile	147.0	91.6 '68	-55.4
U.S.A.	32.0	20.7	-11.3
Mexico	99.7	68.4	-31.3

2. Medical Personnel

Effective medical services require a variety of different kinds of personnel. Hence ratios of nurses to doctors, medical technicians to doctors and something about midwives probably are a better indicator of progress in health than the ratio of doctors to the population, although this is commonly used (partly because it's an available statistic). Suggested ratios are 2 or 3 nurses to one doctor and 4 to 6 technicians to one doctor. Rates of progress require more research.

Number of People Per Doctor

	<u>1950's</u>	<u>Late 1960's</u>	<u>Change</u>
Israel	435	410	-25
Puerto Rico	2335	1010	-1325
Turkey	3295	2260	-1035
Iran	6640	9330	-2690
India	6395	4830	-1565
Pakistan	34300	5350	-28950
Tunisia	6750	7350	590
Morocco	11370	13160	1790
Venezuela	2290	1120	-1170
Peru	4210	1890	-2320
Chile	1900	1810	-90
Colombia	2740	2220	-520
Philippines	12300	1390	-10910
Thailand	7510	8530	1020
U.S.A.	760	650	-110
Mexico	2490	1850	-640

ANNOTATED BIBLIOGRAPHY

American Institutes for Research, EVALUATIVE RESEARCH STRATEGIES AND METHODS, 1970, 160 pp. AIR, 135 North Bellefield Ave., Pittsburgh, Pa. 15213. Price \$5.00.

Adelman, Irma and Morris, Cynthia, SOCIETY, POLITICS AND ECONOMIC DEVELOPMENT, Baltimore, The Johns Hopkins Press, 1967.

Bernstein, Joel, REPORT TO THE ADMINISTRATOR ON IMPROVING AID'S PROGRAM EVALUATION, Feb. 1968, 36 pp. plus attachments. AID/Washington, D.C. 20523. ARC* Catalog No. 353.1, B 531.

Sections of this report are devoted to the meaning, purpose and rationale of program evaluation; motivational problems in getting evaluation carried out; a description of the proposed A.I.D. evaluation system; and actions required to establish this system. Attachment TAB A is titled "The Nature of AID's Assignment"; TAB B "Linking Program Evaluation and Other AID Functions"; and TAB C "What Would the Evaluation Function of Various A.I.D. Offices Be in the Proposed System?" There is also a summary of the principal general conclusions.

Boston University, REPORT OF A.I.D. PERSONNEL -- EVALUATION OF THEIR PERFORMANCE IN AFRICA: PROBLEMS AND SUGGESTIONS, Jan. 10, 1968, 67 pp. Prepared for AID/Washington by the African Studies Center, Boston University, Boston, Mass. ARC Catalog No. AFR 353.1, B 747.

The Report contains information expressed by A.I.D. personnel regarding their work in Africa and some of the frustration and difficulties encountered. There is a summary of the recommendations made by those interviewed on ways of obtaining more effective performance. Data were collected from 61 interviews conducted during the period of 1964 to 1966. Tables give a statistical summary of the replies to questions used in the survey.

Bumgardner, H.L., W. Ellis, R.A. Lynton, C.W. Jung and J.A. Rigney, A MANUAL FOR TEAM LEADERS OF TECHNICAL ASSISTANCE-INSTITUTION BUILDING PROJECTS, June 1971. Developed for AID/Washington by North Carolina State University, Raleigh, N.C.

Borton, Raymond E. (Editor), CASE STUDIES TO ACCOMPANY GETTING AGRICULTURE MOVING: ESSENTIALS FOR DEVELOPMENT AND MODERNIZATION, 1967, 302 pp. The Agricultural Development

*A.I.D. Reference Center

Council, 630 Fifth Ave., New York, N.Y. 10020. Price \$2.25.

Contains 35 case studies on agricultural development. Some cases are purely descriptive; many cover the results achieved and the significant factors contributing to achievements.

Esman, Milton J., THE INSTITUTION BUILDING CONCEPTS - AN INTERIM APPRAISAL. March 1967, 66 pp. Prepared under an A.I.D. Contract csd-763 by the Inter-University Research Program in Institution Building, Graduate School of Public and International Affairs, University of Pittsburgh, Pittsburgh, Pa. 15213. ARC Catalog No. 378.866, I 61.

Based on four field projects in Nigeria, Thailand, Ecuador, and Turkey, the author examines the points he believes are of primary importance in establishing a successful institution-building program. The environment of an institution is studied to determine the factors which, if properly used, would serve to make a program of institutional development successful. In his conclusion the author suggests 10 points which he feels should be used as guidelines by practitioners interested in institution-building theory.

German Foundation for Developing Countries, METHODS AND PROCEDURES OF EVALUATION IN DEVELOPMENT AID. Berlin Conference Report, Nov. 18-22, 1966, 211 pp. Deutsche Stiftung Für Entwicklungsländer, 53 Bonn, Simrockstrasse 1, West Germany. ARC Catalog No. 309.223, G 373.

Contains full transcripts of summaries and presentations on project and program evaluation methods used by nine international agencies and eight donor governments. The reports of six ad hoc working groups formed by the conference are included. These reports discuss the types of divisions within agencies handling evaluation, and present criteria for joint donor/recipient approaches to evaluation. Also considered are the means and methods of evaluating capital aid, training programs and the social impact of development aid. There is a 20-page bibliography.

Hayes, Samuel P., Jr., EVALUATING DEVELOPMENT PROJECTS. Technology and Society Series. UNESCO Document Number SS.65/V.17/A. Second ed., revised 1966, 116 pp. United Nations Educational, Scientific and Cultural Organization, Place de Fontenoy, Paris 7e, France. U.S. Sales Office: UNESCO Publications Center, P.O. Box 433, New York, N.Y. 10016. Price \$2.50. ARC Catalog No. 309.22072, H 418.

This publication was first published in 1959 under the title, MEASURING THE RESULTS OF DEVELOPMENT PROJECTS. It suggests

analytical techniques for measuring social and economic development projects to find out just how effective the projects have been. Describes steps which should be taken before project evaluation begins and identifies the kind of data which project evaluators need. Suggests ways to collect data and how to analyze and interpret them. An appendix provides a brief discussion of methods of sample selection, classifying, coding, tabulating and summarizing data. There is a three-page bibliography.

Herzog, Elizabeth, SOME GUIDELINES FOR EVALUATIVE RESEARCH, U.S. Dept. of H.E.W., Children's Bureau, Washington, D.C., 1959.

Higgins, Benjamin, "The Evaluation of Technical Assistance," INTERNATIONAL JOURNAL, Vol. XXV, No. 1, Winter 1969-70, pp. 34-55. Canadian Institute of International Affairs, 31 Wellesley St. East, Toronto 284, Canada. Single copy price \$2.00. U.S. Department of State Library No. I 638.

The author, a professor of economics at the University of Montreal, draws on his experience with technical assistance missions in ten countries, and with two special evaluation missions for OECD and the UN in Greece and Libya, to outline what he considers to be the main problems of evaluating technical assistance programs. He lists certain basic requirements of the development process indicating that technical assistance is only one factor among many which are necessary for economic development. He describes certain common complaints advanced by donor and recipient governments about technical assistance, and suggests, in broad terms, some of the questions which need to be asked in evaluating such programs.

Higgins, Benjamin, Alexander Stavrianopoulos and Angus Maddison, FOREIGN SKILLS AND TECHNICAL ASSISTANCE IN GREEK DEVELOPMENT, 1966, 169 pp. Development Center of the Organization for Economic Cooperation and Development.

U.S. address: OECD Publications Center, Suite 1305, 1750 Pennsylvania Ave., N.W., Washington, D.C. 20036. Price \$3.50. U.S. Department of State Library No. HC 295.M 24.

The report is an appraisal of the technical assistance furnished Greece from bilateral and multilateral sources during the period roughly between 1954 and 1963. Consideration is given to high-level policy advisors as well as specialized technicians operating at the grassroots level. There is an examination of: (1) the economic and social situation in Greece during the time covered, (2) the skills needed for

rapid growth, (3) how foreign training supplemented Greek skills, (4) the channels of aid, (5) the role of different donors, and (6) the efficiency of technical assistance administration. One conclusion drawn was the importance of utilizing regional planning within the overall framework of technical assistance. Finally, the report considers how Greece, as a donor, has helped other developing countries.

Hirschman, Albert O., DEVELOPMENT PROJECTS OBSERVED, Brookings Institution.

Hyman, Herbert, SURVEY DESIGN AND ANALYSIS, The Free Press, Glencoe, Illinois, 1955.

Hyman, Herbert and Wright, Charles, "Evaluating Social Action Programs", in Paul Lazarsfeld, William Sewell and Harold Wilensky (Eds.), USES OF SOCIOLOGY, Basic Books, New York, 1967.

Jacoby, Neil H., EVALUATION OF AGRARIAN STRUCTURES AND AGRARIAN REFORM PROGRAMS, FAO Agricultural Studies No. 69, 1966.

Jacoby, Neil H., AN EVALUATION OF U.S. ECONOMIC AID TO FREE CHINA, 1951-1965. A.I.D. Discussion Paper No. 11. January 1966, 99 pp. Prepared under Contract to the Bureau for the Far East, AID/Washington, D.C. 20523. ARC Catalog No. CH 309.223551249, J 17.

The report is a comprehensive analysis of the U.S. aid program to Taiwan. In the Preface, A.I.D. Administrator Bell identifies the report as a milestone study which will be of use for years to come. The author develops his own tests for deciding whether aid has or has not been useful. Economics, social and political development are discussed, and there is a summary of lessons learned relative to the U.S. foreign economic aid policy.

Kerwin, Harry W., AN ANALYSIS AND EVALUATION OF THE PROGRAM OF TECHNICAL ASSISTANCE TO EDUCATION CONDUCTED IN IRAN BY THE GOVERNMENT OF THE UNITED STATES FROM 1952 TO 1962. 1964, 285 pp. A doctoral dissertation submitted to the Graduate School of Education at American University, Washington, D.C. ARC Catalog No. IR 370.0955, K 41.

The dissertation gives a detailed historical overview of practically all education programs in Iran and how they were supported by U.S. technical assistance efforts. In the summary chapter the author evaluates the positive and negative factors affecting these programs. These factors are divided into the following five categories: personnel, economic, political, administrative and socio-cultural.

Legum, Colin (Ed.), THE FIRST U.N. DEVELOPMENT DECADE AND ITS LESSONS FOR THE 1970s, 312 pp., Praeger Publishers, Inc., 111 Fourth Ave., New York, N.Y. 10003. Price \$15.00. U.S. Department of State Library No. JX 1977. F 56.

The publication was issued in cooperation with the Vienna Institute of Development. It includes a review of technical assistance activities during the 1960s. The role of both the developed and the developing countries are discussed. Ten leaders concerned with economic development programs explain their views regarding technical assistance and some of the lessons which have been learned. Other authors present their observations and comments. The total input of ideas results in a variety of opinions regarding the best way to proceed with the development decade of the 1970s.

Maynard, Paul J., & Polachart Kraiboon, EVALUATION OF THE MUONG PHIENG CLUSTER AREA, September, 1969, prepared for USAID/Vientiane, Laos by Stanford Research Institute.

Niehoff, Arthur H. (Editor) A CASE BOOK OF SOCIAL CHANGE, 1966, 312 pp. Aldine Publishing Co., 320 West Adams St., Chicago, Ill. 60606.

Nineteen case studies evaluating attempts to introduce change in 16 different developing countries. There is also a chapter on the process of innovation.

Normington, Louis W., TEACHER EDUCATION AND THE AGENCY FOR INTERNATIONAL DEVELOPMENT, 1970, 186 pp. Prepared for the Office of Education and Human Resources, Bureau for Technical Assistance, AID/Washington, by the American Association of Colleges of Teacher Education, One Dupont Circle, Washington, D.C. 20036.

Contains descriptions of technical assistance programs and case studies.

OECD, THE EVALUATION OF TECHNICAL ASSISTANCE, Technical Assistance Evaluation Studies Series, 1969, 134 pp. Organization for Economic Cooperation and Development, Paris. U.S. address: OECD, Publications Center, Suite 1305, 1750 Pennsylvania Avenue, N.W., Washington, D.C. 20036. Price: \$2.90. U.S. State Department Library Catalogue No. HC 60.064.

This report is the first in a series based on lessons learned from the OEEC-OECD technical assistance program which has been in operation since 1969. Part I of this publication is a study of evaluation plus appended case studies prepared by

the OECD Secretariat. Sections are devoted to a discussion of the objectives, types, methods and limitations of evaluation. Part II contains reports on technical assistance evaluation methods used by Sweden, the German Federal Republic and the United States. Part III is comprised of statements regarding the OECD evaluation report made at the OECD Technical Cooperation Committee Meeting, November 8, 1968. A 14-page bibliography lists over 100 publications on evaluation from international agencies, participating OECD countries and non-governmental organizations.

Owens, Edgar, and Robert Shaw, DEVELOPMENT RECONSIDERED, Heath Lexington Books, Lexington, Mass., 1972.

Phillips, Hiram S., HANDBOOK FOR DEVELOPMENT: CHANGING ENVIRONMENTS AND INSTITUTIONS, June 1967, 344 pp. Office of Institutional Development, Bureau for Latin America, AID/Washington. ARC Catalog No. 309.2, P559.

See Chapter VII, "Judging Progress". Also note case studies, Chapters VIII through XI.

Rice, E. B., EXTENSION IN THE ANDES: AN EVALUATION OF OFFICIAL U. S. ASSISTANCE TO AGRICULTURAL EXTENSION SERVICES IN CENTRAL AND SOUTH AMERICA, AID EVALUATION PAPER No. 3 (condensation and 3A complete), Evaluation Staff, Bureau for Program and Policy Coordination, AID/Washington.

An evaluation of official U.S. Assistance to agricultural extension services in twelve countries of Central and South America between 1942 and 1968. The study addresses two questions: was the US effective in building viable extension institutions, and have those institutions had a significant impact on agricultural productivity? The Author concludes that on both counts the programs accomplished far less than expected, partly because the role of extension is rural development was misunderstood.

Schultz, Theodore W., THE ECONOMIC VALUE OF EDUCATION, Columbia University Press, New York, 1963 (69 pp. plus 18 pp. of bibliography).

Sheldon, Eleanor B. and Moore, Wilbert E. (Eds.), INDICATORS FOR SOCIAL CHANGE: CONCEPTS AND MEASUREMENTS, New York, Russell Sage Foundation, 1968.

Smart, Lyman F. (Editor), PROCEEDINGS: REGIONAL CONFERENCE ON INSTITUTION BUILDING. Conference held under the auspices of the Utah International Education Consortium and the U. S. Agency for International Development in

Logan, Utah, Aug. 17-21, 1970.

See particularly the report of Committee G, pages 53-61, titled "Utilization of Project Planning, Review and Assessment of Maturity to Facilitate Maximum Project Results". See also W. N. Thompson's paper, "Ideas and Procedures for the Evaluation of Progress and Maturity in Institution Building", pp. 129-140, and Jackson A. Rigney's, "Guidelines for Achieving the Most from Participation in Overseas Contracts", pp. 141-149.

Spruyt, Dirk J., Francis B. Elder, Simon D. Messing, Mary K. Wade, Brooks Ryder, Julius S. Prince and Yohannes Tseghe, "Ethiopia's Health Program -- Its Impact on Community Health", in the ETHIOPIAN MEDICAL JOURNAL, Vol. 5, No. 3, July 1967, 87 pp. Ethiopian Medical Assn., Addis Ababa, Ethiopia. ARC Catalog No. ET 614.0963 E 84.

The evaluation of public health services made in this report covers the six-year period from 1961 to 1967. Health conditions in three selected health center communities and three matched control communities were studied at the time the health center programs were being initiated and again three to four years later in order to measure program effectiveness. The period between these baseline and resurvey studies was used to carry out several special studies including a functional analysis of each health center program. An analysis of Health Service activities is made, diseases identified, health attitudes studied, and aspirations noted. One of the authors notes that if a program is to improve there must be a critical and honest examination of mistakes as well as recognized successes. As a result of this evaluation study, twelve specific recommendations for improvements in the Ethiopian health program are made.

Suchman, Edward A., EVALUATIVE RESEARCH: PRINCIPLES AND PRACTICE IN PUBLIC SERVICE AND SOCIAL ACTION PROGRAMS, Russell Sage Foundation, New York, 1967.

Thomas, D. Woods, and Judith G. Fender (Eds.), PROCEEDINGS: CONFERENCE ON INSTITUTION BUILDING AND TECHNICAL ASSISTANCE, Sponsored by the Agency for International Development and the Committee on Institutional Cooperation, Dec. 4-5, 1969, 164 pp. Committee on Institutional Cooperation, 1603 Orrington Ave., Suite 790, Evanston, Illinois 60201. ARC Catalog No. 309.223 A 265K.

UNESCO, "Evaluation Techniques", INTERNATIONAL SOCIAL SCIENCE BULLETIN, Vol. VII, No. 3, 1955, UNESCO, 19 Avenue Kleber, Paris 16^e, France.

United Nations, EVALUATION OF PROGRAMMES OF TECHNICAL CO-OPERATION, AGENDA ITEM 15, Document E/4151, May 3, 1966, 92 pp. Report of the Secretary General of the Economic and Social Council, United Nations, New York, N.Y. ARC Catalog No. 309.223, U 58c.

This report is in response to a resolution of the UN Economic and Social Council calling for a systematic and objective evaluation of the impact and effectiveness of technical cooperation carried out by the United Nations family of organizations. Addenda 1-3 of this report reproduce the intensive country evaluation studies carried out in Thailand, Chile and Tunis. The report of the Secretary General summarizes the scope and method of the country studies and his findings, observations and recommendations based on them. The country reports provide information on the deficiencies and shortcomings as well as the successes of technical cooperation programs. Various methods and standards are reviewed by which objective evaluative judgments can be made. It is pointed out that program evaluation will contribute to increased project effectiveness, provide perspective for future programs and assist in the formulation of essential standards for the evaluation process.

United Nations, Food and Agriculture Organization, REVIEW AND APPRAISAL OF PROGRESS IN THE AGRICULTURAL SECTOR, Working paper submitted to Seventh Session of UN Committee for Development Planning ("The Tinbergen Committee"), Geneva, April 1971

United Nations, ECOSOC, FRAMEWORK FOR APPRAISING PROGRESS DURING THE SECOND DEVELOPMENT DECADE, 1971.

USAID/Vientiane, Laos, EVALUATION, JOINT RLG/USAID ACCELERATED RICE PRODUCTION PROGRAM 1967 - 1969, 203 pp. Agriculture Division, USAID/Vientiane, Laos. ARC Catalog No. LS 633.18, U 58.

This in-depth study covering three years of effort to increase rice production in Laos points up the importance of joint host government - U.S. cooperation in project evaluation. Seventeen points in project development are identified, and there is listed a group of actions considered necessary to further increase aid effectiveness. Country background data are given. The project goals and program are discussed and a statistical base for program evaluation is outlined. The use of aerial photography for a land-use inventory is suggested.

U.S. Dept. of State, A.I.D., BUILDING INSTITUTIONS TO SERVE AGRICULTURE: A SUMMARY REPORT OF THE C.I.C.-A.I.D. RURAL

DEVELOPMENT RESEARCH PROJECT, Sept. 30, 1968, 236 pp.
Published for AID/Washington by the Committee on
Institutional Cooperation, Purdue Research Foundation,
Lafayette, Ind.

See particularly Chapter IV, "Effects on Host Institutions,"
Chapter VI, Section 3, "Measurement of Institutional Progress
and Maturity," and Chapter VII, "Basic Factors Conditioning
Success."

U.S. Dept. of State, A.I.D., REPORT ON PROJECT MANAGEMENT
PROBLEMS, Nov. 1969, 38 pp. AID/Washington, D.C. 20523.
ARC Catalog No. 353.1, H 541.

A staff report prepared for A.I.D. management by a special
study group composed of representatives from the Regional
Bureaus and the Auditor General. The study was based on
in-depth interviews of 106 A.I.D. project managers, and other
supervisory U.S. officials in eight recipient countries. The
study teams developed 16 specific findings. For each of these,
they present a brief discussion and a series of recommendations
designed to improve A.I.D. project management systems and
overcome the problems revealed by the survey.

U.S. Dept. of State, A.I.D., PARTICIPANT ASSESSMENT OF A.I.D.
TRAINING PROGRAMS - FIRST ANNUAL REPORT (Office of
International Training, A.I.D.), May 1969.

U.S. Dept. of State, A.I.D., WORLD-WIDE EVALUATION OF PARTICIPANT
TRAINING - SUMMARY OF PRINCIPAL FINDINGS AND PRIMARY
RECOMMENDATIONS FOR ACTION, March 1966.

U.S. Dept. of State, A.I.D., SPRING REVIEW OF NEW CEREAL
VARIETIES - 1969, SPRING REVIEW OF ICI's - 1969, SPRING
REVIEW OF LAND REFORM - 1970, SPRING REVIEW OF POPULATION
PROGRAMS - 1970.

Extensive evaluative documentation prepared for each of the
above topics covering background, issues, analyses, Mission
reports of experience -- and for Land Reform experience also
in non-AID countries -- summaries and recommendations.
Documentation was prepared for Agency-wide evaluations, and
to serve as permanent resource material. A detailed listing
of the papers prepared, including particular countries examined,
is available in the card catalogue of the AID Reference Center
(ARC).

U.S. Dept. of State, A.I.D., UTILIZATION OF ECONOMIC RESEARCH,
PPC/PDA Evaluation Paper #4, October, 1971.

U.S. Dept. of State, A.I.D., SPRING REVIEW OF THE NEW CEREAL
VARIETIES: A PERSPECTIVE, PPC/Evaluation Staff
Evaluation Paper #2, January, 1970.

Webb, Eugene J. et al, UNOBTRUSIVE MEASURES: NONREACTIVE
RESEARCH IN THE SOCIAL SCIENCES, Chicago: Rand McNally,
1966.

Winfield, Gerald F., BEHAVIOR CHANGE FACTORS IN DEVELOPMENT,
1971, 55 pp. Washington Training Center, AID/Washington.

DEPARTMENT OF STATE
AGENCY FOR INTERNATIONAL DEVELOPMENT
WASHINGTON, D.C. 20523

November 15, 1974

National Criminal Justice
Research Service
955 L'Enfant Plaza S.W.
Washington, D.C. 20024

ATTN: Mr. Glomb

Dear Sir:

Enclosed is the copy of AID's Evaluation Handbook
which you requested yesterday.

Very truly yours,



Richard H. Eney
Evaluation Officer
Program Methods and Evaluation

Enclosure:

1. Evaluation Handbook

END