74185

A TRAINING IMPACT EVALUATION METHODOLOGY AND INITIAL OPERATIONAL GUIDE

Prepared for:

Agency for International Development

Africa Bureau

AFR/TR/EHR

Prepared by:
Creative Associates International, Inc.
5301 Wisconsin Avenue, N.W., Suite 700
Washington, D.C. 20015

October 18, 1991

TABLE OF CONTENTS

ACKNOWLEDGMENTS

We acknowledge with thanks the interest and helpful support of Cameron Bonner, Deborah Diaz, Gene Peuse, Bill Renison and Yvonne Andualem of the A.I.D. Africa Bureau in carrying out this study.

We appreciate also the contributions of the people inside and outside A.I.D. who were interviewed during the study. They are:

Jerry Britan, CDIE
Maury Brown, CDIE
Hazel L. Bush, AAI
Joe Carney, LAC
Han Chan, IBRD
Cindy Clapp-Wincek, consultant
Peter Davis, APRE
Graham Donaldson, IBRD
Tom Donnelly, OIT
Martin Hanratty, ENE
Gary Hansen, CDIE
Peter Hansen, United Nations
Hans Heinrich Thias, IBRD
Chris Hermann, APRE
Robin Horn, AFR
Anita Johnson, AAI
Angela E. V. King, United Nations
Krishna Kumar, CDIE
Heather Monroe, AAI
Ron Ridker, IBRD
Barbara Searle, IBRD
Emmy Simmons, AFR
Nena Vreeland, CDIE
Jerry Wolgin, AFR
Mervyn L. Weiner, IBRD retired
Robert Yeuker, consultant

The Creative Associates International, Inc. Staff who participated in the study were Herb Turner, Brenda Bryant, and Andrea Bosch.

INTRODUCTION

In a review of A.I.D. experience with participant training, the Africa Bureau, "... did not find an existing body of comprehensive analysis concerning the impact of participant training on development and accompanying impact indicators, nor any model to ascertain how the Agency's investments in training have contributed to sustainable development. Moreover, many past A.I.D. training efforts have been project-oriented. This is not currently "in synch" with the more expansive A.I.D. and other donor approaches to training and development. With the recent re-orientation of A.I.D.'s Africa Bureau towards a program-CPSP approach and a more wide-ranging approach to human resource development issues, the role of training in the development process needs to be re-examined..." In the design of its most recent participant training program, entitled African Training for Leadership and Skills Project (ATLAS), the Africa Bureau called for the development of a methodology for undertaking a long-term impact study of A.I.D.-funded training on development. ATLAS was the impetus for the undertaking.

Task Summary

In June, 1991, Creative Associates International, Inc. was awarded a task order under PDC-5832-I-00-0095-00 to address training impact evaluation. The purposes of the study are, first, to establish a framework and methodology for evaluation of past investment in A.I.D.- funded training and, second, to propose guidelines and impact indicators that assist in the determination and monitoring of future A.I.D. training investments, including ATLAS, under the Africa Bureau regional and bilateral training portfolio. The project's scope of work calls for these products:

- A methodological framework to evaluate the impact of A.I.D.-sponsored training using an array of performance indicators that relate to DFA strategic objectives;
- A practical set of impact indicators such as rates of return, income, increases in productivity, employment generation, occupational status, organizational position, labor mobility, institutional performance and other broad-based higher-level indicators;
- A prioritized list of recommended African countries and projects and sectors that would serve to test the methodology;
- A questionnaire that corresponds to the analytical framework;
- An implementation and budget plan.

This document, in two parts, contains each of the products listed above.

The development of the impact evaluation methodology was undertaken by a three person team that included an evaluation specialist, a training specialist and a research assistant. The design was prepared during the period June 21 to October 18, 1991 and included the completion of these tasks:

- An extensive literature search which focussed upon a range of evaluative and research models and methodologies, on existing A.I.D. guidance material and on documentation of evaluation studies and compendia;

- A comparative analysis of models and methodologies to determine their relevance, rigor and applicability;

- Interviews within Africa Bureau staff and with staff in other A.I.D. bureaus; discussions with the World Bank, the United Nations, and the African American Institute;

- Articulation of a succinct definition of impact and an explication of the several kinds of impact which the evaluation framework would embrace;

- The tentative formulation of a specific theory of induced change to serve as the basis for the development of an impact evaluation framework;

- Construction of conceptual models for the components of the training process being evaluated;

- Formulation of a comprehensive impact evaluation framework;

- Identification of analytical tools and measurement techniques to be employed within the impact evaluation framework.

Prior to embarking on the work plan's first task, the project team made explicit the criteria to which it would adhere in developing an impact evaluation methodology for the Africa Bureau. The team emphasized the need for a useful, reliable and comprehensive methodology that would go beyond evaluation approaches which are currently available. These precepts emerged and guided the team:

- The elements and actions needed to enhance the evaluability of training are precisely those needed to maximize the impact of training on host country development.

  The purpose of the proposed impact evaluation methodology is to identify and measure the impact of training on development. The goal of the methodology is to increase the impact of training on development. The hypothesized causal contribution between purpose and goal is that a methodology which is capable of measuring impact is also capable of enhancing the probability of achieving impact. If the evaluation methodology does not encompass those elements which contribute to the achievement of impact, then it is not a viable methodology.

  It follows that such a methodology cannot be limited in scope to a minimal prescription of what training results to evaluate and how to evaluate them. It must also identify the preconditions which are necessary for evaluation of impact and provide guidance for creating them. These preconditions address both the formative and summative stages of evaluation and include (a) the formulation of explicit post-training objectives, (b) relating the objectives to host country development plans and operations, (c) identification and understanding of contextual factors, and (d) planned approaches to information collection and analysis.

- To permit effective evaluation and feedback, the cycle of design, training and post-training activities must be integrated through the rigorous application of a single analytical framework which accommodates both linear and non-linear change.

- The theoretical soundness and scientific rigor of the methodology must be defensible and produce an acceptable level of confidence for decision-makers. It must be compatible with established theory found in the literature.

- The methodology must be understandable and implementable. It should not require a sophisticated level of research or analytical skills. The time required and the cost of using the methodology should be within the limited resources available to AID and the host country.

- The methodology must encompass activities and effects in both the public and private sectors and include planned and unplanned effects.

- The methodology must accommodate the total context of program implementation by defining and assessing the relationships among the objectives of the trainee, the host country government, USAID and the implementing agencies.

Overview of the Report

The content of this report is organized as follows:

Chapter I, A Conceptual Approach to the Formulation of an Impact Evaluation Framework for Human Resource Development, describes the creation of a methodological framework for the evaluation of development impact caused by training. There is an orderly progression from an empirical/logical model of the development process through a series of conceptual design components to the evaluation of the training project/program. The components are:

1. An empirical/logical model of the development process called the goal hierarchy

2. The design of the training project/program

3. A specific theory of induced change to characterize and explain the project/program

4. A methodological framework for evaluating impact

5. The analytical tools and measurement techniques

6. The evaluation of the training project/program

Chapter II, Impact and Its Measurement, posits a single, succinct definition of developmental impact. It then explains the phenomena in all its multiple forms and occurrences with a series of ancillary statements of definition, each with an illustrative example. The chapter then addresses a number of definitional and conceptual issues.

Under the heading, Types of Changes that Constitute Development Impact, the concept of impact leverage is introduced. This key concept is discussed in subsequent chapters and is more fully developed in Chapter III.

There is a brief description of three main types of developmental change: (a) primary/secondary - first/second generation, (b) replication - spread effect and (c) multiplier effect.

The section called, Changes that do not Constitute Development Impact, defines the necessary preconditions for impact and describes their role in the development process.

In a section entitled, Impact Evaluation, the report describes the characteristics of impact evaluation and the expectations, constraints and limitations which are inherent in an impact evaluation methodology.

Evaluation Models examines four classes of existing methodologies to determine their relevance and utility for the measurement/estimation of the development impact of training. They are: (1) program evaluation, including the goal attainment model, the goal-free model and the systems model, (2) social science research/experimental design, (3) economic analysis (benefit/cost, cost/effectiveness and rate of return) and (4) the constructivist/participative (fourth generation) approach.

After noting the current approaches to impact evaluation in A.I.D., the report describes:

- A specific theory of induced development change to characterize and explain human resource development programs; and,
- A conceptual model for evaluating the development impact of human resource development programs.

This section asserts the need to build and, over time, to validate a specific theory of induced developmental change caused by human resource development programs in general and training in particular. Such a specific theory would have two necessary functions:

- To provide the basis for planning effective and beneficial investments in human resource development, i.e., for guiding project design;
- To provide the basis for creating a comprehensive impact evaluation framework, which, in turn, has two functions: (a) to measure induced change and by so doing, (b) to validate the specific theory.

There is a first, tentative description of some of the features of a specific theory.

This section also briefly describes the scope of the impact evaluation framework and summarizes the architecture and the constituent elements which derive directly from the theoretical construct tentatively portrayed above. This is a conceptual model, not an operational model.

Chapter III is entitled, A Proposed Approach to the Evaluation of Development Impact. This chapter (a) introduces the proposed impact framework, (b) explains its coverage of the design - implementation -evaluation stages of the training project cycle, (c) characterizes the dimensions within which impact may occur and (d) defines and describes the conditions and the analytical methods and devices which affect the use of the framework. The key sections of the chapter are briefly noted in the following paragraphs.

Impact evaluation and the stages of a training program

The conceptual framework for impact evaluation encompasses the design, implementation and post-training stages of the cycle, draws on, and adapts existing design and evaluation concepts and introduces new approaches. It specifies design and evaluative tools and techniques to be used at the several stages in the project cycle.

The section entitled, Planning/Design Stage, describes how to translate the salient features of the host country's development plans and priorities into a form which the trainee can use in setting his/her own career objectives and in formulating a career pathway to meet those objectives.

The underlying hypothesis is that the more the trainee learns about his/her country's development needs, the greater will be the trainee's contribution.

An empirical/logical model of induced developmental change called the generalized goal hierarchy is introduced here. It displays the progression of developmental effects triggered by a development intervention/investment. Ideally, the generalized goal hierarchy is (a) a valid reflection of the operational realities of the developing countries, (b) universally applicable to a variety of development investments, regardless of sector, geographic location, etc., and (c) useful in predicting as well as evaluating the outcome of new development assistance initiatives.

For each training project an individualized goal hierarchy is prepared. This is an individual career plan, derived from the generalized development goal hierarchy, in which the specific qualifications and interests of the individual are integrated or harmonized with the needs, plans and circumstances of the host country. It is jointly drafted in close collaboration between the trainee, the relevant government ministries, any future employer, USAID and the implementing agents.

The formulation, reformulation and verification of causal hypotheses is central to the methodological framework. Causal hypotheses will inhere within the sequence of developmental change displayed in the generalized goal hierarchy and its derivative, the individualized goal hierarchy. These goal hierarchies will be informed, respectively, by country-specific data on national/sectoral objectives and priorities and by trainee-specific data on personal objectives and priorities.

The section called, Role of the Trainee in the Training Project Cycle, asserts (a) that the viability of the proposed impact methodology is highly dependent upon the trainee's active engagement in all stages of the project cycle, and (b) that exposing the trainee to host country planning and evaluation information could have a powerful and beneficial effect on his/her contribution to host country development.

Effective impact evaluation will require anticipation and observation of the developmental change process as it occurs rather than years later in a conventional, one-time, ex post impact evaluation. Here the role of the trainee is critical.

The section called, Action Stage: Post Training Activities, includes employment, participation in professional activities, self employment, research, networking, the twinning arrangement with the American university/professional society, teaching, etc., and discusses the pre-evaluative and evaluative actions required to assess their developmental consequences.

Impact evaluation and the various dimensions of change

This section describes the continuum within which development impact may occur. It begins with the individual trainee/change agent whose work initiates and sustains the change process.

The second dimension concerns the function of the organization/institution in the development process. This section presents an institution building model which examines that function, delineates the several stages of growth of the organization/institution in the development process and describes the contribution of the graduate at each stage. The institution building model fits within the goal hierarchy structure and defines the linkages between the institution and (a) the subsector/sector system and (b) the target group/beneficiaries.

The subsector/sector dimension is an arena in which resources are mobilized, allotted, invested, processed and distributed. This section examines the possibilities for the design and evaluation of human resource development investments within the subsectoral/sectoral system.

The section entitled, The Target Group/Beneficiary Dimension considers three levels of impact on the target group/beneficiaries: the enhancement of capacity, and consequent increases in performance and improvements in benefits.

In the section called The National Dimension, the key concept of impact leverage as it may be found at national level is more fully developed and examples are provided. Impact at the national level is seen as an aggregate of development effects at prior levels. Finally the limitations of observing the development impact of training at the multinational dimension are considered and placed in perspective.

Conditions affecting the use of the proposed impact evaluation framework

This section considers four main aspects of the impact evaluation methodology and process:

1. Principles, elements and criteria of the evaluation process
2. Causal hypotheses and methods of validation
3. Measurement of developmental change - objectively verifiable indicators
4. Baseline

Impact indicators

Finally, this chapter discusses the limitations of using indicators for purposes of comparability and aggregation. It examines the shortcomings of generic, macro-level indicators for measuring and explaining induced change. These cautionary notions suggest that the broad-based, higher-level indicators, however useful in policy dialogue and overall program planning, may have limited predictive, explanatory or learning value at the operational level. With these notions in mind, the chapter establishes criteria for the formulation of impact indicators and proposes sets of indicators organized by social and economic sectors for use within the methodological framework.

CHAPTER I. A CONCEPTUAL APPROACH TO THE FORMULATION OF AN IMPACT EVALUATION FRAMEWORK FOR HUMAN RESOURCE DEVELOPMENT

This chapter describes the line of reasoning which was followed in deriving a methodological framework for the evaluation of development impact caused by training. The line of reasoning moves in an orderly progression from an empirical/logical model of the development process through a series of conceptual design components to the evaluation of the training project/program. The components are:

- An empirical/logical model of the development process;
- The design of the training project/program;
- A specific theory of induced change to characterize and explain the project/program;
- A methodological framework for evaluating impact;
- The analytical tools and measurement techniques;
- The evaluation of the training project/program.

Each of the components in the progression is briefly explained below. In subsequent chapters, each is elaborated. Chart 1 on the next page displays the progression.

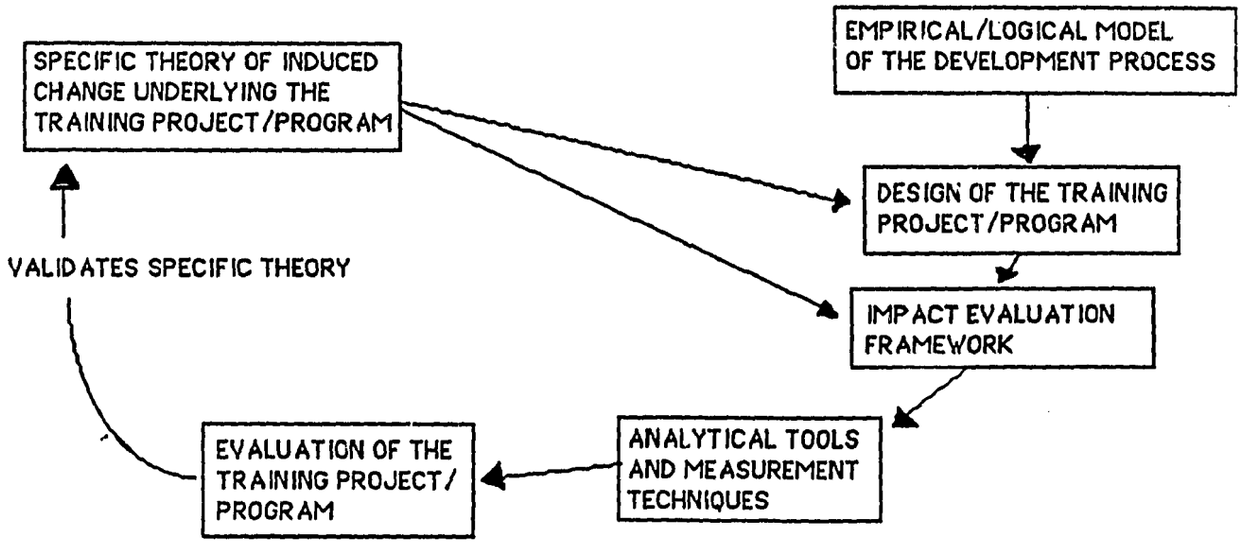
An Empirical/Logical Model of the Development Process

At the outset, the basic reality is that understanding of the process of development is both imperfect and incomplete. It is difficult, comprehensively and accurately to:

- Predict if desired development could be caused to occur, what kinds of causative interventions/investments would be most effective, and how, why, where and when it might occur;
- Discover, retrospectively, what kinds of development did occur and to explain how, why, where and when it happened;
- Measure development changes objectively and reliably.

The first challenge of the investigation was to articulate an empirical/logical model of the development process. The basic requirement was that the model be conceptually capable of explaining and predicting the impact on development of an investment in training. The development model had to fulfill several criteria as follows:

THE LINE OF REASONING FOLLOWED IN THE DEVELOPMENT OF THE METHODOLOGICAL FRAMEWORK



- Descriptive

  The development model should be, at least in part, descriptive, i.e., based upon experience. This criterion can be met only to a limited degree since there is not an agreed, substantive body of research and evaluation findings on the actual occurrence of development, paradoxically, because no viable impact evaluation methodologies exist and very few authoritative impact studies have been completed.

- Responsive to Prescription

  The second criterion called for the development model to be responsive to prescriptive information, i.e., what policy-makers and planners wished to achieve. This criterion could readily be met since most available information was prescriptive in character: policy, planning and design documentation at the macro, sectoral, subsectoral and project levels and supporting feasibility analyses.

- Logical

  The third criterion was that the model be logically sound; that any internal causal linkages be demonstrably necessary and sufficient for the achievement of the next stage.

The empirical/logical model is a goal hierarchy with four major levels of induced change which could be expected to occur as a direct consequence of an investment in training:

- Institutional capacity/performance;
- The capacity and performance of the subsectoral/sectoral system;
- Target group/beneficiary capacity, performance and benefits; and,
- National development.

The structure and functions of the goal hierarchy are described in Chapter III. Modelling of the development process led to the next stage, which was to conceptualize the intervention/investment, i.e., the training project/program.

The Design of the Training Project/Program

The individual or group of trainees in a scholarship is perceived as a discrete development project with an inherent cycle of planning, training and post-training activities. Within this cycle, causative actions can be inserted to change the capacity, and consequently the performance, and finally the contribution to development, of the trainee. Critical to the evaluation of the effectiveness of the cycle are the incorporation of preconditions for evaluation, evaluation and the feedback of findings, inferences and conclusions into replanning. The project cycle is treated in greater detail in Chapter III.

Chart 2, which displays graphically the relationship between the development process and the training project/program, is on the next page. This chart is also used at the beginning of Chapter III to show the basis for deriving the first of two key sets of causal hypotheses within the goal hierarchy model. This set of causal hypotheses are intended to link the trainee/training to development impact.

CONDITIONS AFFECTING DEVELOPMENT IMPACT OF ATLAS TRAINEE-BASIS FOR CAUSAL HYPOTHESES

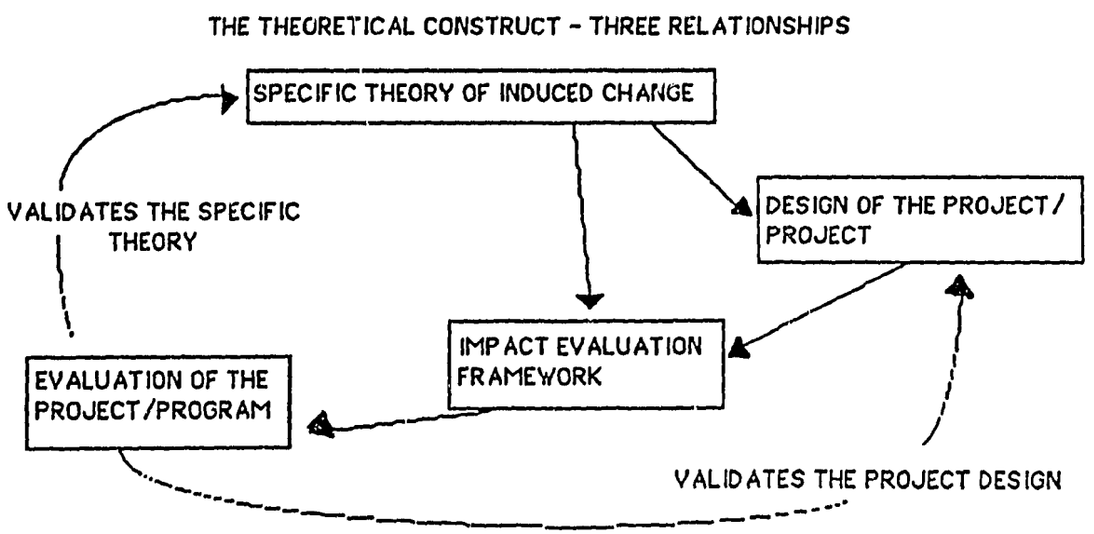
	THE GOAL HIERARCHY	NECESSARY BUT NOT SUFFICIENT CONDITION FOR DEVELOPMENT IMPACT			EXOGENOUS VARIABLES SUSTAINABILITY	INDICATORS OF: •preconditions for development impact •development impact
		PRE-DEPARTURE PLANNING AND DESIGN	IMPLEMENTATION OF TRAINING PROGRAM	POST TRAINING ACTIVITIES		
ASSOCIATIONS OF DEVELOPMENT IMPACT	AFRICA	cross national planning and coordination	-broad based commitment -advocacy -beneficiary evidence	-commitment of resources -mechanism for coordination	•POLITICAL VIABILITY AND STABILITY • SOCIAL VIABILITY AND STABILITY •DEMOCRATIC PRACTICES •PROFESSIONAL MOBILITY •ADEQUACY AND ACCESSIBILITY OF RESOURCES, INFRASTRUCTURE AND SERVICES •VIABILITY OF INTELLECTUAL SERVICES, EG., RESEARCH AND DEVELOPMENT, EDUCATION AND TRAINING, ORGANIZATIONAL DEVELOPMENT	-income/consumption/savings -employment -investment improved economic social, political -improved productivity/production improved life quality
	COUNTRY	awareness, support and participation of national leaders in ATLAS interest and support of key planning and operating agencies viable development plan	-broad based commitment -advocacy -beneficiary evidence	-commitment of resources -appropriate policy choices		
DEVELOPMENT IMPACT	TARGET GROUPS-BENEFICIARIES	-representation at planning stage -target group needs, capacities and performance known to trainee	monitoring and support to trainee reflects unique target group requirement	-work assignment appropriate to unique target group requirements		-income -employment/productivity advancement,consumption -access to economic/social services life quality improvement.
PRECONDITIONS FOR DEVELOPMENT IMPACT	SECTOR/SUBSECTOR SYSTEM	-participation of government, business, industry and academia in planning -access to sector planning -donor coordination -adequacy of needed sector resources and services	capacity for managing flow of resources and actions among institutions and donors	appropriate assignment within sector/subsector coordination within sector and among sectors adequacy of sectoral resources and info.		-sectoral/subsectoral needs for trainee knowledge -sectoral/subsectoral capacity -sector/subsectoral performance
	ORGANIZATION/ INSTITUTION	-participation of top management in planning -strategic plan for use of ATLAS training -plan for org. development	monitors and supports trainee monitors and trainee research/field work trains work units and prepares for reentry of trainee	-ensures appropriate work assignment for trainee -provides support services -provides on job training -provides adequate income and benefits		-demonstrated capacity of institution to utilize trainee knowledge -improvements in institutional productivity and production -improvements in institutional outreach
	INDIVIDUAL ATLAS TRAINEE	trainee actively participates in planning career trainee has mentor host country, employer, AAI, USAID participate in planning access to development plans, priorities, human resource plans, labor market analysis	faculty guidance contact with host country mentor research/fieldwork relevant to career plan and to host country dev plans and priorities	-employment in chosen field -involvement in prof. networks -practical training/ updating/teaching		-personal income -career advancement -contributions to development

A Specific Theory of Induced Change

If the training project/program is to be designed and evaluated, it must derive from a specific theory of induced change. The line of reasoning specified the need for a theoretical construct which would inform the design and evaluation of the project and would permit the formulation of an impact evaluation framework. Three relationships are considered here:

- Between a project design and the specific theory;
- Between the impact evaluation methodology and the specific theory; and,
- Between the project evaluation and the specific theory.

The chart graphically displays the relationships.



The relationship between a project/program design and the specific theory of induced change which explains the project

A development program/project is an attempt to solve a problem or correct a deficiency which inhibits development. In practice, the project/program description usually asserts, either explicitly or implicitly, that the project concept and design are sound - with qualifying assumptions about the anticipated behavior of exogenous factors. History suggests that this assertion is not always supported or supportable. To the extent that the nature and magnitude of the problem are not known and the forces which created and sustained the problem are not fully understood, then it is not possible to define a relevant solution or to predict its effectiveness.

Given this uncertainty, it would seem judicious to base the solution, i.e., the development project/program, on a theory of induced change which is specific to that project/program. A specific theory can set the stage for illuminating the problem within its context, defining the linkage between the problem and the proposed solution, and characterizing and explaining the proposed project/program so that it can be designed and subsequently evaluated. The theory is formulated by the project designer and validated by the evaluator. The means for validation of the theory is the evaluation. The strategy for validation is the evaluation methodology.

Since both the specific theory and the project/program design address the same development issues and both are necessarily tentative, it may be difficult to grasp the differences between them. An attempt is made in subsequent chapters to clarify those differences.

The relationship between a specific theory of induced change and the impact evaluation framework which validates the theory

Just as the project/program design must be based upon a specific theory, so must the impact evaluation framework. The formulation of an overall impact evaluation framework, and the consequent choice of evaluation models, methods and techniques to be employed within that framework, cannot occur in a vacuum. Evaluation methodology is not an end in itself, nor is it self-defining. It must comprehend the basic design elements of the project/program being evaluated and such contextual factors as exogenous and intervening variables, host country and donor values and the unique substantive characteristics of the sector.

It is important to recognize that the methodological framework for impact evaluation described in the next section is intended to validate a very tentative and incomplete statement of theory. The work to date is a beginning attempt to build theory; the theory discussed in Chapter II needs further development. It is not yet ready to be used as an authoritative standard against which to evaluate.

The relationship between the specific theory and the project evaluation

The evaluation validates the specific theory and the project design.

A Methodological Framework for the Evaluation of Development Impact

The line of reasoning which was followed in deriving the methodological framework called for the framework to fulfill two functions:

- To account for any event/influence in the training project/program cycle regardless of when, where and how it occurred, and, in doing so, to recognize that the causes of impact may be multiple and interrelated; and,
- To measure that impact.

The methodological framework is derived from the specific theory of induced change and is subsequently used for a third function: to validate the theory. The methodological framework unifies and guides the project cycle of design, implementation and post-training activities and incorporates the preconditions for evaluation. A fuller description of the methodological framework and the six functions/objectives which it comprises is found in Chapters II and III.

Analytical Tools and Measurement Techniques to be Employed within the Methodological Framework

The six tools which support and are utilized within the methodological framework are described in Chapter II.

Evaluation

The evaluation process is described in Chapter III.

In summary, a comprehensive evaluation methodology must:

- Embody an empirical/logical model of the development process;
- Incorporate the stages of the project/program cycle and hypothesize/account for any event/influence which might affect or contribute to impact, regardless of when, where or how it occurs;
- Help determine a project's evaluability and the evaluability of the critical project variables;
- Identify and describe the relationship between training and other variables affecting impact; and
- Identify and measure impact and impact preconditions, and determine attribution to, and the criticality of, the variables affecting impact.

To produce an evaluation methodology capable of doing these things, requires:

- A clearly stated training program theory and design hypotheses to be tested by the evaluation. These are the internal project/program factors;
- A clearly stated theory and hypotheses regarding the relationship between training and exogenous variables. These are the external dimensions;
- A determination of the extent to which the internal and external factors can be evaluated; and
- Analytical tools and measurement techniques to measure each factor.

CHAPTER II. IMPACT AND ITS MEASUREMENT

The purpose of this chapter is to define impact and to discuss various approaches to the measurement of impact. The chapter concludes with the presentation of the proposed framework for training impact evaluation.

The Succinct Definition of Development Impact

Development has many causes, assumes a variety of forms, takes many pathways, occurs within different geographic areas, sectors and socio-economic levels in a society, and affects different people in different ways. To define the developmental impact of an intervention is to define development itself.

Any definition of developmental impact, to be useful to planners, implementors and evaluators, must take account of this complexity and lend itself to all users in all circumstances. If the dialogue within the development community is to be coherent, the definition must also - in seeming contradiction - begin with a single, unambiguous and universal statement.

Despite the efforts of the Agency to bring clarity and structure to the definition of developmental impact, perceptions of what it is, and the formal/informal usage of the term, tend to be general, all-inclusive, inconsistent and imprecise.

This chapter posits a single, succinct definition of developmental impact and then attempts to illuminate the phenomena in all its multiple forms and occurrences with a series of ancillary statements of definition, each with an illustrative example.

Development impact is the economic, social and political change which results from an intervention and affects the quality of life for a nation or a designated subset of the population.

Development impact can be a single event or a process, i.e., a progressive unfolding of change. The evaluation framework proposed in this report is intended to measure development impact that is caused in sum or in part by training, regardless of how, where or when the impact occurs.

The succinct definition stated above and the further explication of the definition below are consistent with the Africa Bureau definition contained in Program Documentation Requirements for Missions in Africa under the DFA: Country Program Strategy Design, Annex C. page 3, April 20, 1990. This document states:

Results in DFA terminology should refer to people-level (i.e., developmental) impact, not to process indicators of actions undertaken, or intermediate indicators of results that are probably going to be achieved. Thus, increased incomes of rural populations are results (developmental impact), the number of entrepreneurial training courses held is a process indicator, and the number of tons of fertilizer sold through the private sector is an intermediate indicator. Governments issuing revised investment codes are process indicators and dollars of capital invested subsequent to that revision are intermediate indicators. Numbers of jobs added to the economy and income increases generated are results. In earlier Mission Action Plans, many of the benchmarks and targets were stated in process or intermediate indicator terms, not in the language of results (developmental impact). Even in the overall DFA Action Plan, many of the targets and benchmarks are process/intermediate indicators. But results/(developmental)impact are what the DFA is to produce.

CHAPTER II. IMPACT AND ITS MEASUREMENT

The purpose of this chapter is to define impact and to discuss various approaches to the measurement of impact. The chapter concludes with the presentation of the proposed framework for training impact evaluation.

The Succinct Definition of Development Impact

Development has many causes, assumes a variety of forms, takes many pathways, occurs within different geographic areas, sectors and socio-economic levels in a society, and affects different people in different ways. To define the developmental impact of an intervention is to define development itself.

Any definition of developmental impact, to be useful to planners, implementors and evaluators, must take account of this complexity and lend itself to all users in all circumstances. If the dialogue within the development community is to be coherent, the definition must also - in seeming contradiction - begin with a single, unambiguous and universal statement.

Despite the efforts of the Agency to bring clarity and structure to the definition of developmental impact, perceptions of what it is, and the formal/informal usage of the term, tend to be general, all-inclusive, inconsistent and imprecise.

This chapter posits a single, succinct definition of developmental impact and then attempts to illuminate the phenomena in all its multiple forms and occurrences with a series of ancillary statements of definition, each with an illustrative example.

Development impact is the economic, social and political change which results from an intervention and affects the quality of life for a nation or a designated subset of the population

Development impact can be a single event or a process, i.e., a progressive unfolding of change. The evaluation framework proposed in this report is intended to measure development impact that is caused in sum or in part by training, regardless of how, where or when the impact occurs.

The succinct definition stated above and the further explication of the definition below are consistent with the Africa Bureau definition contained in Program Documentation Requirements for Missions in Africa under the DFA: Country Program Strategy Design, Annex C. page 3, April 20, 1990. This document states:

Results in DFA terminology should refer to people-level (i.e., developmental) impact, not to process indicators of actions undertaken, or intermediate indicators of results that are probably going to be achieved. Thus, increased incomes of rural populations are results (developmental impact), the number of entrepreneurial training courses held is a process indicator, and the number of tons of fertilizer sold through the private sector is an intermediate indicator. Governments issuing revised investment codes are process indicators and dollars of capital invested subsequent to that revision are intermediate indicators. Numbers of jobs added to the economy and income increases generated are results. In earlier Mission Action Plans, many of the benchmarks and targets were stated in process or intermediate indicator terms, not in the language of results (developmental impact). Even in the overall DFA Action Plan, many of the targets and benchmarks are process/intermediate indicators. But results/(developmental)impact are what the DFA is to produce.

Changes that Constitute Development Impact

There are innumerable ways in which developmental impact occurs. This section begins with a definition of the concept of impact leverage. Following that definition, there is a brief description of three main types of developmental change.

Impact leverage

Impact leverage is defined as that quality of induced change which causes the greatest development impact with the smallest, most potent intervention/investment. Leverage is conceptually analogous to both the benefit/cost ratio and the multiplier effect (discussed below), but is qualitative as well as quantitative. Impact leverage attempts to find the most powerful, lowest cost solution to a specific development problem which is widespread. Impact leverage is the motivational force which drives development.

Three main types of developmental change

Primary/secondary - first/second generation

This type of developmental change is sequential/linear and thus does not replicate/spread to other persons/areas or multiply geometrically. A first effect becomes the cause of a second effect. The second effect does not occur until the first effect is sufficiently complete to be able to cause the second effect. This is a change in kind.

A graduate might develop and promulgate a new health treatment. This might arrest or cure a farmer's illness which would result in an increase in the farmer's productivity, with, in turn, would increase the farmer's income. Each of these sequential changes is different in kind. In this example, the sequence could branch when the farmer's income increased and cause two parallel secondary effects: an increase in farmer investment in equipment, seed and fertilizer with increased farm output; and, the possibility of the farmer's children entering school rather than working on the farm.

Replication - spread effect

This type of change can spread and/or replicate through a population or geographic area, either on a planned or spontaneous basis. There may be no sequence of further cause and effect, other than that which might emerge from the first group. This is a change in quantity. Replication/spread effect does require an increase in any resources consumed in the replication.

Replication may take the form of a process/technique or a product. A graduate might collect and disseminate information about a new approach such as a farmer's buying cooperative which makes large scale purchases of agricultural inputs (e.g., fertilizer, pesticides) at low unit cost for a large group of farmers. Another graduate might develop a low cost, processed food, enriched with minerals and vitamins for child feeding, formulated from locally grown crops. Planning and systematic communication might cause both the technique (the buying cooperative) and the product (the processed food) to spread to larger groups in wider areas with little or no change in the technique or the product. Informal or unsystematic communication might also cause spontaneous replication.

Multiplier effect

This is a sequential type of change which occurs when a change agent (cause) can produce multiple effects (usually the same effects) and each of the effects, in turn, can cause the production of further multiple effects. This is a change in quantity.

An AFGRAD/ATLAS graduate might become a teacher who trains 20 student teachers. Each student teacher, in turn, may train 20 student teachers. Other examples would be seed propogation or the reproduction of breeding stock in animals or fish.

Changes that Do Not Constitute Development Impact

Impact is defined as direct effects on the quality of peoples' lives. Any prior events, achievements, changes and effects are preconditions to developmental impact. An example is a graduate who becomes the Minister of Agriculture. Although the graduate is positioned to formulate policy, allocate resources and exercise substantial influence on the country's development, his/her position and power do not constitute developmental impact, but are important preconditions to further actions which may impact directly on peoples' lives.

Another example is the change in institutional capacity and performance which might result from the initiatives and efforts of the graduate. Institutional capacity and performance are preconditions for developmental impact, not developmental impact itself.

The methodology described in this report will differentiate between capacity and performance at all levels of the development change process. Capacity will always be a precondition for performance. Performance will also be a precondition except in those cases where performance explicitly means the delivery of goods and/or services to a target group which will enhance the quality of life.

The conventional use of the term, tracking, describes the administrative function of recording the process of participant selection, training and initial post-training employment. This is participant tracking. The methodology described in this report uses the term differently: tracking of induced change describes the systematic pursuit of evidence of developmental impact in the unfolding process of induced developmental change.

Impact Evaluation

This section describes the characteristics of impact evaluation and the expectations, constraints and limitations which are inherent in an impact evaluation methodology.

Characteristics

Developmental impact evaluation attempts to:

- Measure, or estimate, the economic, social and political change induced by an intervention;
- Determine the extent to which the change was attributable to the intervention;
- Establish the extent to which the intervention was critical to the change;
- Discover how and why the change occurred; and
- Assess the role of external factors.

Impact evaluation ultimately is concerned with effects on people: whether these effects are planned or unplanned, desirable or undesirable, transient or lasting, direct or indirect, primary or secondary, immediate or delayed, intermediate or final.

One common definition of impact evaluation is that it encompasses three tasks:

- Measures change and/or achievement of desired objectives;
- Determines the significance of the change/achievement; and
- Establishes attribution/criticality.

The measurement of change may be viewed in two ways:

- Induced change as a single event

  If we evaluate only to determine the achievement of a desired/planned objective, that may imply that the achievement is a single event and the evaluation can also be a single event. The goal attainment model is appropriate for this kind of evaluation. In this model, both the resource inputs and a rigorous design are necessary but not sufficient conditions for achieving the stated objective.

- Induced change as a continuing process

  If we evaluate to identify and measure change induced by an intervention or investment, that may imply that the change is a continuing process and the evaluation must be capable of measuring the nature, direction, rate, and other dimensions of the induced change over time.

  If the design process was not rigorous and there was not an explicitly defined objective, then the appropriate approach may be the systems model and/or the goal-free model. In the latter model, the intervention/investment is a necessary but no sufficient condition. The design of the intervention is not considered a necessary condition.

The proposed impact evaluation methodology is concerned with both of the outcomes noted above: the achievement of any specific objectives which are established at the planning stage; and, the unfolding progression of developmental changes/effects resulting from the intervention/investment, whether these are planned or unplanned, desirable or undesirable, etc. Because of this dual concern, the evaluation methodology proposed here utilizes aspects of several models.

Just as development impact is a progression of changes/effects over time, so must impact evaluation be a progression of observations and measurements over time. Just as development impact may be defined as a series of causal (if - then) hypotheses, so must impact evaluation be informed by the verification/validation of those hypotheses.

The proposed methodology calls for three basic shifts within the conventional goal attainment framework as it has been understood and practiced in AID:

- Much of the effort normally expended in summative evaluation is shifted to the design, training and early post-training stages. This will not only enhance the prospects for impact but will also make the subsequent evaluation process easier, cheaper, quicker and more productive.

- The traditional separation of roles and functions is replaced by a collaborative strategy in which the trainee and the host country stakeholders are invited and encouraged to play active roles in planning, in the collection of pre-impact and impact information and in the evaluation process as well.

- The methodology is interventionist and formative rather than passive and summative in the post-training period. Evaluation is used purposefully to identify obstacles and opportunities for achieving further development impact. The methodology calls for, accommodates and facilitates support services, networking and other post-training activities.

Expectations, constraints and limitations

Expectations

The impact evaluation methodology must take into account a structural issue, i.e., three distinct but closely linked levels of congressional, A.I.D. and host country concern about developmental impact:

- The micro level: the post-training contributions of the individual trainee;

- The country level: the development effects in the host country in both private and public sectors; and,

- The Africa level: the broader, policy-level interests of A.I.D. and the Congress.

In theory, the proposed methodology should attempt to build bridges between the three levels. The methodology does bridge the micro and country levels, using an empirical/logical model of the development process, as explained in Chapters II and III.

It is neither logically nor practically possible to build a bridge between the country and Africa levels. This is because each trainee will pursue a singular line of work in a specialized environment with its own unique contextual (independent) variables. Under these conditions, comparison and aggregation of information about the developmental consequences of the individual cases across countries are not feasible for two reasons. First, the developmental effects/changes for each case will be different both qualitatively and quantitatively. Second, each case will be the result of a unique set of causal factors.

The only possible approach which would, at least partially, accommodate the policy requirements of AID and the congress is to summarize the individual evaluation findings in small groupings with similar experiences. It is possible, for example, that evaluators might find that the efficiency of certain innovative education methods at the primary level was significantly higher than the conventional techniques, and that this finding occurred in several countries. Such a set of findings could then be replicated elsewhere if the findings had external validity, i.e., were found to be transferrable to other settings.

The proposed impact evaluation framework is intended to be universally applicable to human resource development activities. In all evaluation methodologies, and in the proposed impact evaluation framework as well, there are constraints and limitations which reflect the nature and diversity of the environment(s) in which the methodology is used. The more important limitations are:

Limitations inherent in the development process

- Inadequate empirical/theoretical understanding of the development process;

- The difficulty of defining future developmental changes/effects/goals;
- The multiplicity of exogenous factors and limited knowledge of their origin, nature and behavior;
- The time required for significant developmental effects to emerge, and the erosion of information during that time;
- The relative smallness of A.I.D.'s investment in comparison to the host country's needs;
- The diversity of the work of graduates, their work environments and the political, social and economic cultures of their countries; and
- The remoteness of the graduates' work assignments from the target group and macro level development.

Limitations imposed by the programming environment

- Changes in A.I.D. and/or host country objectives, priorities and circumstances; and
- Lack of motivation, incentives and rewards for participation in evaluation activities.

Factors associated with evaluation methods and data

- Inherent deficiencies in evaluation methods, skills and information; and
- Difficulties in achieving objectivity in observation, measurement and inference.

Evaluation Models: Options for Measuring Impact

In general, four classes of methodologies are available as options for the measurement/estimation of the development impact of training. They are: program evaluation, social science research/experimental design, economic analysis (benefit/cost, cost/effectiveness and rate of return) and the constructivist/participative (fourth generation) approach. The comparative relevance and utility of each is discussed below.

Program evaluation models and methodologies

The search for a methodological design for evaluation of developmental impact resulting from training has to confront two basic considerations. First, one must consider the differences between project-oriented training and the more general training of individuals for program/policy level objectives, and the difficulty of evaluating the consequences of the latter. Second, one must consider the interdependent relationship between the design and evaluation functions in both kinds of training.

Given the nature of scholarship programs, the first level of choice of evaluation methodologies is between two antithetical alternatives: the goal attainment and goal-free models. Other models, such as the systems model are considered useful in specific areas such as institution building. [46, 51, 54]

The goal attainment model is widely used to evaluate development assistance projects. It is a simple, straight-forward methodology which attempts to measure progress toward and/or achievement of stated goals. It is highly dependent on the project's design. The goal attainment model is well suited for project-oriented training where the participant training element is integrated with the other resource inputs such as commodities, equipment, expert advice, money, etc.; and, where training is an integral step in a hierarchy of causally-linked objectives which taken together, are readily subject to disciplined processes of design, implementation management and evaluation.

The goal attainment model usually does not employ rigorous analytical methods such as experimental or quasi-experimental design. It requires only limited skills. In practice, evaluation of the results of training, using this model, has been largely limited to such near-term effects as job placement and satisfaction, post-training income and the like.

In scholarship programs, there is often no integration with other inputs aimed at the same program level objective. Indeed, there may not be other inputs at all. There may also be no articulated hierarchical linkage between the training element and the program objective. In certain cases, there may also be no clearly articulated program objective. Because these essential design elements and preconditions for evaluation are not present, the goal attainment model cannot be used effectively without substantial modification.

In its purest form the goal-free model deliberately ignores stated goals, shuns project documentation and avoids contact with project designers and managers - all in the pursuit of maximum objectivity. It employs an outside evaluator who seeks only to identify the kinds and magnitudes of any effects which can be attributed to the investment. It then attempts to determine the merits of the effects: which were planned and which were unplanned, which were beneficial and which were not, who were the beneficiaries and who were not, which effects were cost effective and which were not, etc.

Since goal-free evaluation ignores stated goals, the argument could be made that it could be used for projects and programs with only very generalized objectives or none at all, e.g., scholarship-type programs.

Conversely, the absence of evaluation preconditions in a scholarship-type program would appear to coincide nicely with the requirements of the goal-free model which would purposefully ignore them even if they existed. Thus it could be argued that the goal-free model - by coincidence, if not by intent - could be used for scholarship programs.

Scholarship programs and training stipends may appear to fall into this goal-free category when specific results at the macro or sector levels have not been stated. This probably occurs because education and training have usually been seen as an intrinsic good, and since developing countries almost invariably lack educated and trained people, it follows that the more education/training the better.

It is easy to understand why one does not usually see an explicit statement of specific planned developmental effects at the sectoral or macro levels which are expected to result from educating individuals. To define specific desired results would require the construction, for each individual case, of a hierarchy of causally linked objectives to connect the training and the desired program level effects. This hierarchy would have to be solidly based on predictions about and/or control of a number of host country and individual circumstances, events and decisions.

There are several reasons why the goal-free model is only rarely used and why it is not recommended for human resource development programs:

- The goal-free model is perceived as discrediting, or at least discounting the importance and validity of the pre-evaluation functions of planning, design, project review and approval and implementation monitoring;
- It may alienate program and project staff whose technical competence and dedication are spurned by an outside evaluator as distractions in the pursuit of objectivity;
- It excludes project staff from the evaluation process, thus inhibiting the feedback of findings into redesign and improved execution;
- It makes more difficult the use of any evaluation techniques which require preevaluation action, e.g., the early collection of pretreatment data, the formulation of hypotheses about causal factors within or outside the design, etc.

It should be noted that there is not yet a precise, commonly agreed definition for the goal-free model. The model is still evolving and remains controversial.

Social science research/experimental design[19, 35, 91]

First, most social science research methodology is designed to identify, explain and predict the patterns and relationships which exist in populations. The analytical models (e.g., experimental or quasi-experimental design with or without random selection) and the analytical tools (e.g., correlation, regression and probability functions, hypothesis verification by quantitative/statistical means) normally applied for these purposes are usually based upon populations with one or more common characteristics.

Scholarship trainees, by and large, do not constitute a population from which patterns and relationships need to be, or can be randomly drawn. Both the purpose and the structure of a scholarship program emphasize the independent situation of each individual participant. The key variables for each participant are unique: country, sector, professional interests, technical content of the training, future employment, AID's interest in the individual participant's contribution to host country development, etc. The uniqueness of each participant's situation far outweighs any commonality within the group. There appear to be only two significant commonalities. First, all training is intended to contribute to host country development. Second, post-training networking and other support services may be available to all trainees.

The validating power of random selection is not available to a scholarship program since the program uses purposeful selection. Because of the uniqueness of the individual trainees, experimental/quasi-experimental design would not appear appropriate.

Another consideration is the requirement in social science research/experimental design to control (eliminate the influence of) independent variables, i.e., to isolate the treatment from its context. There are several problems in using the treatment/control method.

The power of the treatment/control comparison is greatest when the treatment is instantaneous (an inoculation) or short-term (a course of therapy). Where the treatment time is brief, the changes induced by environmental factors in the treated and the control are minimal. The contrast wanes as the treatment period increases since time and circumstance induce changes in both the treated and

control. In the case of long term training, the control may receive schooling from a different source or may benefit greatly from on-the-job-training/work experience while the treated was absent.

Just as time is critical, so are distance and location. If treatment and control are nearby, the influence of cultural, economic and other variables is minimal. When the trainee is in America and the untrained control is in East Africa, both will change in different ways and at different rates.

When the treatment is training and the control did not receive the same kind of training, the difference may not be significant for yet another reason. The training was given to fill a void in the host country human resource resource base. By definition, the trainee's knowledge is clearly of value to the host country. The use of a treatment/control comparison is not needed - as it is in a medical experiment -to verify the value/utility of the treatment.

Finally, while the experimental design is rigorous in the measurement of induced change, it fails to fulfill the need to identify and understand the intervening and exogenous factors which are the context for a scholarship program. A major objective of an impact evaluation methodology is to learn how, and to what extent, contextual factors influence the contributions of the graduates and, conversely, how the efforts of the graduates affect the contextual factors.

Experimental design is value free, substance free and, by virtue of the treatment - control paradigm, is free of the confounding effects of intervening and exogenous factors. When the context is purged, as it is in experimental design, internal validity is preserved, but it is preserved at the expense of external validity. External validity - determining whether, and under what conditions the experience can be transferred to another setting - is negated, i.e., it is not possible to learn how and why the change occurred.. Experimental design does not illuminate the mechanisms of dissemination, receptivity/resistance, adaptation, disruption, spread and/or multiplier effects, etc., all of which are parts of development theory and are of value to A.I.D.

The sum of these considerations is that the conventional social science research models and treatment/control techniques are not readily applicable to a scholarship program.

Economic analysis

At first glance, the benefit/cost ratio, including the social rate of return, might be seen as a valuable and rigorous addition to an impact evaluation methodology. Review and analysis of the literature in this area suggests a number of limitations and obstacles which are briefly noted here.[26, 34, 43, 92]

To be valid, both benefit/cost analysis and rate of return calculations must adhere to precise standards. First, both benefits and costs must be strictly delineated to ensure that the definitional boundaries, i.e., the lines of inclusion and exclusion, are clearly drawn and that the basis for comparability is established. Second, the defined benefits and costs must both be expressed in monetary terms. Finally, invalidating factors and impurities must be eliminated. These include such considerations as:

- Discounting and the distinction between real and nominal discount rates;
- Present value;
- Sensitivity;
- Fluctuations in valuation of monetary units;
- Opportunity costs of capital as a shadow price;

- Marginal rate of time preference;
- When inflation is present, maximizing net present value versus maximizing internal rate of return; and
- Dealing with distributional aspects of benefit/cost analysis via the social welfare function versus the social value function.

The several requirements noted here are intended to ensure that the benefit/cost ratios and rates of return are standardized. To the extent that they are standardized, it is then possible to make comparisons within and across geographic areas and economic and social sectors, specialized types of training, academic disciplines and other categories. It will also be possible to aggregate the data within and across these same categories.

It is apparent from the above that this technique imposes a substantial demand for strict definitions, quantitative information and sophisticated economic research and analytical skills. The literature indicates that most of the research in this area has centered on benefit/cost ratios and internal rates of return to the individual, i.e., the income of the graduate divided by the cost of the education or training. The research which has been done on rates of return to others such as employers, consumers, sectors and society in general, (which would be defined as developmental impact) is exploratory and, at the same time, controversial.

In situations where quantification in monetary terms for both benefit and cost is not possible, the alternative of cost/effectiveness analysis merits attention. There are three widely accepted definitions of cost/effectiveness analysis:

- Any analytic study designed to assist a decision maker identify a preferred choice among many possible alternatives;
- A comparison of alternative courses of action in terms of their costs and their effectiveness in attaining some specific objective; and
- A benefit/cost analysis without monetary valuation of program outputs.

Evaluating the impact of ATLAS trainees would normally call for the identification and measurement of economic and social effects/benefits such as:

- Decreases in morbidity;
- Access to potable water;
- Increases in wheat production; and
- Reductions in fertility.

These particular effects can readily be measured in quantitative terms (but not as easily in monetary terms), i.e., number of people affected, kilos of wheat, etc. The number could then be divided into the cost of the ATLAS training grant and thus produce a cost/effectiveness ratio.

At a superficial level, it may appear that cost/effectiveness analysis might be a useful element in the proposed impact evaluation methodology.

The dissimilarities in the two techniques are substantial. The formulation of benefit/cost ratios and rates of return poses formidable problems of definition, information collection and monetarization. It also requires a high degree of sophistication in economic research methods.

The cost/effectiveness technique is used to determine first, whether the objective is worth achieving; second, which among alternative objectives should be achieved; and third, the best way to achieve any desired objective. The technique cannot be used in those areas of qualitative change where quantification is not feasible.

The similarities of the two techniques are perhaps more important than the dissimilarities in that both are intended primarily for ex ante analysis rather than ex post evaluation. They are assessment/valuation devices, used at the planning/design stage prior to project/program approval. They are devices for comparison of alternatives, not for measuring induced developmental change. They help a decision maker to make a rational and optimal choice among alternative strategies for achieving an objective before approving funding. They are not appropriate for measuring progress toward, or achievement of that objective. Nor are they pertinent to determining how and why the objective was achieved or not achieved. They are not designed to explain the effects of exogenous variables on goal achievement.

A key development goal of the Africa Bureau is to improve the performance of African institutions and organizations to plan and promote sustainable development in Africa. The chosen strategy is to strengthen the leadership and technical abilities and enhance the professional performance of individuals serving in African public and private sector activities, including universities, research centers, and other key development institutions. The creation of a cadre of educated Africans was the objective. In the specific case of ATLAS, it was not only the chosen strategy, but was chosen as the only viable and relevant strategy. Therefore, it is a given that there is no alternative strategy against which to compare ATLAS. Thus a device for assessing alternative choices is not needed at this moment in history.

The primary challenge for the evaluator is simply to discover the kinds and magnitudes of development change resulting from ATLAS training and to determine the factors associated with that change. For this task, neither benefit/cost nor cost/effectiveness techniques are relevant.

Constructivist (fourth generation), participative and similar evaluation models

Constructivist (fourth generation evaluation) approaches are described in sharp contrast to positivist (goal achievement and social science) models. While the latter emphasize scientific rigor and objectivity, the former place a premium upon values and context. The positivist attempts to control/eliminate external influences in order to observe causal processes without distractions or confounding. The constructivist regards external influences/contextual variables as determinants and studies them in order to understand their influence on attitudes and behavior. The positivist compares the results against the original objective, measures the kind and magnitude of change and draws inferences about what happened and how. The constructivist examines the perceptions of stakeholders about the present situation in an attempt to find areas of agreement and disagreement.

The constructivist approach implies active and substantial participation by host country stakeholders, a viewpoint which supports the processes of design and evaluation envisioned in the training project cycle. The impact evaluation framework proposed in this report incorporates elements of the constructivist model in order to strengthen the mutual commitment - by the trainee, the employer, the host government planners, USAID and the implementing agents - to the government's development objectives.[4]

A.I.D.: Current Approaches to Impact Evaluation

A.I.D. has had a long term interest and a serious concern for the achievement of development impact. Over the years the Agency has made a substantial effort to formulate methodological guidance and conduct evaluations as a means for informing policy and program planning and improving operations. This investigation has examined closely prior Agency experience as well as the efforts of other bilateral and multilateral donor organizations. Of particular interest were the impact evaluation series begun by the Agency in the late 1970s; the A.I.D. publication, *An Approach to Evaluating the Impact of A.I.D. Projects;* the CDIE impact studies in Nepal, Kenya and the Philippines and the related methodological work. The research and consequent development of an empirical goal hierarchy by Albert L. Brown and Edmond C. Hutchinson under PPC sponsorship in 1975 and 1977 proved invaluable.

In virtually all cases, prior Agency methods and evaluative experience were found to serve purposes, follow definitions and cover program areas which were different from those specified in this contract task order, thus confirming the conclusions of the ATLAS design team. Nevertheless, the Agency's prior experience enriched the work which is summarized in this report.

A Theoretical Construct and a Conceptual Model for Selecting and Designing an Impact Evaluation Framework

This section develops two closely related, sequential ideas:

- A specific theory of induced development change to characterize and explain human resource development programs; and
- A conceptual model for evaluating the development impact of human resource development programs.

The theoretical construct

This section reiterates the Chapter I summary statement on theory and further develops the argumentation. It also attempts a first, tentative description of some of the features of a specific theory.

This report asserts the need to build and, over time, to validate a specific theory of induced developmental change caused by human resource development programs in general and training in particular. It further asserts that such a specific theory has two necessary functions:

- To provide the basis for planning effective and beneficial investments in human resource development, i.e., for guiding project design; and
- To provide the basis for creating a comprehensive impact evaluation framework, which, in turn, has two functions: to measure induced change and by so doing, to validate the specific theory.

These assertions may appear difficult to support since A.I.D. and its predecessor agencies, as well as other bilateral and multilateral donor organizations, have financed the training of hundreds of thousands of participants in the past half century, and have attempted to evaluate the results of the training - all seemingly in the absence of an explicitly stated, and commonly agreed theory. One possible explanation is that theories of training, explicit and/or implicit, and evaluation methodologies have not been systematically placed in concert. It is the purpose of this investigation to bring theory and methodology together.

The asserted need for a theory is based upon the following arguments.

The formulation of an overall impact evaluation strategy, and the choice of evaluation models, methods and techniques to be employed by that strategy, cannot occur in a vacuum. Evaluation methodology is not an end in itself, nor is it self-defining. It must comprehend, within some coherent framework:

- The origin, nature, magnitude and intensity of the problem;
- The resources required for its solution;
- A strategy for utilizing those resources; and
- The key contextual elements affecting the project such as exogenous variables, host country values and the unique substantive characteristics of the sector.

A development program/project is an attempt to solve a problem or correct a deficiency which inhibits development. If the nature and magnitude of the problem are not known and if the forces which created and sustained the problem are not fully understood, then it is not possible to define a relevant solution or to predict its effectiveness.

Given this uncertainty, the only rational and systematic approach is to design the development project/program within a theoretical construct which sets the stage for:

- Illuminating the problem within its context;
- Defining the linkage between the problem and the proposed solution; and
- Characterizing and explaining the proposed project/program so that it can be designed and subsequently evaluated.

The project is formulated by the designer within the framework of the specific theory. The evaluator evaluates the consequences of the project within that same framework, and in so doing, also validates the specific theory which underlies the project.

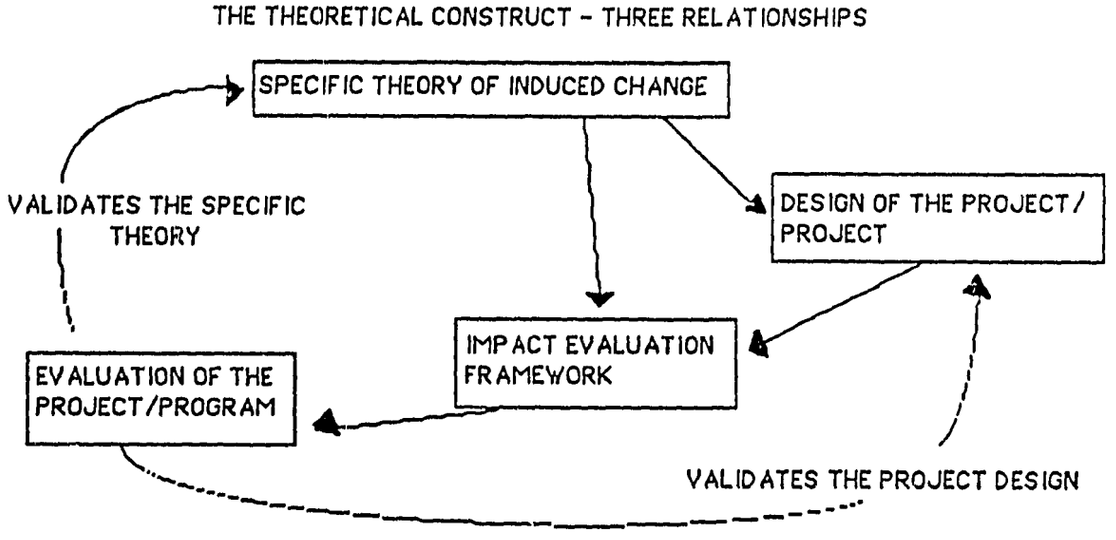
It is useful at this point to restate the classic definition of theory as it applies to the development of human resources: a set of interrelated assumptions, principles and/or propositions to guide/explain economic and social actions; a framework for formulation and testing of hypotheses.

To clarify and summarize, the specific theory of induced change has two functions. First, theory guides project/program design. Second, theory guides the development of the impact evaluation methodology.

The impact evaluation framework has one function: to provide the means for evaluators to measure induced change, i. e., to guide the evaluation.

The evaluation of the project/program has three functions. First, evaluation measures induced change. Second, evaluation validates the project design. Third, evaluation validates the specific theory of planned change.

These relations are multiple and circular. To ensure that they are clearly understood, they are illustrated in Chart 3 (repeated from Chapter 1) on the next page.



More specific reasons for the development of a theory are:

- It has the potential for increasing the capacity to predict, explain, measure and account for changes induced by an intervention/investment, a capacity which is now limited and highly dependent upon judgement;
- It can accomodate the key concept of impact leverage;
- Many of the elements which might inhere in or support a theory, already exist in A.I.D., explicitly or implicitly, and have shown evidence of improving the design, execution and evaluation of development interventions. These are:
  - models for characterizing and analyzing key segments of the society such as institution building models, sector analyses and planning models and macroeconomic models;
  - frameworks for causal modelling such as the project logical framework matrix, the goal hierarchy, decision tree diagrams, and the like;

- methods for observing contextual settings and identifying and assessing exogenous variables, such as surveys and feasibility analyses;
- approaches to identifying and integrating diverse stakeholder interests;
- means for problem identification, diagnosis and measurement, and for the establishment of baselines, benchmarks and objectives;
- methods for the formulation and testing of causal hypotheses; and
- an enormous body of substantive literature on educational investment, performance and theory, covering both empirical and research orientations.

Without a theory, human resource development will continue to be, as it has for decades, an act of faith, reflecting the maxim that education is an intrinsic good and therefore, the more the better. Without a theory, donors and host countries will continue to tinker at the margins of the existing system, e.g., to refine administrative procedures, to calibrate arrangements and to adjust budgets.[12, 13]

The arguments against the development of a theory tend to revolve around uncertainties and costs:

- What is a theory, what it will look like, how will it differ from an impact evaluation methodology?
- Why is it needed?
- What can (and cannot) it do?
- Who will use it and for what purposes?
- How difficult will it be to construct and to test?, and finally,
- How can one determine if it will make a difference, and how much of a difference?

The benefits and costs of developing a specific theory of induced change cannot be known in advance, and perhaps never. Despite these uncertainties, this report recommends that a priority and adequate resources be assigned to the development of a theory. To this end, the following passages will attempt to outline a tentative definition of how a theory might be evolved. The theory is described in terms of three elements. These are:

- The problem;
- The solution/intervention - the development goal and project purpose of the scholarship program; and
- The context, including exogenous variables, values and substantive needs and characteristics.

Each element is accompanied by a brief commentary on the extent to which it is evaluable, i.e., can be validated. Generally, these elements are relevant both to a total scholarship program and to

individual training grant projects. This section uses ATLAS to exemplify higher education as a causative intervention/investment. It is illustrative only, it does not presume to be definitive or exhaustive.

The problem

The creation of a specific theory begins with the identification of the problem. The extent to which the theory can explain the project/program and predict its effects depends, in the first instance, upon the amount and quality of information on the origins, nature, magnitude, intensity and other dimensions of the problem to be solved as well as the factors which have previously inhibited a solution.

In the illustrative case of ATLAS, the project documentation briefly notes the need for highly qualified technical and managerial personnel throughout Africa, but does not provide any specific categorization or quantification of the need. It also does not identify the nature or intensity of the constraints which historically and currently inhibit the creation of such a group. Thus, there is not yet a basis for determining the level of domestic and donor resources and effort required to solve the problem/deficiency, or for calculating at what rate and when it will be solved. If the problem is open-ended, both in terms of coherent understanding and of resources and time required, then the solution is also open-ended and the possibility of a viable theory is limited.

The solution - the development goal and project purpose

The specific theory of induced change permits a relevant and feasible solution to the specific development problem. The problem and the solution should equate within a specified time frame. There should also be an internal structure of causal linkages which is viable and sound.

In the specific case of ATLAS, the purpose (trained people) is the proposed solution to the problem (lack of trained people). The goal to which ATLAS contributes: increased capacity at the institutional level, is the proposed solution to the problem of inadequate institutional performance.

In attempting to equate the human resource problem and the training solution, two broad issues arise: First, education is not, taken alone, a sufficient condition to cause development. Second, a scholarship program is only one part of the education intervention. A brief commentary on these two issues follows:

Insufficient solution

Higher education enables, but may not be sufficient to cause development. Some subsectors/sectors are able to act as engines of development: creating consumable products, earning revenue, generating employment and driving other sectors forward. Mineral exploitation, high yield, high quality agricultural products, efficient export manufacturing and other subsectors often assume this role. Unlike these leading subsectors, higher education is intended, not to drive, but to fill deficiencies and to provide services. Scholarship programs tend to be predominantly conservative, risk-avoiding enterprises; evolutionary rather than activist or disruptive; assumed to be critically important but not necessarily a spearhead of development.

Partial solution

The scholarship program is only a piece of a piece of the total solution. First, a highly educated elite is a necessary, but not sufficient condition for political/social/economic development. Second, a scholarship program is a small part of that necessary, but not sufficient condition. The educated elite is not sufficient because other necessary conditions must also be present. These are: capital

investment and financial markets, physical infrastructure, energy sources, transport, natural resource endowment, export markets and foreign currency reserves, human resources created by other means, etc. Together, these elements constitute a necessary and sufficient set of conditions for development to occur.

If higher education is chosen as the solution, then one must look at the defined scope, character and purpose of the chosen solution, i.e., a scholarship program, to determine what evaluation should be expected to accomplish, and what kinds of evaluation are necessary and plausible. At least two levels of evaluation merit consideration:

- Effectiveness

  At the Africa-wide level it is not possible to evaluate whether a scholarship program such as ATLAS has met known, specific human resource requirements unless those requirements have been articulated in explicit and verifiable terms. At the country level, the possibility for determining the extent to which the program has met manpower requirements depends upon whether such requirements have been set and whether they are valid and realistic. These are effectiveness, rather than impact questions, but they have some importance for determining the significance of the investment relative to the deficiency, both at the overall program and country levels.

  Another dimension of effectiveness is the extent to which the graduate's work is consistent with host country development plans and priorities as well as the graduate's own professional interests and personal ambitions. The answer to this question is a precondition for the evaluation of developmental impact noted in the following paragraph and discussed at length in Chapter III.

  A specific theory of induced change might facilitate the formulation and validation of project design hypotheses about the practical relationship between human resource planning on one hand and investment in higher education among competing sectors and disciplines on the other.

- Developmental impact

  Given the situation-specific nature of scholarship grant projects, the pursuit of evidence of developmental impact must focus on the individual trainee as the unit of analysis. This fact determines the design of the impact evaluation framework. It implies tracking of the developmental effects, using a development process model and causal hypotheses. Within the specific theory, this kind of impact evaluation is viable.

  Conversely, the comparison and aggregation of information on impact within and across countries and within and across disciplines and sectors does not appear feasible, although summaries of similar individual experiences may be useful to decision-makers. A specific theory might be able to provide a basis for identifying commonalities not previously apparent. If this showed promise, then comparative inferences might be possible. This question is considered in Chapter III.

The two issues discussed in the preceding paragraphs are essentially limiting in character. They should be seen in contrast to the concept of impact leverage, which presupposes the possibility that a single, simple solution can bring about widespread and profound changes, if the solution is

specific to a pervasive problem. Impact leverage is defined in Chapter II and discussed in greater detail in Chapter III.

Context

Context includes exogenous variables, values and substantive needs and characteristics. The role of theory in this area is potentially strong, since the interaction of an intervention/investment and its environment are central to development.

Theory is a proposed explanation of induced change expressed in conceptual/abstract terms. There are different kinds of theories and differing levels of abstraction and specificity. The challenge of theory building proposed in this report includes the need to define the kind and level of abstraction most appropriate to human resource development programs.

In order to be universal in its applicability, broader, more general theory is intended to be context-free, value-free and substance-free. Social science research methodology espouses this view. A narrower theory, to be explanatory/predictive in its defined area of interest, must remain neutral and yet must be capable of explaining and predicting change which is specific to a sector, or influenced by local values and local exogenous factors.

The test of theory building posed in this report is to define a theoretical construct which has universal applicability in a defined area and at the same time is capable of guiding decision making on concrete/operational matters at the project level.

Exogenous variables

The context includes exogenous variables as well as intervening/process variables, both of which may significantly affect implementation and the achievement of project/program objectives. Training programs operate within an unlimited number of contextual variables. The impact evaluation framework must identify the most critical of these and assess their influence on the development effects of the training.

At the planning stage for conventional projects, assumptions about the behavior of exogenous variables are cited in the logical framework matrix (a specific theoretic device) and further explored in the several feasibility studies prior to approval. Causal hypotheses, formulated during the planning stage and reformulated on the basis of new information during the post-planning period, complete the preparations for subsequent evaluation. In the evaluation process, the main issue is to determine attribution and criticality, i.e., to differentiate the causal hypotheses from the exogenous and intervening factors.

Values

The identification and diagnosis of a development problem or deficiency, and the preliminary design of a solution tend to be normative (i.e., prescriptive) processes. Normative theory embodies values. It deals with aspirations and desires as well as inhibitions and anxieties, with commercial motivations as well as cultural passivity. Project/program design in the normative mode states what the designers would like to achieve: what <u>should</u> be accomplished. What should be accomplished may not necessarily coincide with what can be accomplished. Normative theory is embodied in fourth generation and participative evaluation models which place emphasis on qualitative assessment.

Normative theory, as applied to scholarship programs, determines whether the program/project which is undertaken is a true reflection of the values of the key stakeholders. In the specific case

of ATLAS, the stakeholders are the host country planners (who are presumed to represent the interests of the target group/beneficiaries), the trainee (who is the change agent), USAID, the implementing agents (AAI and the American universities) and finally, the evaluators. Given that the values of these parties are divergent, the resolution of this divergence at the earliest possible stage is crucial.

Causative theory is descriptive, rather than prescriptive. It accepts the values which were embodied in the design and attempts to determine if the design is plausible, i.e., if and how the stated objectives can be achieved. At the planning stages, it attempts to predict how the causative linkages and the processes will operate. At the evaluation stage it attempts to explain how they actually did work. It is concerned with both planned and unplanned change. It permits the exercise of judgement to interpret evaluative information.

At the evaluation stage, it is important for the evaluators to know about, and weigh the role that values played at earlier stages. It is also necessary for the evaluators to take the values of the stakeholders into consideration when interpreting the developmental effects of the training program in the host country environment.

Substantive needs and characteristics

As noted above, a specific theory should be universally applicable to any differences in substance. The specific theory must also recognize that differences in substantive characteristics have differing influences among individual projects, e.g., the substantive characteristics of agriculture as a science versus the substantive characteristics of industrial manufacturing, health, mining, nutrition, etc. When that specific theory is expressed in the form of a conceptual project design, and the conceptual model is then operationalized, substance becomes a major consideration.

Although the specific theory of induced change itself is substance-free, it recognizes that the training project will be heavily weighted toward the unique substantive content of the trainee's field of study. For example, the theory deals with the project cycle in abstract terms, defining the functions of each stage in the cycle and unifying them through the device of the goal hierarchy.

At the concrete, operational level of the training project, the theory acknowledges, but makes no provision for the fact that the cycle of problem identification, design, implementation, evaluation and feedback in a substantive/technical subsector/sector, e.g., cereal grain production, requires two kinds of expertise: the diagnostic/analytical skills used in design and evaluation and substantive knowledge of the specialized technology of cereal grains. Few people possess both skills. The combined skills of the host country, USAID, and contractors generally are adequate to the task.

Conversely, the specific theory should spell out the conditions which determine the evaluability of the project but then acknowledge that the evaluability is not significantly affected by the degree of substantive/technical content. What is challenging for the evaluator is the fact that both the project design and the substantive/technical content are independent variables and each may have its own significant effect upon the outcome. Thus the evaluator must add to the evaluation plan the issue of substantive relevance: First, was the design of the project/program relevant and appropriate to the substance (cereal grain technology); and second, was the substance more relevant to the host country problem than other alternatives, e.g., cereal grain technology versus irrigation systems versus storage and marketing versus plant disease control.

A final word on theory. Theoretical devices and conceptual thinking are widespread in A.I.D., as they are in other development organizations. But there still does not yet exist a coherent construct to guide the purposeful creation of human resources or their effective employment within the development process.

A Conceptual Model for Evaluating Development Impact

This section briefly describes the scope of the impact evaluation framework and summarizes the architecture and the constituent elements which derive directly from the theoretical construct. This is a conceptual model, not an operational model.

The scope of the impact evaluation framework

The scope of the impact evaluation framework encompasses all dimensions and levels of impact within the host country society. It delineates the defined differences between impact and the preconditions to impact. It embodies the total cycle of design, training and post-training activities, draws on, and adapts existing design and evaluation concepts and introduces new approaches. It specifies training project design and evaluative tools and techniques to be used at the several stages in the cycle. The architecture and the main elements of the methodology are briefly summarized here. Subsequent chapters describe the methodology and the design and evaluation tools in greater detail.

The architecture and key elements

The architecture of the impact evaluation framework is comprised of six major functions/objectives and six analytical tools/measurement techniques. The six major functions are:

Integration of the project cycle

A trainee in a scholarship program is a discrete development project with an inherent cycle of design, training, post-training activities, evaluation and feedback. The impact evaluation methodology facilitates the integration of the cycle, making it continuous and self-reinforcing rather than a series of discrete stages. Each of the analytical tools/measurement techniques listed below contributes to the integration of the cycle.

The integrity and explicitness of the project design

This function is to ensure that the project objectives are necessary and relevant to the problem/deficiency to be solved; are explicit and objectively verifiable; and are achievable. It also concerns the coherence and logic of the internal structure of causal linkages.

The evaluability of the project

This function calls for the incorporation of preconditions which will enhance the possibility for determining impact, attribution and criticality. These are the preconditions for the last function listed. below.

The relevance and effectiveness of the training

This function relates the training program to the needs, plans and priorities of the host country and the professional interests and personal aspirations of the trainee.

Impact on host country development

This function is the ultimate objective of the scholarship program, and its achievement is the central function of the methodology.

The identification and measurement of impact, and the determination of attribution and criticality

This is the purpose of the impact evaluation framework and the focus of its elements. It includes the concept of impact leverage.

To fulfill these six functions, the methodology draws on six analytical tools. As noted above, each of these analytical tools contributes to one or more of the functions. Conversely, each of the functions is supported by more than one element. The elements are:

Joint planning

This is a consultative process involving the trainee, the host country government, any known future employer, USAID and the implementing agents. Within this process, the project is designed and the preconditions for subsequent evaluation are set. The joint planning process requires full access to host country development plans and priorities, manpower analyses, labor market information and other relevant sources.

Trainee commitment

Trainee commitment to host country development plans and priorities is established in the joint planning process noted above. It is a necessary but not sufficient condition for impact.

Goal Hierarchies

The methodology utilizes first, a generalized goal hierarchy which reflects the anticipated host country development process and second, an individualized goal hierarchy which reflects the planned and actual contributions of the trainee to host country development. The two goal hierarchies systematically link host country needs to the trainee's career pathway. The devices are intended to function at the level of the individual trainee and for ATLAS as a total program.

Causal hypotheses within the structure of the goal hierarchies

The methodology specifies the formulation and reformulation of causal hypotheses at the project design stage to disaggregate and delineate the intended process of developmental change. Two independent but interrelated sets of causal hypotheses are specified: (a) generic hypotheses linking the trainee/training to development impact and (b) situation-specific hypotheses which track the process of induced developmental change.

The hypotheses are modified and fortified to include simple networking of the change process and consideration of both intervening and exogenous variables. The methodology uses hypothesis verification as an integral follow-on at the evaluation stage.

Observation, intervention and information collection, simultaneous with the occurrence of induced change

This process begins when the ATLAS trainee is selected and continues into his/her career. It affords the opportunity to enhance the probability of achieving impact and of identifying and measuring impact.

Joint participation in the evaluation process

This is the counterpart of the joint-planning process noted above and would involve the same technique and the same stakeholders.

N. B. *The relationship between the methodological framework and the analytical tools/ measurement techniques (above) requires a word of clarification. Each of the six functions/objectives of the framework is supported by more than one of the six tools; in some cases, all six of the tools serve to support the same function. Since the linkages between functions and tools are each multiple sets, it is not possible to portray the relationships clearly in graphic form.*

CHAPTER III. PROPOSED APPROACH TO THE EVALUATION OF DEVELOPMENT IMPACT - WHERE THE CAUSE OF CHANGE IS TRAINING

This chapter (a) introduces the proposed impact evaluation framework, (b) explains its coverage of the several stages of the training project cycle, (c) characterizes the dimensions within which impact may occur and (d) defines and describes the conditions and the analytical methods and devices which affect the use of the framework.

The methodology proposed here is an integration of new ideas and proven techniques. It draws from the goal attainment model in its emphasis upon measuring achievement of stated objectives. It specifies a goal hierarchy device - analogous to the project logical framework matrix - to guide the design and evaluation of training projects and integrate the project cycle. It abstracts from the systems model for the design and evaluation of institutional capacity and performance. It utilizes a major theme of the fourth generation evaluation paradigm in defining a joint collaboration among stakeholders throughout the project cycle. It may even be seen as resembling one feature of the goal free model in that it places no constraints on the range of possibilities for developmental impact, although in practical terms, the specialized education of many graduates may self-define a circumscribed universe in which developmental impact can be expected to occur. Finally, the methodology utilizes a modified hypothesis formulation - reformulation - verification technique derived from Aristotle's hypothetical syllogism, as the basis for design and evaluation.

The impact evaluation framework described here is not limited to a prescription of what training results to evaluate and how to evaluate them. It goes beyond that passive level by deliberately introducing into the design of training programs those elements which not only permit impact evaluation, but more importantly, will increase the explicitness and relevance of post-training goals and will enhance their compatibility with the host country's development plans and operations in both private and public sectors.

Perhaps the most important idea which can be made explicit within the framework is that there is a wide range of possibilities for greater or lesser impact leverage: that the selection, training and employment of the individual can have a profound effect on the nature and magnitude of the resultant development impact.

In addition to the ideas summarized above, it should be noted that the formulation/reformulation/ verification of causal hypotheses, within the framework of the goal hierarchy device, integrates the training project cycle. Hypothesis formulation and reformulation would occur on a continuing basis during the design and implementation of training and the early post-training stages. Hypothesis verification is central to the evaluation and feedback stages. Since causal hypotheses are critical at all stages in the cycle, this chapter begins with a summary display of the necessary developmental conditions from which causal hypotheses are derived (see Chart 4).

CONDITIONS AFFECTING DEVELOPMENT IMPACT OF ATLAS TRAINEE-BASIS FOR CAUSAL HYPOTHESES

	THE GOAL HIERARCHY	NECESSARY BUT NOT SUFFICIENT CONDITION FOR DEVELOPMENT IMPACT			EXOGENOUS VARIABLES SUSTAINABILITY	INDICATORS OF: •preconditions for development impact •development impact
		PRE-DEPARTURE PLANNING AND DESIGN	IMPLEMENTATION OF TRAINING PROGRAM	POST TRAINING ACTIVITIES		
AGGREGATIONS OF DEVELOPMENT IMPACT	AFRICA	cross national planning and coordination	-broad based commitment -advocacy -beneficiary evidence	-commitment of resources -mechanism for coordination	•POLITICAL VIABILITY AND STABILITY • SOCIAL VIABILITY AND STABILITY •DEMOCRATIC PRACTICES •PROFESSIONAL MOBILITY •ADEQUACY AND ACCESSIBILITY OF RESOURCES, INFRASTRUCTURE AND SERVICES •VIABILITY OF INTELLECTUAL SERVICES, EG., RESEARCH AND DEVELOPMENT, EDUCATION AND TRAINING, ORGANIZATIONAL DEVELOPMENT	-income/comsumption/savings -employment -investament -improved economic social, political -improved productivity/production -improved life quality
	COUNTRY	awareness, support and participation of national leaders in ATLAS -interest and support of key planning and operating agencies -viable development plan	-broad based commitment -advocacy -beneficiary evidence	-commitment of resources -appropriate policy choices		
DEVELOPMENT IMPACT	TARGET GROUPS-BENEFICIARIES	-representation at planning stage -target group needs, capacities and performance known to trainee	-monitoring and support to trainee reflects unique target group requirement	-work assignment appropriate to unique target group requirements		-income -employment/productivity -advancement,consumption -access to economic/social services -life quality improvements
PRECONDITIONS FOR DEVELOPMENT IMPACT	SECTOR/SUBSECTOR SYSTEM	-participation of government, business, industry and academia in planning -access to sector planning -donor coordination -adequacy of needed sector resources and services	capacity for managing flow of resources and actions among institutions and donors	appropriate assignment within sector/subsector coordination within sector and among sectors adequacy of sectoral resources and info.		-sectoral/subsectoral needs for trainee knowledge -sectoral/subsectoral capacity -sector/subsectoral performance
	ORGANIZATION/ INSTITUTION	-participation of top management in planning -strategic paln for use of ATLAS training -plan for org. development	monitors and supports trainee monitors and trainee research/field work trains work units and prepares for reentry of trainee	-ensures appropriate work assignment for trainee -provides support services -provides on-job training -provides adequate income and benefits		-demonstrated capacity of institution to utilize trainee knowledge -improvements in institutional productivity and production -improvements in institutional outreach
	INDIVIDUAL ATLAS TRAINEE	trainee actively participates in planning career trainee has mentor host country, employer, AAI, USAID participate in planning access to development plans, priorities, human resource plans, labor market analysis	faculty guidance contact with host country mentor research/fieldwork relevant to career plan and to host country dev plans and priorities	-employment in chosen field -involvement in prof. networks -practical training/ updating/teaching		-personal income -career advancement -contributions to development

The impact evaluation framework includes the following components. First, it encompasses all project/program phases including planning, implementation and post-training. Second, it considers contextual factors which have the power to affect change. These exogenous factors are in addition to training, which is the primary causal factor under investigation. The contextual factors are political viability and stability, social viability and stability, adequacy and accessibility of resources, infrastructure and services, viability of intellectual services, e.g., research and development, education, organizational development, and others. Third, it has the ability to evaluate impact in one or more dimensions of change including individual, organization, sector, beneficiary group, country or multinational. Fourth, it has the capacity to evaluate the impact of training regardless of the ways A.I.D. Missions use or do not use training (e.g. ATLAS) as a component of their country assistance programs. Finally, it defines the impact indicators which are acceptable measures of impact.

In the discussion which follows, each component of the impact evaluation framework is reviewed: stages, contextual factors, dimensions of change and impact indicators.

Impact Evaluation and the Stages of a Training Program

The conceptual framework for impact evaluation encompasses the design, implementation and post-training stages of the cycle, draws on, and adapts existing design and evaluation concepts and introduces new approaches. It specifies design and evaluative tools and techniques to be used at the several stages in the cycle.

The planning/design stage

This section describes the several elements which comprise the design stage and set the stage for subsequent evaluation. The concept which unites these elements is the need to translate the salient features of the host country's development needs, plans and priorities into a form which the trainee can use in setting his/her own career objectives and in formulating a career pathway to meet those objectives. The underlying hypothesis is that the more the trainee learns about his/her country's development needs, the greater will be the trainee's contribution.

The first element is a generalized goal hierarchy: a simple model of the major stages of the development process into which can be placed information on the resource requirements, objectives, priorities and strategies of the host country. The goal hierarchy permits the formulation of causal hypotheses which explain how the process of development might be expected to occur. The hierarchy, informed by explicit host country planning information, makes it easier for the trainee to plan a post-training career which links his/her personal and professional interests to the needs of the host country.

The second element is an individualized goal hierarchy/career pathway which the trainee derives from the generalized goal hierarchy after he/she is fully aware of key information about the host country's needs, objectives and priorities. The individualized goal hierarchy/career pathway is a step-by-step charting of the personal aspirations, interests and objectives of the trainee. The formulation of causal hypotheses, initiated in the generalized goal hierarchy, is explicated, i.e., sharpened and made more specific to the post-training employment of the alumnus. The preparation of the two goal hierarchies, and the formulation of causal hypotheses sets the stage for the subsequent evaluation of developmental impact.

The third element is the formulation, reformulation and verification of causal hypotheses within the framework of the generalized and individualized goal hierarchies. It is a simple, qualitative version

of the technique. It is inductive rather then deductive. It eschews the accoutrements of the statistical version of hypothesis testing in that it does not presume to measure confidence levels, rejection regions, significance probabilities and the like.

The fourth element is the process of joint consultation and planning which results in the preparation of the individualized goal hierarchy and the causal hypotheses within that hierarchy. The process is centered around the participant and supported by the host country, any known, future employers, USAID and the implementing agents.

The four elements of the planning/design stage facilitate the full expression of A.I.D. policy interests and priorities, including special policy emphases such as Women in Development, environmental protection, etc.

These four elements are described below.

The generalized goal hierarchy

The generalized goal hierarchy is based upon three sources:

- Host country macro and sector development plans;
- The designs of development projects and programs which directly support host country development plans; and
- Evaluations which illuminate any differences between planned and actual development results.

The generalized goal hierarchy identifies and displays the progression of developmental changes and effects triggered by a development intervention/investment. Ideally, the generalized goal hierarchy is a valid reflection of the operational realities of the developing countries; universally applicable to a variety of development investments, regardless of sector, geographic location, etc.; and, useful in attempting to predict as well as evaluate the outcome of new developrnent assistance initiatives.

The generalized goal hierarchy is an empirical logical model of induced developmental change. Its only purpose at the planning stage is to be a bridge between the macro and sectoral information contained in the host country development plans and the micro aspirations and plans of the individual trainee.

The generalized goal hierarchy is a diagnostic/planning device which describes *how* development might occur at the institutional, sectoral, target group and macro levels. Its purpose is to illuminate the possible dimensions, levels, and pathways for developmental impact. Although it is intended to be universally applicable to all developing countries, to be plausible, and therefore most useful, it must be informed by/consonant with the kind, direction and magnitude of development plans and realities of each country.

Moving from host country development plans and priorities to personal career planning is an enormous leap. It requires interpretation and adaptation which may be beyond the capacity of most trainees. The generalized goal hierarchy is proposed as an intermediate device to help bridge the gap.

The host country development plan, and adjuncts such as manpower inventories and labor market analyses, identify *what* human and other resources are needed for development. They also contain the information which is needed to formulate an individualized goal hierarchy/career pathway.

As changes occur over time during the ATLAS training cycle, any significant changes in host country circumstances and priorities should be reflected in the generalized goal hierarchy and consequently in the individualized goal hierarchy.

A more detailed explanation of the structure and functions of the generalized goal hierarchy follows.

AID development assistance incorporates a general logic system in which progress proceeds from stage to stage in an ordered and sequential manner, i.e., from input to output to purpose and beyond purpose to developmental effects in organizations/institutions to subsectoral and sectoral systems to target groups and to broader national societal benefits. The sequence is intended to be causal, although not necessarily linear, with attention given to the influence of intervening/process variables and to exogenous variables.

At each stage in the hierarchy at least two general types of development change can be observed. These are antecedent changes in activities, characteristics and capacity and consequent outcomes, performance and benefits. The stages can be classified and described in operationally relevant terms as a basis for planning and design, for incorporating pre-evaluative elements at the planning stage and for guiding the evaluation process.

The causal relationships within the hierarchy described here are demonstrable either empirically or logically. A simplified version of the generalized goal hierarchy looks like this:

Levels in the Goal Hierarchy	Developmental Change/Impact
National/macro	Societal/national group benefits
Target group	Target group benefits Target group performance Target group capacity
Subsector/sector system	Subsector/sector system performance Subsector/sector system capacity
Organization/institution	Institutional performance Institutional capacity

Following are definitions of the goal hierarchy levels and developmental changes/effects.

- Organizational/Institutional Level

  An institution is a significant organization which is a part of a subsector/sector system. It may be a ministry or a primary sub-unit, an autonomous agency, a private firm or other entity. Depending upon how the sector being assisted is defined, it may also be a discrete subsystem of the sector system, e.g., credit, marketing, small-scale manufacturing, a consumer cooperative, cereal grain research.

Institutional capacity is the potential of an organization. Institutional performance is the creation and delivery of policies, services and/or products by the organization.

A model of organization/institution building is described in Chapter III. The model is an integral part of the generalized goal hierarchy.

- Subsector/Sector System Level

  A sector system is that functional segment of an economy - composed of facilities, activities, institutions and relationships - which directly supports a development goal at the national/macro level. A sector system is the combination of, and the interrelationship among organizations, practices, channels and policies which moderate sector performance. Some development projects may leapfrog directly from institutional performance to the target group beneficiaries. However target group activities (and the success of the institutional products) are normally influenced by the sector system.

  Sector system capacity includes changes in the number, type, volume or quality of system activities brought about by the training project. System performance includes policies, services and products emanating from multiple institutions operating as a single sector system. Sector performance includes both the effects of the training project and non-project influences within the sector system which together are necessary and sufficient to achieve a change in target group behavior.

- Target Group Level

  A target group is an identifiable class of people which the project (the graduate) is expected to influence/affect in a predictable way. This may be all people of similar characteristics or some more restricted subgroup defined by location, occupation, sex, income or other distinguishing characteristics. Target group capacity is the behavior, knowledge, attitudes or social organization which the project is expected to alter. Target group performance is the proximate result of the effective application of target group capacity, and usually takes the form of increased productivity and production. Target group benefits are the desired result of target group performance, e.g., farmers change their cultivation practices (capacity) to increase yield per hectare (performance) to receive a higher income (benefit).

  In the preceding example, the causal chain was linear and the benefit accrued to the farmer only. Development is often non-linear, has multiple effects and requires a broader benefit and benefit incidence: the capacity of one target group should result in the production of policies, services and/or goods which benefit other target groups and/or contribute to a different societal/target group goal. Using the same example, farmers (target group 1) may change their cultivation practices to increase yield per hectare to receive a higher income. The change in cultivation practices may increase farm employment (target group 2) and increase the quality and quantity of farm produce available to urban consumers at lower prices (target group 3).

- National/Macro Level

  National level goals represent those benefits desired for broad national groups and the society as a whole. These are the goals most commonly stated in national development plans or articulated by national leaders. National group benefits are changes in the characteristics of broad groups or systems to which the target group belongs, but which transcend the target group. This impact class defines and gives content to national goals in terms of the benefits to be conferred on particular classes of citizens. Societal benefits represent the national aspiration for economic growth, improved social relationships, general well-being, participation in the international order and national policy. They thus represent the goals from which lower order goals should be derived and to which the efforts of the graduate should ultimately be directed.

  Another way of understanding the goal hierarchy is to see it as a series of enabling mechanisms, cascading upward, and at each level, adding a new dimension of capacity and performance from the preceding level:

  - At the level of the graduate, the enabling mechanism includes the technical, behavioral and attitudinal capabilities which were acquired during training.
  - At the institutional level, the enabling mechanism includes the enabling mechanism of the graduate (above), as well as the enabling mechanism of the institution, e.g., its capacity, resources, doctrine, program, equipment, budget, leadership, staff, etc.
  - At the sectoral level, the enabling mechanism includes the enabling mechanism of the graduate and the institution (above) as well as the enabling mechanism of the sector/subsector, e.g., the sectoral capacity to absorb, utilize and multiply its activities, resources and relationships and to deliver goods and services to the target group (and to the macro level).
  - At the target group level, the enabling mechanism includes, all the prior enabling mechanisms (above), as well as the capacity of the target group to absorb, use and benefit from the accumulated goods and services.
  - At the national/macro level, the enabling mechanism includes all the prior enabling mechanisms (above) as well as the capacity of the society to foster policies and programs which provide for capital formation, equity in taxes and income, social welfare, export marketing, currency stability, etc.

- There is a third perspective in which one can view the generalized goal hierarchy. The uses of the goal hierarchy are bidirectional.

  From the bottom up, it is a formative, planning framework for the causal hypotheses which characterize the development continuum; the explicit assumptions about the behavior of exogenous and intervening variables; the indicators; and, the kinds and sources of information needed to verify/validate the hypotheses. From the top down, the hierarchy is a summative evaluation framework within which the causal hypotheses are verified/validated and attribution and criticality are pursued.

Logically and conceptually the goal hierarchy is linked to project level objectives. The goal hierarchy attaches schematically to the project logical framework matrix at the purpose level, with the attainment of project purpose triggering the causal chain represented in the goal hierarchy.

This is illustrated in the specific case of ATLAS as follows:

The purpose statement for ATLAS is: To strengthen leadership and technical abilities and enhance professional performance of individuals serving in African public and private sector entities, including universities, research centers, and other key development institutions.

The goal statement for ATLAS is: To improve the performance of African institutions and organizations to plan and promote sustainable development in Africa.

The first level in the generalized goal hierarchy coincides with the goal statement of the ATLAS project logical framework and represents the first step in the unfolding progression of developmental effects which will emerge from ATLAS.

At this preliminary stage, the generalized goal hierarchy is conceptual, not operational. It is not yet a device for maximizing or optimizing, although with the incorporation of magnitudes, values, time and other specific information, it might begin to take on such a characteristic.

The individualized goal hierarchy/career pathway

The individualized goal hierarchy is a planning device which describes how the trainee's personal contribution to host country development and personal career achievement might occur. It illuminates the possible dimensions, levels, and pathways for such personal achievements and contributions.

This is an individual career plan, adapted from the generalized development goal hierarchy, in which the specific qualifications and interests of the individual are integrated or harmonized with the needs, plans and circumstances of the host country. It is to be jointly drafted in close collaboration between the trainee, the relevant government ministries, any future employer, USAID and the implementing agents.

The individualized goal hierarchy is intended to integrate the design and evaluation functions of the training project. Specifically, it would build into the design stage the necessary preconditions for the subsequent evaluation of post-training impact. The integrity and utility of the individualized goal hierarchy as a predictor of developmental impact at the planning stage would obtain, in the first instance, from the fact that it was derived from the empirically based generalized goal hierarch.

Its value would also depend importantly upon the nature of the pre and post-training collaboration between the trainee, the relevant host country ministries and any future employers. Prior to this collaboration, the trainee may have little or no knowledge of the host country's macro or sector development plans, its human resource inventory, the current and projected state of the labor market, or its experience with prior development initiatives.

The building blocks of both the generalized and individualized goal hierarchies are causal hypotheses, formulated in a tentative fashion by the graduate to predict the sequence of initiatives,

activities, events and achievements which he/she planned for the post-training employment. These causal hypotheses, provisionally stated at the planning stage, will be reformulated during the training and post-training stages as the trainee learns more about his/her field of study and about changing host country circumstances. The individualized goal hierarchy will contain three elements necessary for evaluation:

- Baseline data describing the participant's pre-training qualifications and the host country circumstances in his/her area of interest;
- The participant's career objectives and their relation to host country development objectives; and,
- The causal hypotheses which will illuminate the developmental change process.

The two goal hierarchies are not formulae to be followed nor forms to be completed; rather they are means for thinking through the process of induced developmental change and how that process might be influenced by the graduate. The goal hierarchies are not ends in themselves, but are only one of several means of increasing the probability that the graduate's contribution will be relevant and substantial. The goal hierarchy device is a necessary but not sufficient means for forecasting development changes/effects.

The goal hierarchy device may look intimidating and tedious. It is not so intended. It should be seen for what it is: the framework for career planning.

The nature of causal hypotheses and their formulation/reformulation at the design, implementation and early post-training employment stages are described in greater detail below. The subsequent verification of causal hypotheses in evaluation is treated later in this chapter.

The formulation/reformulation of causal hypotheses

Hypothesis formulation, reformulation and verification is not a discrete evaluation methodology but is an analytical device inherent in all evaluation methodologies. It attempts to predict or explain some phenomena. It describes the conditions under which the phenomena is expected to occur.

The formulation, reformulation and verification of causal hypotheses will be applied within a framework. The framework is the sequence of developmental change displayed in the generalized goal hierarchy and its derivative, the individualized goal hierarchy. As noted above, these goal hierarchies will be informed, respectively, by country-specific data on national/sectoral objectives and priorities and by trainee-specific data on personal objectives and priorities.

The basic building block is a simple conditional statement of presumed cause and effect, i.e., an if - then statement. A series of these conditional statements are formulated in an attempt to disaggregate and trace the continuum of change from one major level of development to the next, e.g., from institutional capacity to institutional performance to sectoral capacity to sectoral performance. Causal hypotheses at the planning stage are called predictive hypotheses. At the evaluation stage, they are explanatory hypotheses.

To enhance the power of the technique, each conditional statement is supported by:

- Independent sets of indicators of cause and effect;

- Notations of what data is needed to support the indicators; and,
- Assumptions about the relevance and behavior of intervening (process) variables and exogenous factors.

This technique, which will be described in greater detail in subsequent paragraphs, is consistent with the approach specified in the Bureau guideline, Program Documentation Requirements for Missions in Africa under the DFA: Country Program Strategy Design, April 20, 1990, page 8, which states:

Here it will be important to: (a) identify the intermediate steps that lie between project/non-project outputs, program targets, and the strategic objectives of the program, (b) clearly and convincingly draw the analytical links between them, and (c) put in place systems which will measure that interim progress.

It is also consistent with the approach proposed in AID Program Design and Evaluation Methodology Report No. 5 of March 1986, Sections 4.3.2 and 4.3.3, pp. 22 - 25, which discusses how to identify causal chains and how to develop working hypotheses.[46]

Causality, attribution and criticality are dimensions of the same phenomena. They are interrelated and form the basis for the conditional statement/causal hypothesis. They are briefly noted here. Following these notes is a description of the hypothesis formulation/reformulation technique.

- Causality

  The concept of causality rests on the basic premise that each level of developmental change in the goal hierarchy can be shown to be necessary, albeit not sufficient to the achievement of the next level. Since each causal linkage is subject to external factors which may be unpredictable and beyond control, the basic model of causality must include assumptions about the desired behavior of those external factors. Necessary and sufficient conditions for causality are present when the first level of necessary developmental change exists and when the assumptions about the desired behavior of the external factors are valid. Thus the causal model is built upon two parallel hypotheses: one specifying the intended causal linkage and one specifying the behavior of exogenous variables.

  Each hypothesized causal link can be seen as a simple input-output relationship which requires verification and a search for alternative explanatory hypotheses.

  Causality cannot be proven on the basis of logic. The hypothesis verification technique is therefore limited. It is capable of presenting persuasive evidence of causality only to the extent that the supporting data is comprehensive and reliable.

- Attribution

  The concept of attribution identifies the cause of an observed effect. It is the converse of causality. It moves in the opposite direction from causality. Like causality, it cannot be proven. Causality and attribution do not constitute a dilemma or a zero-sum game, nor are they mutually exclusive. Their relationship can be rigorously treated by the testing techniques described below. In addition to these verification techniques, the relationship between causality and attribution can be illuminated if comprehensive and reliable data can be found.

- Criticality

  The concept of criticality determines whether the effect could have been achieved by any means other then the hypothesized cause. Criticality is a correlate of attribution. It is the converse of causality. It moves in the opposite direction from causality. Like causality, it cannot be proven. Causality and criticality do not constitute a dilemma or a zero-sum game, nor are they mutually exclusive. The techniques for examining their relationship, cited below, are the same as those for attribution.

The search for attribution and for criticality has two dimensions:

- To determine that ATLAS training is the direct/primary cause of the observed effect; and

- To identify and eliminate other possible causal agents.

The techniques which are described in later paragraphs are useful in pursuing these two dimensions.

The causal hypothesis technique and the dimensions which characterize it, have to be seen in a practical perspective: the ability to impute causality and/or association between a training program and induced developmental change in an underdeveloped society.

The relationship between independent/causal variables (e.g., training) and dependent/resultant variables (e.g., developmental change) is a gradient. This gradient ranges from the unattainable extreme of pure causality to the other easily attained extreme of zero correlation. Between these extremes is a gray area where the independent and dependent variables are linked by various combinations and degrees of causality and association. Although causality cannot be proven and measured, association can - if the data are quantitative and/or objectively verifiable.

The only non-statistical way to deal with the gradient is hypothesis formulation, reformulation and verification. There is a diagram (Chart 5) of a sequence of causal hypotheses on the following page. The diagram shows the influence of intervening and exogenous variables at each causal hypothesis as well as the use of independent sets of indicators for separately verifying cause and effect.

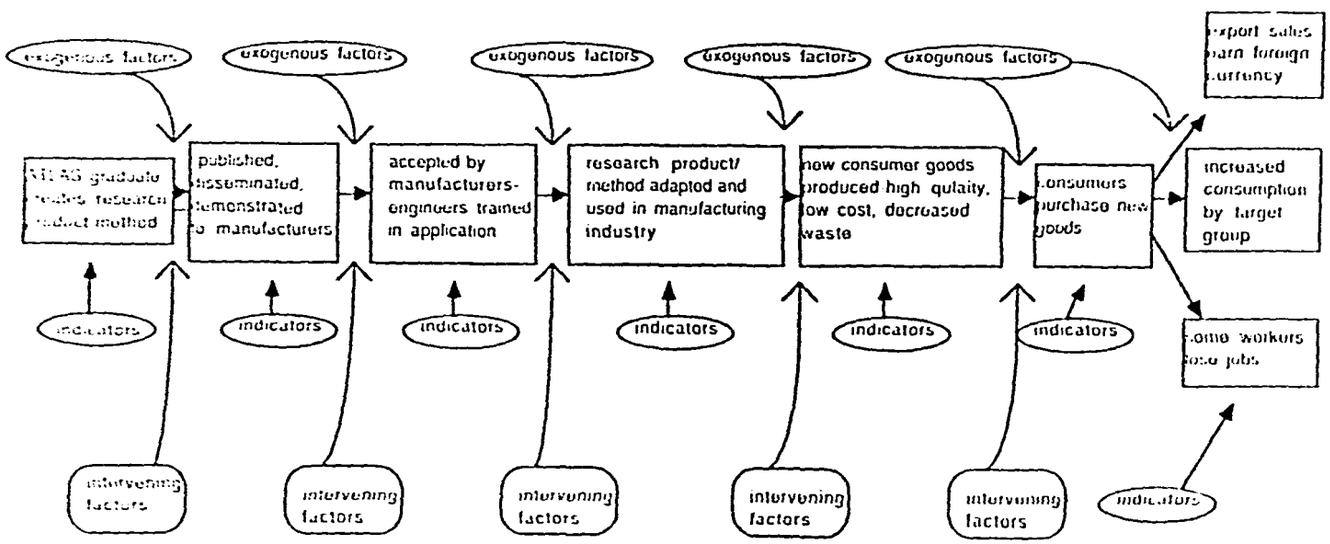
It is important to note that there is no such thing as information which predicts. Predictions are made using descriptive information about variables which experience has shown to be highly correlated, i.e., A and B have tended in the past to be associated, A is proportional to B, A has tended to precede B. When a causal hypothesis is first tentatively formulated, during the joint planning process prior to training, it should have several qualities:

- It should describe a short span of change and time, i.e., a foreseeable and achievable change which is credible;

- It should be stated in terms which are as precise, explicit, finite and objectively verifiable as practicable; and

- It should attempt to describe how and why the change will occur, i.e., the process by which the causal agent will operate to bring about the planned effect.

MODEL OF HYPOTHESIS FORMULATION-REFORMULATION-VERIFICATION



Causal hypotheses are usually stated in "if - then" terms, e.g., if a new high yield, disease resistant strain of wheat is available (cause), then farmers will adopt it (effect). This causal hypothesis has a relatively short span of developmental change.

A second example, with a broader span might be: if a new high yield, disease resistant strain of wheat is available (cause), then the export earnings of the agricultural sector will increase (effect). In this example, the hypothesis overreaches, i.e., it is much harder to validate for the obvious reason that its wide span first, opens the door to multiple uses/effects of the wheat; second, brings into play many more intervening, external variables; and third, greatly reduces the possibilities of direct attribution of the effect to the cause.

It follows from these examples that the impact evaluation methodology emphasizes the shortest practicable spans in order to increase the accuracy and power of the hypothesis verification process.

It should be underscored that the reformulation of causal hypotheses on the basis of new information and experience (i.e., the revision of project design) is not in conflict with the basic requirement for rigorous and careful planning at the outset. It may seem wasteful to devote a great amount of effort in formulating the individualized goal hierarchy/career pathway at the planning stage, only to reformulate it later. It might also seem that reformulation means abandoning the original, carefully defined objectives. Neither of these perceptions is true.

There are two reasons why rigor and discipline are needed both at the planning and subsequent stages. First, the scarcity of development resources, both in the host country and AID, requires that they be carefully husbanded. This can best be done by disciplined planning with continuing adjustment to host country realities over time. Second, we have only limited knowledge and capacity to diagnose the present and predict the future. There is no proven development theory upon which to base the design of individual ATLAS training grants. This means drawing on the only body of authoritative knowledge available: empirical knowledge based upon systematic observation. Given these realities, each training grant must be seen as existing in a formative mode in which knowledge about both the problem and the solution progressively increases. Careful design and reformulation are not mutually exclusive, rather they are mutually reinforcing.

During the training period, and in the early stages of post-training employment, as the trainee's knowledge and awareness increases, each hypothesis should be periodically reformulated to bring it into closer conjunction with host country circumstances.

The role of the trainee in the training project cycle.

The development of an impact methodology for ATLAS must take account of an operating policy issue: to what extent should the host country (and/or AID) control/influence the trainee's training program and post-training employment - to what extent should the trainee's freedom of choice prevail?

There may be an inherent tension between the host country's pursuit of its development goals/priorities and the trainee's freedom to choose his/her field of study and post-training employment. This tension may be present at the planning stage, during training and/or in the post-training period. The two sides of the argument look like this:

- One side of the argument

  Since the host country and AID are both investing scarce resources in the training and since the explicit purpose of the training is the achievement of those host country development goals which AID is supporting, it is in the host country's and AID's interests to try to exert some control/influence over the selection of the trainee, the content of the training and the post-training employment. This is not merely a theoretical argument. Both AID and host countries have exercised explicit and implicit controls and attempted to impose obligations on trainees for many years in selection criteria and practices, in predeparture processes and in the form of visa requirements, post-training covenants and agreements, etc.

  At the operational level, the AID policy determination on participant training states that "All feasible steps should be taken to ensure that AID sponsored trainees return to work in their home countries and in positions where their training is utilized effectively. The timely return of trainees and their continued employment in fields relevant to development will be major criteria for evaluations of training programs."

  At the ethical level, trainees are being given a free MS and/or PhD and an enhanced lifetime earning capacity worth many times that amount. It is not unreasonable to expect some sense of trainee commitment to contribute to his/her country's development effort.

- The other side of the argument

  The converse argument is that in a free society, the individual has a right to choose whatever training and post-training employment he/she wishes without coercion. If the trainee feels that the government or A.I.D. is encroaching upon his/her freedom of choice then he/she can reject the training offer. At a more practical level, the trainee can accept the training and then go his/her own way after completion of the training, on the assumption that the host government is unwilling or unable to impose its will once the trainee has entered the labor market.

  It could be argued that there is a risk and a potential cost both to the host country and A.I.D. in relying too heavily on the laissez faire approach. For example, would the two sponsors consider it a success if a returned trainee established a thriving, profitable, employment-creating enterprise for the manufacture and marketing of a trivial product (e.g., jewelry or pinball machines) when the host country development plan assigns priority in the manufacturing sector to scarce and essential products such as irrigation equipment, food preservation and processing facilities, etc.

The viability of the proposed ATLAS impact evaluation methodology is highly dependent upon the trainee's active engagement in all stages of the design, implementation, evaluation and feedback cycle. Exposing the candidate to planning and evaluative information which is available only from host country government sources could have a powerful and beneficial effect on his/her view of the future and of the contribution that he/she could make to host country development either in the private or public sectors. Even at a minimum, pretraining consultation and joint planning could identify and help to reconcile divergent interests.

From the standpoint of impact evaluation, perhaps the most critical element will be the requirement to anticipate and observe the developmental change process as it occurs and to collect and analyze information as it becomes available, rather than years later in a conventional, one-time, ex post impact evaluation. Here the role of the trainee is critical.

The power of the hypothesis verification technique, and the difficulty/ease of using it are highly dependent upon the extent to which the observer can:

- Predict and anticipate the induced change;
- Directly and simultaneously observe how, where and when it occurs, and record the observation; and
- Directly and simultaneously observe the circumstances and context in which the change occurs, identify and assess the external factors which may have affected the change, and draw inferences about the nature and magnitude of the changes, and about the causal and associated factors.

The importance of the time dimension is displayed here:

Observation/information collection simultaneous with occurrence of developmental change	One-time impact evaluation x years after trainee returns home
anticipate next change	not relevant
accurate information	loss of memory, loss of data
all sources of information are in place and available	some sources lost through attrition
can observe exogenous/intervening variables	observers largely unavailable
can observe replication/spread effect, multiplier effect, first and second generation effects	difficult to identify, measure, attribute
knowledge of substance and values	difficult to reconstruct

Each of the above three criteria call for the observer to be intimately familiar with/engaged in the change process and, at the same time, be capable of objectivity toward the process. Only one person can meet these three criteria.

Traditional approaches to evaluation make a sharp distinction between the participant, who may be deeply committed, intellectually and emotionally, to the activity and the observer/evaluator, who must bring detachment and objectivity to the evaluation process.

This report proposes an unconventional solution to this apparent dilemma: the one person who can meet the three criteria is the trainee. The trainee should be given a central and active role in all stages of the evaluation of impact. At the same time, the trainee should be trained/equipped to fulfill that role by means of:

- Orientation in the benefits and risks/hazards involved;
- Incentives for undertaking the additive task;
- Selective, short-term training in evaluative and information collection techniques; and
- Support, in the form of advice, cross checking and validation by others.

It should be noted that projects and programs can be evaluated even if they are poorly planned. This is the worst of two worlds since a poorly planned project is not only difficult to evaluate, but has weaker prospects for achieving the desired change. The creation of a new approach to impact evaluation recommended in this report presents a rare opportunity to devise a set of evaluation preconditions for use at the planning stage which will facilitate evaluation and at the same time increase the probability that impact will occur.

The training/implementation stage

This section briefly considers two sets of actions which arise within the training process. The first includes those initiatives which should increase the probability that the trainee will contribute substantially to host country development. The second describes pre-evaluative actions which will set the stage for subsequent impact evaluation.

The methodology calls for of several kinds of actions to occur during the implementation/training stage. The purpose of these actions is to ensure that the training experience is functionally linked to the needs of the host country and that the newly-acquired knowledge will be effectively utilized.

These actions are initiated by the trainee. They require continuing contact between the trainee and the stakeholders (probably through a single representative of the stakeholder group). They also require the support and participation of the stakeholders. These actions form a continuum with the post-training support and networking envisioned in the ATLAS project paper.

- Adjustments to the program of study as the trainee learns more about the discipline and his/her own professional interests;
- Modifications and clarifications of the individualized goal hierarchy as well as continual updating of information about the host country's development plans, labor market situation, etc.;
- Applications of academic experience to the host country situation. Trainees may conduct research or write papers that are directly linked to the home setting.

Although there are several reasons for continuing contact with the stakeholder group during the training period, the ultimate argument is that it is crucial to ensure that the graduate enters into a supportive work environment which needs, and will effectively utilize, his/her newly-acquired knowledge. Stated in more negative terms, the actions noted above are intended to avoid the isolation and neglect which frequently plague returning trainees and diminish their effectiveness.

Prior to training, the trainee is a latent change agent. Training is intended to raise his/her potential, as a change agent, to a level which is necessary, but not sufficient, to meet certain host country requirements. Measurement of both levels of potential - before and after training - will serve as the baseline(s) for subsequent evaluation of development impact and for identifying the impact leverage which that particular training investment produced.

The pre-evaluation activities at this stage are intended to identify and record information on at least three characteristics of the trainee which the training program can be expected to alter. These three characteristics should be measured before and after the training program. They are:

- Level of knowledge in the discipline;
- Attitude concerning:
  - the quality, relevance and utility of his/her training;
  - the American political, economic and social system and the values, rewards and sanctions embodied in that system;
  - the host country development effort and his/her role; and
- Behavior, including initiative, assertiveness, receptivity, etc., as manifest in the joint planning and consultation, training experience, etc.

The action stage: post training activities

The graduate's return calls for reconvening the stakeholders to assess the training experience and to launch the graduate into his/her post-training activities. These activities may include employment, participation in professional activities, self employment, research, networking, the twinning arrangement with the American university/professional society, teaching, etc. The stakeholders should also confirm the graduate's role in the preevaluation activities and take the appropriate measures to prepare the graduate for this role.

Impact Evaluation and the Various Dimensions of Change

This section describes the continuum within which development impact may occur. It begins with the individual trainee/change agent whose work initiates and sustains the change process.

The individual dimension

In addition to the educational and other qualifications normally applied in scholarship programs, the methodology described here calls for the trainee to be willing and able to:

- Study host country development plans and other relevant material;
- Participate actively in joint planning;
- Prepare an individualized goal hierarchy/career pathway;
- Participate actively in post-training networking and other follow-on activities; and
- Engage fully in the impact evaluation process, including systematic information collection.

The effective utilization of participant skills newly acquired in a scholarship program has often been inhibited by several factors:

- The returned participant may not have had adequate knowledge of, or access to host country development plans, priorities and programs either before or after training;

- There may ha 'e been very little systematic, long-term planning by the host country for the utilization of the participant's post-training skills; and

- Any sense of obligation and/or responsibility to work in areas of developmental priority may have been outweighed by the belief that the participant was a free agent and must be given complete freedom in choosing his/her post-training employment.

These factors may reflect weaknesses in the host country's capacity to plan and administer. They may simply result from a lack of awareness of how to engage participants in the development effort. They may be a manifestation of an extreme laissez faire attitude toward the post-training employment of returned participants.

In contrast, the proposed individualized goal hierarchy, and the process by which it is to be developed and implemented, are intended to be systematic and purposeful. Effective use of the goal hierarchy would enhance the trainee's contribution to host country development through an intensive process of joint planning, before, during and after training. If the joint planning is handled constructively, it should be possible to achieve three targets:

- To maximize both the relevance and the utilization of the participant's training;

- To define the participant's relationship to the host country's development priorities in both the private and public sectors in constructive rather than coercive terms, i.e., to provide a concrete incentive for the participant's full engagement; and

- To preserve the freedom of the individual to make informed career choices without arbitrary government pressure.

At a minimum, the methodology must seek to maximize the probability that the choice of trainee, training and post-training employment will fulfill the needs of the three parties. This can best be accomplished by employing the methodology at the planning stage. Indeed, the application of the impact evaluation methodology at the planning stage is a central precept of this report.

The organization/institution dimension

The organization/institution plays a critical role in the development process within the framework of the generalized and individualized goal hierarchies. This chapter examines that role, delineates the several stages of growth of the organization/institution in the development process and describes the contribution of the ATLAS graduate at each stage.[22, 63, 78, 90]

This section presents a model of institution building. The last part of this section identifies the performance of the organization/institution. At that point the developmental pathway is not necessarily linear. i.e., the developmental effects of the goods and/or services produced by the organization/institution may take several directions (see Chart 6).

EFFECTS CAUSED BY INSTITUTIONAL PERFORMANCE - NON-LINEAR PATHWAYS

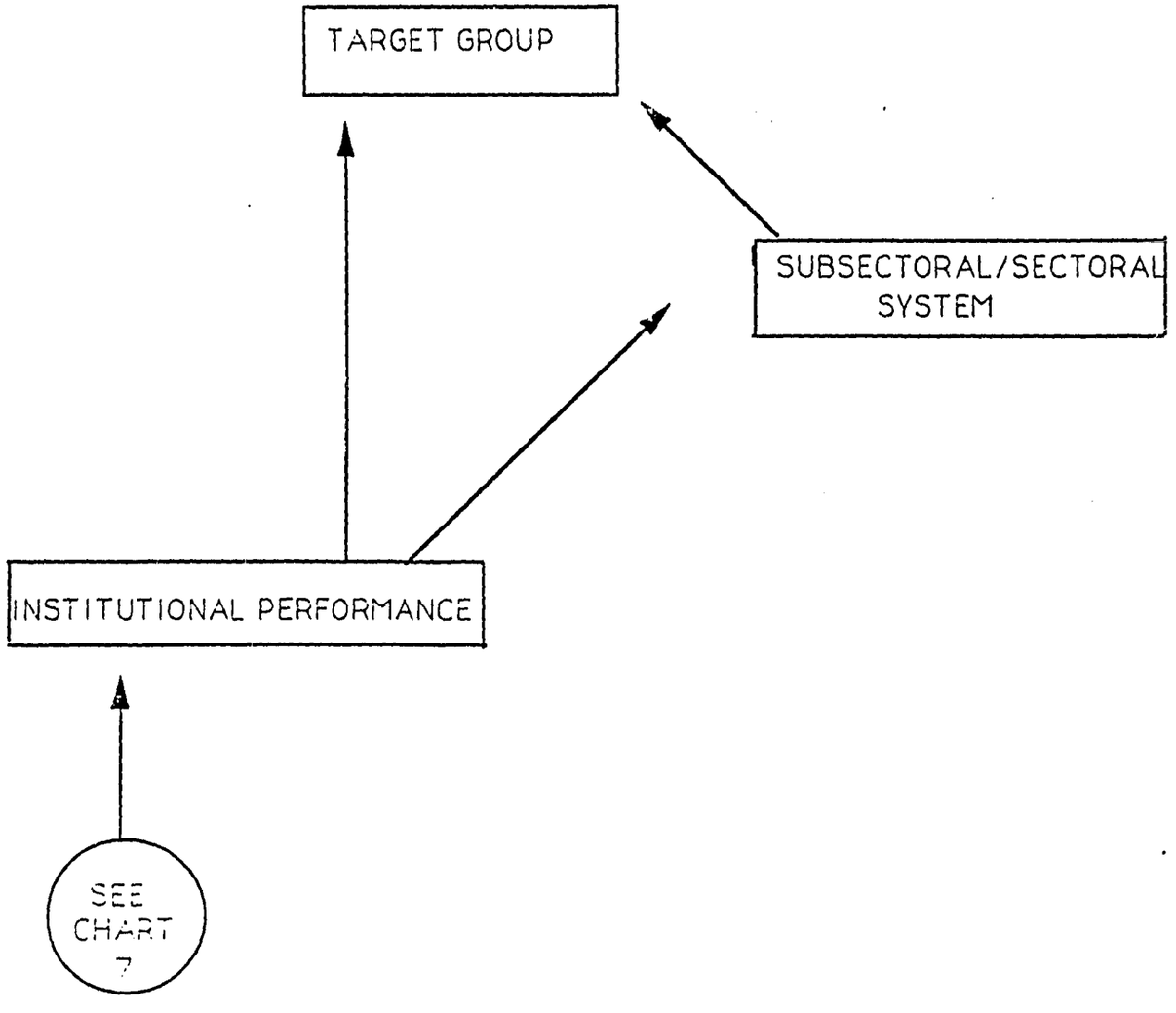


chart 6

First, organizational goods and services may be absorbed into a subsector/sector system and go no further, i.e., not have a substantial, visible effect on the lives of a target group which could be attributed back to the training. In this case, the goods and/or services produced by the institution may increase productivity, improve quality and/or reduce costs within the sector system in ways which have only marginal impact distributed over a large population, e.g., more reliable public utility services, lower cost highway construction and maintenance, improved administration within a government ministry.

Second, organizational goods and services may be absorbed into a subsector/sector system and be transformed into goods and services which directly support or are directly delivered to the target group, thus causing developmental impact, e.g., hospital management, educational planning and school administration, food storage and processing.

Third, they may be delivered directly to the target group/beneficiaries without entering into the subsector/sector subsystem, e.g., family planning clinics, vocational schools.

Fourth, they may take both of the two latter paths, i.e., the organization/institution may deliver its goods and/or services into a sub/sector system and directly to a target group. In each case, the concept of impact leverage will operate to govern the extent to which the institution's goods and services cause development to occur.

This section considers the kinds of changes a graduate might induce in an organization/institution, how the organization becomes an institution and the institution's interaction with its environment. The chapter articulates an empirically-based institution building model. The model has a hierarchical structure which is an integral subset of the generalized goal hierarchy. For institution building projects, the model would constitute the logical framework matrix.

The discussion at each level of the institution building model includes an explanation, examples of causal hypotheses within each level and between levels, illustrative indicators and assumptions about exogenous factors. At each level the model defines capacity (the antecedent) and performance (the consequent). The discussion is linked to the diagram of the institution building model on Chart 7.

The elements of the organization - the beginning of the institution building process

The elements necessary to the functioning of an organization include:

- Trained/qualified leadership and staff;
- Doctrine, operational policies and procedural systems;
- An organizational structure;
- An approved program;
- An approved budget;
- Equipment and material;
- An evaluation and feedback system; and
- Linkage to other organizations.

THE INSTITUTION BUILDING MODEL

CHART 7

GOAL HIERARCHY

LOGICAL FRAMEWORK MATRIX

Target Group/ Beneficiary Level

Institutionalization Based Upon Experience (lasting)

Increased Benefits to Target Group
Increased Production by Target Group
Increased Capacity of Target Group

Subsector/ Sector Level

Contribution of Institutional Goods/Services to The Subsector/ Sector

Institutional Level

Institutionalization - Based Upon Expectation (perishable)

Acceptance and Utilization of Goods/ Services by Target Group

PROJECT PURPOSE LEVEL

Institutional Performance: The Production and Delivery of Goods/Services

- Field demonstrations and Pilot Projects completed
- Completed Research
- Training courses completed
- Advisory Services delivered

PROJECT OUTPUTS LEVEL

INSTITUTIONAL CAPABILITY

- Trained/Qualified Leadership and Staff
- Doctine, Operating Policies and Preocedural Systems
- Functional Organizational Structure
- Approved Program
- Approved Budget
- Evaluation and Feedback System
- Equipment and Material

These elements and their interrelationships are preconditions to institutional performance, i.e., they add up to a potential capacity to perform. The existence of the elements and, to some limited extent, their interrelationships, can readily be confirmed at this level. The relevance and utility of the elements themselves, the viability of their interrelationships and such qualities as creativity, competence, skill, etc., must remain largely speculative, pending the next stages. In logical framework matrix terms, this is the output level and falls within the realm of manageable interest.

At this stage, the organization needs a capability for assessing and forecasting the effective demand for its products/services as well as methods for dissemination, marketing and distribution. It must also have the capability for evaluating its performance at the next levels in the model (see below) and feeding back the findings and conclusions of the evaluation into changes in the organization's elements. Most of the organizational elements noted above have a direct role to play in forecasting demand, dissemination of the goods and/or services and in evaluation and feedback.

As noted earlier in this Chapter, the formulation, reformulation and verification of causal hypotheses plays a major role in the ATLAS project cycle. Following is an example of two sequential, causal hypotheses within the organizational capacity level:

- If the ATLAS graduate is able to identify deficiencies in the organization's existing doctrine of limiting its program to research only, then he/she may be able to persuade the organization's leadership to equip and redirect the organization (retraining of staff, redesign of program, revision of procedures) to take on new functions such as research clearing house and extension/demonstration of research products from other sources.

- If the leadership is persuaded and acts, then the program, the composition of staff resources, the procedures and other elements of the organization would change and its overall capability would expand.

An example of a causal hypotheses between this level and the next level might be an continuation of the first example: the addition of new clearing house and extension/demonstration functions might produce a wider and more relevant/useful range of products and services for end users within the sector or the target group/beneficiaries.

Indicators of the existence of the organizational elements are usually obvious and information collection is relatively simple, e.g., staff size and qualifications, budgetary resources and program are easy to identify and document.

Assumptions about the behavior of exogenous factors at this level would mainly be concerned with the predictability of budgetary resources, effective demand, the effects of the labor market on staffing, the availability and cost of equipment and material, etc.

The contributions of the graduate at this level may be visible, relevant and substantial. They may also be readily evaluable, in the sense that observable changes in organizational capability can be clearly attributable to the efforts of the graduate, e.g., the formulation of operational procedures and the training of staff in new technologies. It should also be possible to determine, at least in approximate terms, whether, and to what extent the contribution depended upon the training. The several techniques for testing causal hypotheses cited in Chapter III, will be useful in sorting out the issues of attribution and criticality.

The changes described at this level do not constitute developmental impact but are early preconditions for such impact.

The production and delivery of goods and/or services by the organization

At this level, the organizational capacity (antecedent) is manifest in the performance (consequent) of the organization. Although performance marks the first step in the process of institutionalization, at this stage the organization is still an organization, not yet an institution. Organizational performance normally means goods and/or services produced (and/or value added) and delivered to a client/target group/beneficiary group. The goods and/or services might include:

- Research reports or products;
- Field demonstrations and pilot projects;
- Training courses;
- Advisory or material services;
- Basic needs such as food, clothing, shelter; and
- Manufactured products.

The organization's production and delivery of goods and/or services are intended to fulfill a developmental need. The performance criterion is the extent to which the goods and/or services are responsive to the goals and priorities in the host country's development plan.

While the goods and/or services may be critical to development, they are preconditions to developmental impact, rather than impact itself. They represent a next higher level of precondition for impact. This will be seen in the discussion of the next two stages in the institution building model.

An example of a causal hypothesis within this level might be: if, on the basis of his/her training and research, the graduate can improve and adapt an agricultural product, and the method for cultivating it, for application to host country circumstances - and if the product/method have a higher benefit/cost ratio than the existing host country equivalent -then the organization's product will be improved.

An example of two sequential causal hypotheses linking this level to the next might again be a continuation of the prior example. First, if the improved agricultural product were seen by the extension service as viable, then they would include it in their extension program and field demonstrations. Second, if it were promulgated by the extension service, and the farmers saw the product and the method for cultivation as economically advantageous, then they would adopt it.

The formulation of indicators at this level is straight forward since goods and services usually can be identified, observed and assessed in quantitative and/or qualitative terms. The behavior of exogenous factors would normally not play an important role except in those cases where market forces cause unexpected shortages of resources which are critical to the creation of the goods and/or services. Intervening (process) variables, such as bureaucratic frictions and inertias, would be more likely to disrupt production.

The contributions of the graduate might be visible, relevant and substantial, as in the prior stage. They may also be evaluable, in the sense that the amount and quality of the goods and services could be shown, at least in approximate terms, to be directly attributable to the efforts of the graduate and to his/her training.

In a logical framework matrix for an institution building project, this level would constitute the project purpose and the managerial interest would be problematic.

The next levels in the institution building model are the absorption of institutional goods and services into the subsector/sector system and the acceptance and utilization of the institution's goods and/or services by the target group/beneficiaries. Although these further effects of the institution's performance contribute to higher levels in the goal hierarchy, it should not be forgotten that the sectoral system and the target group/beneficiaries thus include levels of developmental effects which are within the scope of the institution building model.

The subsector/sector dimension

This section examines the possibilities for the design and evaluation of human resource development investments within the subsectoral/sectoral system.

The subsector/sector system is an arena in which resources are mobilized, allotted, invested, processed and distributed. The subsector/sector is defined more by substance, function and tradition than by formal organization or centralized control. The subsectoral/sectoral system can be viewed as a broadly but vaguely defined market system.

The system is composed of institutions and individuals, policies, relationships, standards and conventions, channels and activities, many of which tend to be loosely interlinked. It consumes and produces, it adds or subtracts value, it employs and serves/exploits the target group/beneficiaries. These constituent elements and traits moderate subsectoral/sectoral performance.

It is a bridge between the institutional and national/macro levels in the society. The separate treatment in this chapter of the organizational/institutional dimension and the sectoral dimension is necessarily arbitrary since the latter subsumes the former. The same overlap occurs in the next section where the national/macro dimension subsumes the subsectoral/sectoral. Also, recall the non-linear pathways where the institutional model enters into (i.e., overlaps) the sectoral system and the target group level.

The broad scope and loose definition of a sector, the complexity and uniqueness of individual sectors and the absence of formal structure and central control of sectors all conspire against the synthesis of a single theoretical or conceptual construct from which a methodological formulation for impact evaluation can be drawn. Within the contractual constraints of time and resources, it has been possible to consider only the simplest and most incomplete construct. Some tentative thoughts on the subject are summarized here.

The characteristics of the subsectoral/sectoral system which affect the possibilities for design and evaluation of human resource development and scholarship programs might be seen as centrifugal and centripetal forces within the sector:

Centrifugal	Centripetal
disparate interests and objectives within the sector	common substantive traits and technologies. market incentives
lack of common standards, no central leadership or control	regulatory mechanisms, market forces, cabinet ministry can provide some guidance and services
lack of communication and reliable information	government services and local organizations can assist
lack of policies/incentives for savings and investment versus consumption	macro and/or sectoral planning and statutes/regulations, competition in financial markets
weak coordination of donor and foreign investment resources	government and PVO efforts in planning and coordination

Similarly, it may be useful to contrast the conditions in the organizational/institutional dimension which affect design and evaluation with those at the subsectoral/sectoral levels.

Organizational/Institutional	Subsectoral/Sectoral
single organization	multiple organizations
monolithic or formal structure	multiple, diverse, informal, sometimes structured
single leader	nominal leader, largely decentralized
single mission	multiple missions
single product or set of products	multiple, diverse and competing products
specific market	broad and diverse markets
unified records	no unified records
single or few clientele	multiple clientele

The implications of these two sets of categories for the planning and evaluation of human resource development and scholarship programs are briefly discussed here:

- Trainee Entry Point

  The organization/institution, with its specific, defined mission, determines the trainee's entry point into the subsectoral/sectoral system (and/or directly into the target group level). The organization's output is the sectoral system's input. To the extent that the trainee's contribution to the organization's goods and services has been identified, that contribution can be further tracked to see what consequent changes it has made in the sectoral system's capacity and performance. The analytical tools and measurement techniques described in Chapter III and elsewhere are appropriate for this task.

- The Upward Direction

  Elements and linkages within the subsectoral/sectoral system are intended to have a substantial influence on the capacity, performance and benefits of the target group/beneficiaries. The objective of the scholarship program is to modify one or more of these sectoral system characteristics so that the target group changes its behavior. The logic of the goal hierarchy follows the sequence of change upward. Since the goods and services delivered by the institution are part of the subsectoral/sectoral system, that institutional product will affect a policy (e.g., import controls, tax regulations, interest rates), a commodity (e.g., school texts, fertilizer and herbicide), a service (e.g., mobile health clinics, a marketing mechanism, farmer cooperatives, housing subsidies) which may already exist in the sectoral system. These changes within the sector system will in turn permit the desired changes in the target group. The significance of these changes, and their potential for reaching the target group/beneficiaries can be largely explained by the concept of impact leverage, defined in Chapter II and discussed below.

  Looking upward in the goal hierarchy helps to identify the conditions which are necessary to change the behavior of the target group. At the same time, it should be increasingly easier to discern those other elements of the sector system, taken together with the contributions of the trainee, which will be sufficient to change target group behavior.

- The Downward Direction

  The subsectoral/sectoral system is expected to provide the conditions needed to bring about the desired change in target group behavior. The goods and services delivered by the organization/institution into the sectoral system, of which it is a part, is one of those conditions. The downward direction is intended to discover the other necessary conditions and to assure their adequacy.

  The goal hierarchy is a device for identifying the resources, actions, events and conditions which are believed to be essential. These include planned actions taken as part of the scholarship program. They also include conditions which are external to the program and which may be unpredictable and uncontrollable.

  In moving downward from the target level, each condition needed to bring about the desired target group change is identified as a sectoral system outcome. Similarly, each condition needed to strengthen sectoral system performance is identified as an institutional product. At each level there must be consideration

of both the necessary and sufficient conditions, whether these are produced by the original scholarship investment or are external to that investment.

The target group/beneficiary dimension

This chapter considers two levels of institutional impact on the target group/beneficiaries - antecedent and consequent.

It will be useful to recall again the non-linear pathways where the institutional model enters into/overlaps the sectoral system and the target group level. These pathways, and the kind and magnitude of induced change which they convey will be substantially affected by the impact leverage inherent in the institutional goods and services.

The first (antecedent) level is the provisional/tentative acceptance and use of institutional goods and/or services by the target group, based upon the expectation that they will be beneficial. At this level, the organization takes on the characteristics of an institution, but these characteristics are based upon the expectations of the target group and are therefore perishable. The characteristics of institutionality at this level are:

- The establishment of tentative communication and support links and relationships with suppliers, peer institutions and target groups/beneficiaries;
- The creation of predictable and stable sources of financing;
- The creation of effective demand and markets for goods and/or services;
- The development of internal resources, policies and methods;
- The ability to interact constructively with the operating environment and to adjust to externally imposed change; and
- The ability to evaluate its own performance and feed back the evaluative findings into institutional improvements.

The achievements at this level are generally that the clientele accept and utilize the goods and/or services of the institution on faith. For example, farmers buy new high yielding seeds and fertilizer and practice the new cultivation methods which the agricultural extension service recommends, although the farmer's behavior is predicated on benefits which are anticipated, but not yet realized. Until the benefits are realized and are seen to be sustainable over the long term, this level should conservatively be seen as a precondition for developmental impact.

The formulation of indicators of target group/beneficiary acceptance and utilization does not appear to be difficult, nor does the collection of information to support the indicators. The changes in target group behavior often can be expressed in quantitative form, e.g., numbers or percentages of farmers using the new production factors and methods, numbers or percentages of mothers attending pre-natal classes or using recommended contraceptive methods.

Assumptions about the behavior of exogenous variables may be more difficult, e.g., acceptance and utilization may be affected by unexpected fluctuations in farmer income, reductions in mobile health/population services occasioned by recession or problems of insurgency, etc.

The contributions of the individual graduate - and of his/her training - may be less clear, in part because it may have been a relatively small element with limited leverage at the institutional

production level or the subsectoral/sectoral level or because other causative forces and agents entered the process of developmental change and confounded the problem of attribution and criticality.

The second (consequent) level reflects the unqualified acceptance of the institutional goods and/or services by the target group/beneficiaries based upon the affirmation, through experience, that the goods and/or services are beneficial.

At this level, institutionalization is complete in the sense that the institution's interaction with its environment is constructive, mutually advantageous, stable and mature. The characteristics are parallel to those of the prior level except that they are no longer tentative:

- Communication links and relationships with suppliers, peer institutions and target groups, beneficiaries are productive and stable;
- Financial resources are adequate and reliable;
- Internal resources, policies and methods are established and functional;
- The institution is in equilibrium with its environment; and
- The institution practices self evaluation and self improvement.

The achievements at this level go far beyond the achievements cited at the prior level. The latter were limited to the acceptance and utilization of institutional goods and services by the target group/beneficiaries. At this level there are three closely linked kinds and levels of developmental impact.

First, there is the increased capacity of the target group. This may take the form of enhanced knowledge, understanding and skill. It may involve modifications in attitude, outlook and behavior. It may encompass the potential for greater personal productivity or the willingness to borrow money and take risks in order to increase the productivity of land, labor and/or capital. Increased capacity is the precondition for increased performance.

Second, is the actual increase in the target group productivity and output which results from and is directly attributable to the increase in capacity.

Third, are the benefits directly attributable to the target group's productivity and output. This level of benefit is subject to at least two constraints which will be treated later:

- The existence of equity or economic and social justice, i.e., the avoidance of exploitation or preemption of the rewards of increased productivity and output by others; and
- Benefit incidence or equitable distribution of benefits within the target group to those most deserving and in greatest need.

The national dimension

This section examines the possibilities for planning and evaluation of the contribution of human resource development at the national level. It is useful at the outset to differentiate between two kinds of benefits at the national level.

First, societal goals and benefits represent the national aspiration for economic growth, improved social relationships, general well-being, participation in the international order and national policy. These goals are sometimes stated in quantified form in national planning documents, sometimes articulated in inspirational oration by a national leader. They are largely political goals although they may be achievable mainly through economic means, if they are achievable at all. Rarely are they stated with sufficient precision or explicitness to serve as a basis for evaluation. They may be contradictory within the total set of national goals. Individually, they may be unattainable within the constraints of available resources. They may change substantially as governments change. All-in-all, they do not lend themselves as criteria for a rigorous impact methodology.

Second, national group benefits are changes in the income, health, education and other characteristics of well-being of groups which are broader than the target group. This class of impact beneficiaries may be identified in the host country development plan as meriting special attention and benefits because of its distinctive social or economic status. Enhancement of the well-being of that group would be consistent with the development priorities of the host country.

Impact leverage at the national level

It is commonly accepted folklore that the developmental consequences of a small technical assistance project cannot be identified or measured at the national level; that effects at the national level which are of a kind with the project's purpose are not readily attributable to the project because they cannot be disentangled from other causal factors; that small projects are suboptimal by definition; that the effects of small projects are invariably overwhelmed by exogenous factors which operate at national level; that the collection of data needed to establish attribution and criticality would be prohibitive in cost and suspect in reliability; etc. These beliefs may be largely true. Although the problem of evaluation at the national level is indeed daunting, never-the-less, it should be confronted. This section suggests how that confrontation might occur. It relies on the concept and uses of impact leverage.

The impact evaluation framework embraces the concept of leverage as a means for increasing the probability of both achieving impact and identifying and measuring impact. Leverage is defined here as achieving the greatest development impact with the simplest, smallest, most potent intervention/investment.

Leverage is conceptually analogous to, but operationally different from, the benefit/cost ratio. The benefit/cost ratio is a quantitative device which requires monetarization of both benefits and costs. Leverage, as it is perceived here, is a concept rather than a ratio. It is expressed in qualitative terms. It seeks to find the most powerful, lowest cost solution to a specific development problem which is widespread.

Leverage is also conceptually linked to the three ways in which impact occurs: primary/secondary - first/second generation, replication - spread effect, and multiplier effect. These are patterns for the propagation of impact. Leverage can give dimension and force to these patterns.

The relationship of impact leverage to benefit/cost ratios and to the three patterns of impact can be better understood if leverage is seen as an impelling force which drives the progression of ~~development effects~~.

These definitions are abstract. The following discussion is intended to make the real meaning of leverage clearer.

Even the smallest project has a development effect. At the national level, that effect may be microscopic and therefore difficult to identify and to measure - but it exists. To understand what

leverage is, and what role it might play in human resource development programs, three classes of leverage are examined here.

At one extreme is maximum leverage. This is illustrated by the cases of Sir Alexander Fleming, whose discovery of penicillin saved countless millions of lives and Norman Borlaug whose genetic transformation of cereal grains largely eliminated undernutrition and malnutrition throughout the developing world. Although events of this magnitude occur only at rare intervals, and although they may not be a practical model for scholarship programs, there is something valuable to be learned from this class of leverage, i.e., that the solutions to large, pervasive problems might be simple in scope and originate from one person's efforts.

A simple, low cost method for manufacturing enriched, processed food for preschool children from locally grown crops might reduce national child mortality rates. A single radio teaching product/technique might bring about a widespread increase in adult literacy. A low cost building material, and the technique for using it for sweat-equity, low income housing construction, could improve the prospects' for home ownership nation-wide. While these examples are far less dramatic than the Fleming-Borlaug triumphs, they do indicate the possibilities for a graduate to affect national development. The examples also illustrate that tracking development effects to the national level may be feasible.

The implications of this class of leverage are far greater for the evaluation of scholarship programs than for their design. It seems unlikely that one could plan purposefully to achieve higher-order leverage. Never-the-less the odds might be improved by using the goal hierarchy structures to optimize choices, both for the selection of candidates and the formulation of their career pathways.

At the other extreme is minimal leverage. In this class of leverage, the graduate's efforts result in a product or service which, however useful, simply vanish into the organization and/or the sectoral system and cannot be identified, and consequently not measured, either at the target group/beneficiary level or the national level. The examples cited in Chapter III, The Organizational Dimension, illustrate this class of minimal leverage at the national level. In these examples, the graduate's goods and/or services may increase productivity, improve quality and/or reduce costs within the sector system in ways which have only marginal impact distributed over a large population, e.g., more reliable public utility services, lower cost highway construction and maintenance, improved administration within a government ministry, better library facilities in public universities.

The implications of this class of leverage are far greater for the design of scholarship programs than for their evaluation. Anticipating post training leverage at the planning stage might be strengthened through use of the goal hierarchy structures for the selection of candidates and the formulation of their career pathways.

Between the two extreme classes of leverage described above is a middle ground where impact is less dramatic and pervasive than the first extreme class but more significant, and therefore easier to discover then the second. Examples of this middle ground might be:

- Simplification of credit/banking/cooperative systems which could facilitate personal savings, investments, lending and loan repayment operations;
- Discovery of a new/adaptation of existing means for disease control for livestock which could increase farmer income and reduce food costs for consumers; and
- Design of a simplified system for management of public health, nutrition and

population services which would sharply reduce administrative overhead, costs and delays in delivering treatment and thus permit wider coverage.

Impact may not be readily visible in this middle class of leverage, but it is susceptible to being made visible if the simultaneous impact tracking/data collection techniques suggested in Chapter III are followed.

Four further thoughts on development leverage are offered. First, the central lesson to be learned from the Flemming Borlaug examples is not the breadth of the impact, but rather the specific, sharply focussed nature of the solution.

Second, ideas are not always implemented simply because they are meritorious, and impact does not always occur simply because the causal agent (e.g., a product, a service, a method) is valuable and relevant. Change occurs when the conditions which permit and facilitate change are also present. For a causal agent to create development effects, there may have to be a means for dissemination/distribution; effective demand in the form of purchasing power; a willingness to accept innovation and take risks; a clear advantage over existing alternatives; etc.

Third, impact at the national level in a sub-saharan African country resulting from the work of a scholarship graduate is not nearly as remote a phenomenon as would be impact in India. Thirty two sub-saharan countries have populations of less than 10 million and of these, 11 have about 1 million or less.

Fourth, impact leverage, and its influence at all levels of the development hierarchy, may be a more critical issue for both donor and recipient than the pursuit of developmental effects at the national/macro level. Observation of both is best approached in the manner outlined for the three lower levels of the goal hierarchy: through the formulation/reformulation of causal hypotheses by the joint planners and verification at a later stage by the evaluators. This process is described in Chapter III.

The multinational dimension

This section briefly examines the multinational dimension at two levels. First, it raises a number of basic questions at the policy and program level to determine the possibilities for design and evaluation. It then looks at the methodological aspects which might emerge from those policy and program questions.

Policy and program concern

- Is it possible to intervene/invest at the subsaharan African level and to create change at that level? Does subsaharan Africa constitute an political/economic/social system which is capable of absorbing and responding to interventions/investments - as a system?

- Are the interventions - and consequent development effects - which are appropriate at the national level different and distinct from those at the subcontinent level?

- Is there such a thing as a subsaharan African agricultural sector, a subcontinent-wide education sector?

- What different United States policy interests are served at the multinational and national levels in Africa?

If these questions were posed in conjunction with the European Economic Community rather than subsaharan Africa, some of the answers to these questions might begin to be affirmative, albeit still tentative, in coming years. Although there is a modest level of trade among African countries, and even some transport infrastructure, there is not yet enough of an organic or systemic entity to permit affirmative answers to the questions above.

A specific theory of induced developmental change cannot explain the functioning of a system if there is no existing system. Subsaharan Africa does not yet constitute a political/economic/social system. Similarly, a specific theory cannot guide and predict the design and evaluation of a project/program if A.I.D. does not implement such activities at the subcontinent level. For these basic reasons, the methodology does not embrace the project/program design at the multinational level.

Methodological concern

- Is there a need for evaluative information on subsaharan Africa as a geopolitical unit versus evaluative information on individual countries? What is that need? How, by whom, and for what purposes would such information be used?

- Is it possible to measure induced change in subsaharan Africa as a geopolitical entity other than by adding up the measurements of change at national levels?

The absence of a subcontinent system and the concomitant lack of A.I.D. program activity at the subcontinent level means that the methodological framework proposed in this report is not capable of producing evaluative information which reflects developmental changes beyond the country level.

It may be useful to mention here the policy of the APRE Bureau of not comparing program performance among APRE missions. The Bureau compares the current and past performance of a country program. This policy reflects the Bureau's recognition of the uniqueness of individual country circumstances.

As noted in Chapter II, the only feasible means for collecting information on issues common to more than one African country is to identify and summarize individual evaluation findings from individual countries in small clusters with similar experiences. The result of introducing new cultivation methods and/or price and market incentives and structures in cereal grain production in East African countries might provide lessons for wider application within those same countries and in neighboring countries as well.

Conditions Affecting the Use of the Proposed Impact Evaluation Framework

This section considers four main aspects of the impact evaluation methodology and process:

- Principles, elements and criteria of the evaluation process;

- Causal hypotheses and methods of validation;

- Measurement of developmental change - objectively verifiable indicators; and

- Baseline.

Principles, elements and criteria of the evaluation process

Evaluation is the retrospective measurement and analysis of the results of an intervention. Evaluation measures induced change, compares the change against some standard and draws inferences from the comparison. Evaluation attempts to determine what happened, how and why.

At the evaluation stage, the individualized goal hierarchy (and the generalized goal hierarchy from which it was derived) would be the explicit and authoritative criteria against which to evaluate. Its authority would reflect two facts: First, that it was the direct result of a deliberate joint planning process based on national development priorities, and second, that it was an accurate and explicit formulation of the causal hierarchy of events, linkages and objectives which are expected to link the training with jointly agreed developmental changes. Even though the formulation of the individualized goal hierarchy may be informed and disciplined, it must be recognized that its criteria/objectives are limited by the ability of the planners to forecast the future.

Because developmental impact may be complex, unpredictable and not necessarily finite, the evaluation process must also look beyond the objectives which were explicitly formulated and reformulated at earlier stages in the two goal hierarchies. As defined in Chapter II, impact evaluation must be seen as open-ended in that it seeks to discover developmental effects which are planned or unplanned, desirable or undesirable, transient or lasting, direct or indirect, primary or secondary, immediate or delayed, intermediate or final. Here the ultimate criteria for evaluation are the development goals of the host country and A.I.D. and the welfare of the target groups/beneficiaries.

An intervention/investment can produce developmental effects which are substantially different in kind, intensity and magnitude. We present two extreme cases in simplified form.

First, an intervention/investment can produce an effect which progressively and substantially grows, replicates, spreads and multiplies over time within the development process. This impact is readily observable long after the immediate event and in a wide area/audience. Causality, attribution and criticality may be easier to determine at higher levels in the development hierarchy.

Second, an intervention/investment can produce an effect which is absorbed into the development process but does not, in any substantial way, grow, replicate, spread of multiply over time within the development process. While this effect may be valuable, even indispensable, it is not readily observable, except in the immediate aftermath and within a circumscribed area/audience. Causality, attribution and criticality may be more difficult to determine, even at the lower levels in the development hierarchy.

A third example, which is not necessarily mutually exclusive from the others, is the case where the intervention/investment creates a product and/or service which is sustained by a stable, effective demand. The product and/or service may ultimately reach a specific level in the goal hierarchy and not replicate, spread or multiply as in the above example. Sustainability, in and of itself, is not an objective, but it is important as a characteristic and as a measure of market value.

The methodology calls for the examination of two kinds of independent variables which are capable of affecting the contributions of the ATLAS graduate to the development process and confounding the evaluation of induced change. These are intervening/process variables and exogenous variables.

Intervening variables usually operate within the change process, e.g., between the graduate's efforts and the developmental effects which result from those efforts. The graduate's attempts to improve the capability of his organization, and consequently to expand its production of goods

and/or services will be affected by such intervening variables as institutional inertia, lack of equipment and budgetary resources, shortage of trained staff, the absence of institutional program doctrine and direction, internal battles over turf, etc. The graduate's attempts to disseminate research findings or market a product or service may be inhibited by a poor communications and transport system.

Exogenous variables reside in the political/economic/social environment, are independent of the training project, are generally beyond control and often are unpredictable and unexpected. A graduate's newly-created small business may be swamped by a large-scale multi-national corporation investment in the same area. A graduate's field trials of a new cultivation method may be destroyed by flooding in the wake of a typhoon. A graduate's attempt to reduce morbidity in farm animals may be thwarted by a national foreign currency shortage which halts the import of needed pharmaceuticals.

Causal hypotheses and methods of validation

It is important to recognize that it is not possible conclusively to prove a causal hypothesis on the basis of logic. It is possible only to increase the degree of confidence or credibility in a hypothesis. For example, it is not possible to verify the use of fertilizer (cause) and conclude that it resulted in increased crop yield (effect). It is also not possible to verify the existence of increased crop yield (effect) and from that fact, to conclude that it was caused by the use of fertilizer.

The predictive causal hypotheses formulated at the planning stage take the form: if A, then B. For evaluation purposes, the predictive hypotheses must be seen, and restated, in reverse order, as explanatory hypotheses: if B exists, it was caused by A. Evaluation requires the separate and independent verification of both the presumed cause and the presumed effect. If that can be done, then confidence in the empirical validity of the hypothesis is increased.

Although the science of logic is not applicable here, the weight of evidence and objective judgement, based upon experience, is applicable.

Once the effect B is observed, the evaluator must search for the causal agent, A., e.g., farmers will increase crop yields (effect), if they use fertilizer. If both increased crop yields and the use of fertilizer can be confirmed independently, then causality is at least tentatively demonstrated.

If the explanatory hypothesis is more narrowly stated, attribution and criticality also can be tentatively proposed: farmers will increase crop yields if, and only if, fertilizer is used. This kind of hypothesis statement begins the process of eliminating other possible causal agents.

There are several techniques useful in the verification of hypotheses, i.e., in causal inference. These are usually employed in combination.

The first technique is called the null hypothesis. Although it is not possible to prove a causal hypothesis, it is possible to disprove one. This opens the door to a logical means for increasing the level of confidence in a hypothesis. The most viable method is to create a null hypothesis, which is the negative complement of the original hypothesis, and then to disprove or discredit it, thus increasing confidence in the original hypothesis. Applying this approach to the prior example, the hypothesis is that the use of fertilizer will result in increased crop yields. The null hypothesis would be that there is no causal relationship between fertilizer and crop yields. A clearer statement of the null hypothesis is in terms of treatment and control: the difference in the crop yields of farmers who use fertilizer and farmers who do not use fertilizer is zero. The second, clearer version of the null hypothesis can be readily tested with empirical data. If it can be disproved or discredited, the original hypothesis gains in currency.

A second technique is to formulate hypotheses or explanations which compete with the causal hypothesis. For instance, a competing hypothesis in our example would be that new seed strains caused the increases in crop yields. Other competing hypotheses might be that treatment with herbicides and pesticides, or higher market prices for farm produce caused the increases in crop yields. If a competing hypothesis can be shown to be plausible or persuasive, then the original hypothesis loses currency. If the competing hypothesis can be discredited, then the original hypothesis gains currency.

Two additional tests of internal validity may be useful. First, the two variables,(i.e., the causal agent and the effect), must be shown to be systematically related to each other. This means that their nature, magnitude, proportionality and direction can reasonably be seen to be within a common, interrelated system of variables. Second, change in one variable must be shown to precede change in the other variable, thus establishing the direction of causality. In the case of an established positive correlation between health and income, the evaluator must determine if increased income permitted the person to purchase better health care and consequently enjoy better health or whether the achievement of a healthy state permitted the person to work more efficiently and consequently enjoy increased income.

The relationship may be found to be circular, with an increase in each variable causing the other variable to increase, and so on. Usually, the initial direction of causality between the two variables can be established by the precedence of change.

Measurement of developmental change - objectively verifiable indicators

Given the Agency's interest in indicators of developmental change over the past three decades, this section will not reiterate definitions and guidance which are already well known. Instead, several key aspects of measurement which are generally pertinent to human resource development and of specific interest for training programs will be noted.

Quantitative versus qualitative measurement

The perpetual confusion over the relative merits of quantitative and qualitative measures needs to be resolved:

- If the criterion is the validity of the measurement, then there is no generic superiority of one over the other.
- If the criterion is disinterested objectivity, neither quantitative/qualitative measurement is, in and of itself, necessarily more objective.
- If the criterion is reliability (i.e., can be reconfirmed through repetition), there is no difference as long as the observer is as accurate and meticulous in collecting numbers as he/she is in observing behavior.

The questions of validity, objectivity and reliability can be summed up in two ways:

- A number is no more valid or precise than the behavior which it represents;
- The validity, accuracy and reliability of both quantitative and qualitative measures can be confirmed equally well by the method of redundancy and congruence, i.e., by repeated measurement by other observers.

Comparability and aggregation

Quantitative measures are often assumed to be more useful than qualitative measures because of the belief that they can be compared, and sometimes aggregated, within and across sectors and within and across countries. Comparability and aggregation of information about developmental impact have long been pursued by the Congress, Agency decision-makers and the staffs supporting them. The desire to compare and aggregate is understandable when the investment is large and the outcome is both uncontrollable and unpredictable. In the underdeveloped world, real comparability and aggregation tend to be illusive - unsupported and insupportable by evidence. Developmental effects which appear on the surface to be comparable usually are not because of how and why they occurred.

The basic problem is that seemingly comparable results may conceal substantial differences in the underlying circumstances and causative agents. A per hectare yield of the x strain of wheat, which is the same in eastern Pakistan and northern India, may appear to have rich meaning for evaluators, technicians and policy makers. A closer look may show that the yield in eastern Pakistan was mainly resource-driven, i.e., the result of new cultivation practices, seed strains and newly available, low cost inorganic fertilizer. The yield in northern India was largely market-driven, i.e., the result of farmer response to the incentives of higher wheat prices and guaranteed, pre-negotiated crop purchases by a marketing cooperative. Comparing and aggregating information on these crop yields and transmitting it to Congress and the Agency's leadership would be a disservice.

Generic, macro-level indicators versus situation-specific indicators

The desire for information on developmental effects which can be compared and aggregated (see above), manifests itself in the search for broad, generic indicators for use in overall program planning, resource allocation and policy formulation. At the other end of the spectrum, situation-specific indicators are required at the design, monitoring and evaluation stages of individual projects and programs. If both of these approaches are believed to be effective measures of developmental change and achievement, then clarification is not only needed but urgent.

Generic, macro-level indicators, such as those used in Washington policy and program guidance documents, and, to some extent, in mission planning documents, are most valuable as statements of policy direction and priority. They are not operationally useful for the actual measurement of developmental change for several reasons.

First, they usually do not embody values, substantive needs and characteristics or contextual (exogenous) factors, except at a generalized/abstract level. Being virtually value-free, substance-free and context-free means that they carry information which has almost no learning/explanatory value for the user.

Second, by their very existence, macro-level indicators imply that they subsume and integrate lower level information which is comparable and aggregated. Comparability and aggregation of information about developmental effects - whether at the planning level or at the results level - is a chimera and cannot be taken seriously.

The measurement capability of situation-specific indicators is much different. Because they are usually formulated by operational staff, and are part of the project/program design, implementation, evaluation, feedback cycle, situation-specific indicators can be value-rich, relevant to the substantive characteristics and closely indicative of contextual factors. In the hands of a knowledgeable professional, their potential capacity for valid, reliable and accurate measurement of developmental change is very high. This is particularly true if host country personnel are directly involved.

For the reasons noted above, the impact evaluation methodology presented here relies solely on situation-specific indicators.

Indicators

The term indicator, like the term impact, is used within the Agency in ways which are inconsistent and imprecise. Clarification is needed.[58, 66, 69, 85]

An indicator is an explicit and objectively verifiable measure of induced change and/or results expected. It illuminates and elaborates the various aspects of a project objective. Indicators have no existence of their own. They derive from, and are mirror images of some developmental change, objective or result. If the change/objective is simple and quantitative (e.g., fertility rate or crop yield), it may be possible to characterize that change with one or two simple quantitative indicators. If the change is complex and at least partly qualitative (e.g., farmer acceptance of new technologies and institutional services), it may be necessary to formulate a larger set of qualitative and quantitative indicators to comprehend all of the relevant dimensions of the change.

Indicators should have the following characteristics:

- Comprehensivenes

  Each developmental change/objective should have a set of indicators sufficiently comprehensive to illuminate all of its relevant aspects.

- Validity

  Each indicator should accurately and proportionately reflect variations in the quality, quantity, intensity, level, etc., of the change.

- Specificity

  Indicators should reflect only the change being observed and should exclude extraneous factors.

- Independence

  In a hypothesized causal relationship, cause and effect must each be separately and independently measured by independent sets of indicators.

- Objective verifiability

  The indicator must present evidence (either qualitative or quantitative) which is unambiguous and incontestable, i.e., which will receive the same interpretation by more than one observer.

- Targeting

  Indicators must be explicit in terms of magnitude, quality, location and time.

- Corroboration

  A limited amount of redundancy in indicators can serve to corroborate the measurement of change, particularly when surrogate or proxy indicators are

used. Redundancy is insurance against the effects of unforseen variables and misleading signals in the measurement process.

- Supportability

  Reliable data to support the indicator must be obtainable easily at reasonable cost.

Baseline

A training program requires two kinds of baseline:

- A record of the kind and level of knowledge of the trainee in his proposed field of specialization prior to and after training. This measurement should occur at the beginning of the joint-planning process. The baseline should include the trainee's awareness and familiarity with host country development needs, plans and priorities. This latter item can subsequently be compared to what exists at the end of the joint-planning process.

- The kind and level of capacity and performance of the organization/institution in which the graduate will be employed - if that is known in advance. Chapter III contains guidance on institutional indicators as well as examples.

It may be useful to note the relation between baseline information and indicators of developmental change. Ideally, the baseline includes all of the change (dependent) variables which may be affected by the treatment. At the time the baseline is established/measured, the indicators for these variables may be set to zero if no activity or development is present. Usually however, there has been some prior activity. In that case, the indicators should be set to register that existing level.

Impact Indicators

A.I.D. has a long and rich history of formulating and using indicators. Guidance and listings of indicators are abundant although not yet organized, uniformly understood or employed effectively. Within the limits of this task order it was not possible to search exhaustively, to assess, or to structure the Agency's storehouse of experience and material on indicators. The search did discover, however, that much valuable material has been lost. For instance, it was not possible to find the extensive work on indicators done by the PPC evaluation office in the mid and late 1970s; the LAC Bureau was unaware of, and could not find the computerized data base of performance and achievement indicators developed by the bureau's evaluation division in the late 1970s.

In this chapter we discussed the limitations of using indicators for purposes of comparability and aggregation. We examined the shortcomings of generic, macro-level indicators for measuring and explaining induced change. These cautionary notions suggest that the broad-based,higher-level indicators, however useful in policy dialogue and overall program planning, may have limited predictive, explanatory or learning value at the operational level. With these notions in mind, the following criteria were used for selecting the indicators which follow.

First, in keeping with the succinct definition of development impact specified in Chapter II, the listing includes those indicators which measure changes in the lives of the target group/beneficiaries. It excludes indicators of the pre-conditions for impact found at the institutional level and at the subsector/sector level. It excludes broad, generic statements of national aspiration which have no practical utility in project/program design and evaluation.

Second, it emphasizes the situation - specific, i.e., areas and populations most likely to be affected by the work of the graduate.

Several other characteristics should be noted. First, the indicators are illustrative. The indicators which are stated in general terms can be reformulated to measure:

- A single state or level of development, i.e, a benchmark;
- An incremental amount or percentage of change over a defined period, i.e., a rate of change, a trend;
- A change from one level to another level; and
- Final achievement/non-achievement of a stated target.

Second, there is some overlap among categories because the classification of target group/beneficiaries covers some individuals in the other sectoral categories.

- Agricultural and Food Production, Processing and Distribution
  - Farmer access to physical factor inputs; purchase and utilization of physical factor inputs,
  - Farmer access to land ownership/rental; purchase/rental of land,
  - Farmer access to capital; borrowing for capital improvements/crop production costs; repayment performance,
  - Farmer access to markets; sales; profits,
  - Land productivity; yield per hectare,
  - Labor productivity; yield per man year,
  - Capital productivity; yield per capital investment; yield per crop production costs,
  - Use of new technologies, diversification of crops,
  - Agricultural output, trends and rates, percentage of GDP,
  - Value added through processing, enrichment, preservation, packaging, services,
  - Domestic and export earnings from traditional and non-traditional farm produets,
  - Farm household income, earnings from off-farm labor.[53, 62]
- Health and Nutrition
  - Life expectancy,

- Infant, child and maternal morbidity and mortality,
- Birth weight,
- Birth attended by trained/qualified health worker,
- Food consumption (broken out by classes of food),
- Access to physician/nurse/midwife/hospital/etc.,
- Incidence of preventable diseases,
- Deaths averted.

• Family Planning

- Access to family planning services by women of child-bearing age,
- Acceptance and use of family planning methods by women/men, including contraceptive prevalence rates, couple years of protection, etc.,
- Crude birth rate,
- Total fertility rate,
- Birth attended by trained/qualified health worker.

• Education

- Literacy rate,
- Enrollment in primary/secondary/tertiary school:

  a. total
  b. female
  c. pupil/teacher ratio
  d. dropout rate/completion rate
  e. performance in standardized tests,

- Qualifications of teachers at each level of school:

  a. levels of qualifications
  b. percent of total teacher population for each level
  c. rate of change of a and b,

- Education expenditures relative to other sectors, per student, per teacher, per school.

• Industry/Manufacturing

- Domestic sales of products to target group/beneficiaries,

a. clothing
b. shelter and furnishings
c. medicine and health/sanitation products
d. food (included in Agriculture and Food Production and Health/Nutrition headings above),

- Non-agricultural jobs created, broken out by occupational categories,

- Wages paid, broken out by occupational categories.

- Commerce and Trade (per capita where possible)

  - Consumption by target group/beneficiaries, broken out by classes of products, commodities and services,

  - Employment of target group/beneficiaries, broken out by occupational categories,

  - Value added in export marketing, new products entering export markets,

  - Trade balance, trends and rates,

  - Investment, public and private, formation of new enterprises,

  - Access to capital, levels and trends in borrowing by private entrepreneurs, by sector. Number of new banks, branches, credit unions, cooperatives. Interest rates and terms of lending,

  - Prices and availability of consumables, trends and rates of change.

- Target Group/Beneficiaries

  - Income in monetary and non-monetary forms with characteristics: a. level, b. stability/fluctuations, c. sustainability

  - Consumption

    a. energy consumption, absolute and as percentage of income
    b. food consumption broken out by calories, protein, etc. stability/ fluctuations, as percentage of income
    c. clothing and footware
    d. housing, fuel and utilities
    e. health care (see health care above)
    f. education
    g. ~~transport~~
    h. other consumer durables
    i. taxes

- Savings

  - Savings, dissavings, debt and investments.

- Amount, rate and interest earnings and payments

  N B difficult to collect for individuals but may be available for target group from public records or sampling.

- Women in Development

  - Risk of dying by age,
  - Life expectancy,
  - Maternal morbidity and mortality,
  - Education,
    a. Persistence in grade _ as percentage of cohort
    b. Females per 100 males (by primary, secondary, tertiary)
  - Employment, levels by occupational categories and age, relative to male employment,
  - Earnings, levels by occupational categories and age, relative to male earnings.

BIBLIOGRAPHY AND END NOTES

1. Academy for Educational Development, Training for Development Review of Experience. Draft. John A. Gilles, 1991

2. Academy for Educational Development, Vocational Educational and Training, Review of Experience. Draft. Dennis R. Herscabach, Franceds B. Hays, and David P. Evans, 1991.

3. African American Institute AFGRAD Phase III, Cooperative Agreement AFR-0045-A-00-5020-00, Oct. 1988 - Sept. 1989.

4. African American Institute The Study of AFGRAD Alumni: Training High Level Human Resources for African Development, 1963-1980., May 1988.

5. Aguirre International, The Fourth Annual Report of The Caribbean and Latin American Scholarship Program, April, 1990.

6. American Technical Assistance Corp. Preliminary Design of an Evaluation Methodology Beyond the Specific Project Level, Albert L. Brown and Edmond C. Hutchinson, 1975-1977.

7. Bernbaum, Marcia CLASP II Impact Evaluation, May 1989.

8. Bernoff, Robert; Girault, Emily and Rowe, Mary Budd The Application of Temporary System Concepts to Effective Planning and Management of Short Term Educational Programs.

9. Blomqvist, Ake "Higher Education and the Markets for Educated Labour in LDCs: Theoretical Approaches and Implications", World Bank Discussion Paper, The World Bank, January, 1987.

10. Brinkerhoff, Derick W. Improving Development Program Performance: Guidelines for Managers, Boulder, CO: Lynne Rienner Publications, 1991.

11. Brinkerhoff, R.O. Achieving Results from Training, San Francisco: Jossey-Bass, 1987.

12. Chen, Huey-Tsyh Theory Driven Evaluation, Beverly Hills: Sage Publications, Inc., 1990.

13. Chen, Huey-Tsyh and Rossi, Peter "Evaluating with Sense: The Theory Driven Approach" in Evaluation Review, vol. 7, no. 2, Sage Publications, April, 1983.

14. Clark, Peter Action Research and Organizational Change, London: Harper and Row Publishers, 1972.

15. Diambomba, Miala, Universities and Development in Africa: Problems and Challenges for Planning in Prospects for Educational Planning: A Workshop Organized by IIEP on the Occasion of its XXVth Anniversary, Francoise Callods (ed), IIEP, LAI, 1991.

16 Dietz, Thomas and Pfund, Alicia "An Impact Identification Method for Development Program Evaluation" in Policy Studies Review, vol. 8, no. 1, Fall, 1988.

17. Eicher, J.C. "Education Costing and Financing in Developing Countries: Focus on Sub-Saharan Africa", World Bank Staff Working Paper #655, 1984.

18. Flad, Harry Robert *Evaluating the Impact of Training: A Model for Use within an Organizational Framework*, unpublished doctoral dissertation, Boston College, 1989.

19. Freeman, Howard E., Rossi, Peter H., and Wright, Sonia R. Evaluating Social Projects in Developing Countries, OECD 1979.

20. GAO Progress in Implementing the Development Fund for Africa GAO/NSIAD - 91 - 127, April 16, 1991.

21. Guba, E.G. and Lincoln, Y.S. Fourth Generation Evaluation, Newbury Park, CA: Sage Publications, 1989.

22. Gulley, James Lee *Evaluating de Impact of International Training for Development in Zaire: A Case Study of the Studies and Planning Service, the Dept. of Agriculture,* unpublished doctoral dissertation, American University, 1987.

23. Haddad, Wadi; Carnoy, Martin; Rinaldi, Rosemary; and Regel, Onporn Education and Development Evidence for New Priorities, World Bank Discussion Papers (bibliography).

24. Hope, K. "Indicators of the State of Society" in M. Bulmer (ed.) Social Policy Research, London: Macmillan Books, 1978.

25. Hinchliffe, Keith Higher Education in Sub-Saharan Africa. London: Croom Helm, 1987.

26. Kless, Steven J. "The Economics of Education: A More than Slightly Jaundiced View of Where We are Now" in The Prospects for Educational Planning.

27. Labat-Anderson Institute *Private Sector Needs Assessment Human Resources Development Assistance Project*, USAID/Mali, Feb., 1989.

28. Labat-Anderson Institute, TIMS Report (Training Information Management System), Functional Specifications and Implementations Plan, 1991.

29. Lynton, R. and Pareek, U. Training for Development, West Hartford: Kumerian Press, 1978.

30. J. Lukomski, M.F. Participant Training: An Evaluation of the USAID/Lisbon Participant Training and a Suggestion for a Standardized A.I.D. Participant Training Planning, Monitoring and Evaluation Methodology, 1983.

31. Macadam, Robert Systems Agriculture - An Alternative Approach to Agricultural Education, Research and Extension for Rural Development, Perspective in Education, vol. 4, no. 1, 1988.

32. Macadam, Robert and Bawden, Richard "Challenge and Response: Developing a System for Educating More Effective Agriculturalists" Prometheus, Vol. 3, no. 1 June, 1985.

33. Paul, Samuel; Steedman, David; and Sutton, Francis "Building Capability for Policy Analysis" World Bank Working Paper, 1989.

34. Psacharopolous, G. "Higher Education in Developing Countries: A Cost Benefit Analysis", World Bank Staff Working Paper #440, 1990.

35. Sanguinetty, Jorge and Kimmel, Paul "A Comparative Method to Evaluate the Developmental Impacts of AID assisted Training", Washington, D.C., 1985.

36. Schwartz, Hugh and Berney, Richard Social and Economic Dimensions of Project Evaluation InterAmerican Development Bank, 1977.

37. Simpson, Carl, *JTPA Evaluation at the State and Local Level, Vol. IV: A Guide for Gross Impact Evaluations*, ERIC documents, March 1986.

38. Steele, Roger E. (ed.) Programs for Building Professional Capacity in Asia: A Seminar Report, Winrock International Institute for Agricultural Development, 1989.

39. Sudma, Seymour and Bradburn, Norman Asking Questions, San Francisco: Jossey-Bass Publishers, 1989.

40. Sullivan, Louis and Roskens, Ronald *A Report to the President: Child Survival and Aids in Sub-Saharan Africa: Findings and Recommendations of the Presidential Mission to Africa*, January 4-18, 1991.

41. Tan, Jee-Peng, Lee, Kiong Hock, Mingat, Alain "User Charges for Education: The Ability and Willingness to Pay in Malawi", World Bank Working Paper #661, 1990.

42. Teachers Corps Research Adaptation Framework, Evaluating Programmatic Impact in Education, Florida State University, 1976.

43. Thompson, Mark S. Benefit-Cost Analysis for Problem Evaluation, Sage Publications, 1980.

44. Tilak, Jandhyah *Education and Its Relation to Economic Growth, Poverty and Income*

Distribution, World Bank Discussion Papers, 46 Research #4, The World Bank, 1989.

45. *USAID Administrator's Evaluation Agenda, FY 1991-1993, January, 1991.*

46. USAID, *An Approach to Evaluating the Impact of AID Projects*, March 1986, no. 5.

47. USAID/Africa Region, *Assessment of Program Impact Reports for Fiscal Year 1990.*

48. USAID/AFR Human Resource Development Assistance (HRDA) Project Paper. 1987.

49. USAID, *AID and Education: A Sector Report on Lessons Learned*, January 1984, no. 12.

50. USAID, *An Assessment of the Impact of AID's Participant Training Programs in Nepal, March 1990.* A Special Study, no. 68.

51. USAID, *AID Evaluation Handbook*, April 1989, no. 7.

52. USAID Audit Reports 85-08; 87-05; 9-000-87-5 and 9-000-87-7.

53. AID Evaluation "Occasional Paper no. 33 Indicator of Households Income for Use in the Evaluation of Agricultural Development Projects, Oct. 1989.

54. USAID Evaluation Occasional Paper No. 11 Review of Participant Training Evaluation Studios, Moser, T., and Elmer, L. 1986

55. USAID, *AID Evaluation Occasional Paper #3, Flexible Design Approaches*, January, 1986.

56. USAID/Botswana Limited Country Development Strategy Statement FY 1991, August 1989.

57. USAID/CDIE/AFR Impact Evaluation of AID Policy Reform Programs.

58. USAID/CDIE Tracking Program Performance - A Review of Objectives and Indicators, Management Systems International, February, 1991.

59. USAID Conducting Group Interviews in Developing Countries Program Design and Evaluation Methodolgy Report No 8, April 1987.

60. USAID, *Conducting Key Informant Interviews in Developing Countries*, Dec. 1989, no. 13.

61. USAID Development Finance Institutions: A Discussion of Donor Experience, July 1990, no. 31.

62. USAID, Development Fund for Africa--The Impact of Rural Credit Projects in Africa: A Synthesis Report, March 1990.

63. USAID, *Effective Institution Building: A Guide for Project Designers and Project Managers Based on Lessons Learned from the AID Portfolio*, March 1982, no. 11.

64. USAID, *Fresh Start in Africa: AID and Structural Adjustment in Africa*, August, 1990.

65. USAID/Ghana Country Program Strategic Plan, FY 1992-1996, July 1991.

66. USAID, *Guidelines for Data Collection, Monitoring and Evaluation Plans for AID--Assisted Projects*, April 1987.

67. USAID/Guinea Country Program Strategic Plan FY 1992-1996, September 1991.

68. USAID Handbook 10 - Participant Training, 1988.

69. USAID, *Indicators for Measuring Changes in Income, Food Availability and Consumption and the Natural Resource Base*, Sept. 1989, no. 12.

70. USAID/Kenya - Country Program Strategic Plan FY 1990-1995, March 1990.

71. USAID/Kenya Country Training Strategy: 1989-1990.

72. USAID Program Overview, Latin America and the Caribbean/Education and Human Resources Development, 1991.

73. USAID/Malawi Country Development Strategy Statement-Action Plan FY 1991, December 1989.

74. USAID/Mali Action Plan: FY 1990-1991, May 1989.

75. USAID/Mali Country Development Strategy Statement, FY 1990-1994, May 1989.

76. USAID Policy Determination (PD-8) *Participant Training*, July 13, 1983.

77. USAID Policy Paper *Basic Education and Technical Training*, Dec. 1982.

78. USAID Policy Paper *Institutional Development*, Mar. 1983.

79. USAID Policies and Programming in Education, vol. I, II, *Creative Associates International, Inc.*

80. *USAID* Prism: *An Agency-wide Program Performance Information System for Strategic Management: A Plan for Design and Management, 1991.*

81. *USAID Project Identification Document*, Africa Regional African Capacity Building Initiatives (698-0536), 1990.

82. USAID Project Impact Evaluation Report, *Ethiopia: Alemaya University of Agriculture; no. 71, Washington, D.C., June 1989.*

83. *USAID*, Methodologies for Assessing the Impact of Agricultural and Rural Development

Projects: A Dialogue, June 1989, no. 11.

84. USAID, *A Manual for Evaluating Small Scale Enterprise Projects*, November, 1985, no. 6.

85. USAID, *Rapid, Low-Cost Data Collection Methods for AID*, December, 1977, no. 10.

86. USAID, *To Education in Nepal: A 20 Year Beginning*, May 1981, no. 19.

87. USAID, *Universities for Development: Lessons for Enhancing the Role of Agricultural Universities in Developing Countries*, August 1989.

88. USAID U.S. Assistance for Africa, The Development Fund for Africa, May 1989.

89. Van Bruggen, Johann "The Problems and Possibilities of the Impact Evaluation of Institutes for Curriculum Development" in Studies in Educational Evaluation, vol. 13, 1987.

90. Van Reenen, Gert-Jan and Waisfisz, Bob *Final Report on Institutional Development*, Institutional Development Advisory Services, The Hague, April 28, 1988.

91. Yeuker, Robert unpublished Proposal for doctoral dissertation, Geroge Washington University

92. Ziderman, Adrian "Social Rules of Return to Manpower Training Programs: The Policy Context", Phree Background Paper Series, Document no. Phree/88/04, World Bank, Nov. 1988.

93. Assorted other documents from USAID, IBRD and other multilateral and bilateral donor governments and organizations

Interviews/contacts to date: (9/15/91)

Graham Donaldson, IBRD

Barbara Searle, IBRD

Ron Ridker, IBRD

Han Chan, IBRD

Hans Heinrich Thias, IBRD

Mervin Weiner, IBRD retired

Peter Davis, APRE

Krishna Kumar, CDIE

Emmy Simmons, CDIE

Maury Brown, CDIE

Martin Hanratty, ENE

Robin Horn, AFR

Joe Carney, LAC

Jerry Briton, CDIE

Nena Vreeland, CDIE

Cindy Clapp-Wincek, consultant

Chris Hermann, APRE

Gary Hansen, CDIE

Tom Donnally, OIT

Gene Peuse, AFR

Bill Renison, AFR

Heather Monroe, African American Institute

Hazel Bush, African American Institute

Anita Johnson, AAI

Peter Hansen, United Nations

Robert Yeuker, consultant

Angela E.V. King, United Nations

Jerry Wolgin, AFR

ATTACHMENTS

IMPLEMENTING A TRAINING IMPACT EVALUATION METHODOLOGY

An Initial Operational Guide

International development aims to improve the quality of life among people in the developing world. Education and training, treated as an integral part of the development process, make contributions to the achievement of improvements in the quality of life. The improvements, however, are difficult to define and measure. It is, therefore, equally difficult to determine an appropriate level of investment in education and training or to argue for support of the education and human resources sector within the context of the development strategy of the Agency for International Development. Our task is to propose a solution to problems associated with the measurement of training results.

In this document, we offer a strategy for the implementation of a training impact evaluation. The purposes of the document are to:

- Present a methodology for training impact evaluation;
- Provide guidelines for the implementation of the training impact evaluation;
- Present a questionnaire that serves as a basis for the implementation of a training impact evaluation;
- Propose a strategy for testing the guidelines and questionnaire in a field setting.

Each of these components of the strategy is discussed below.

The Methodology for the Training Impact Evaluation

The methodology for the training impact evaluation consists of a theoretical base and an operational framework that is derived from the theoretical base. Each of these two elements of the methodology will be discussed in brief. The reader should refer to the main body of this report for an indepth discussion.

A Theoretical Base

The theoretical base for the training impact evaluation methodology is derived from a precise definition of impact, from a theory of planned change, from a multi-level model of the system that is being changed through development activity, and from a theory of training that is congruent with both the planned change theory and the multi-level model of the system being changed.

First, we define **impact** as changes in the quality of life of a target population. International development projects and programs seek, ultimately, to make positive improvements in the quality of life of the people benefitting from assistance.

DEVELOPMENT IMPACT=IMPROVEMENTS IN THE QUALITY OF LIFE

This narrow and specific definition of impact limits the number of induced changes that can be called impact and it gives rise to the notion of development impact preconditions. Many of the

changes that are caused by development interventions such as training are preconditions for impact, but they are not impact. In other words, many improvements, such as building a road, are necessary components of development and, joined with other improvements, will lead to impact. These preconditions are critical to development and to the achievement of impact. It is essential to understand and measure both preconditions and impacts if we are to undertake a thorough and useful training impact evaluation. Therefore, we must have the capacity to define preconditions as well as to formulate indicators to measure both impact and preconditions.

Second, our methodology calls for the creation of a specific **theory of induced change.** In brief, that theory holds that change in social systems can be planned and managed. It can be caused, measured and understood. Social systems are seen as existing in a state of relative equilibrium until an event or events causes disequilibrium. During the state of disequilibrium, changes occur in the system. Such changes can be introduced through training as well as through other means. When a change has been consolidated, the system returns to a state of equilibrium and continues to perform in new ways. We must be able to imbed a training impact evaluation in the context of a changing environment.

Third, we provide a **model of the system** being changed. It is a model of a system that is comprised of six related levels of organized society in which development efforts intervene. We call this model the generalized goal hierarchy. The six-level hierarchy includes the individual, the institution, the sectoral systems, the target group, the country, and the multi-country/continent levels. This model suggests that while the individual is the unit of change being targeted in training, the individual nearly always will perform in an institutional context and that institution will nearly always perform in a system of institutions before impact can be caused. This model stresses the importance of an individualized goal hierarchy that ties the trainee to the country development plan. The model also suggests that changes at the level of nations or continents; that is, changes at the macro level, be viewed as aggregations of changes among target groups.

CONTINENT } aggregate of countries
COUNTRY } aggregate of target groups

TARGET GROUP

SECTORAL SYSTEM } preconditional level of impact
INSTITUTION } preconditional level of impact
INDIVIDUAL } preconditional level of impact

The focus of the model is at the level of the target population receiving the benefits of development. However, it is important to consider all levels to evaluate training impact.

Fourth, we rely on a **social systems theory of training.** This theory of training holds that in order for training to cause a change in the social system at the institutional and sectoral systems levels, training must be purposively related to the institution or system that is being changed. In simple terms, training must start and end in that institution or sectoral system. Training must be planned within that context; it must be managed so that it continues to be related to that context, and it must be followed up with specific actions taken in that target institution or sectoral system. Put another way, training must be a part of the system it is intended to change. A training impact evaluation must have the capability to determine that the training system and the target of change are well joined to produce results. This relationship between training and targeted social systems constitutes a set of preconditions for training impact.

An Operational Framework

The operational framework translates the theoretical bases of the methodology for training impact evaluation into a structure for applying the theory. The framework appears on the following page. Each column and cell in the framework is discussed below.

Levels of Intervention

The framework begins with the levels of intervention that are affected by development efforts. At the bottom of the column is the individual who is the trainee participating in the ATLAS program. The next level is the organizational or institutional level in which the trainee will work or otherwise participate once his or her training is completed. Above the institution is the sectoral level which is comprised of the numerous organizations and institutions involved in the development effort. Next in the hierarchy is the target group that receives the benefits of the institutional and/or sectoral outputs. At the latter two levels of the country and Africa as a whole, training is unlikely to make a direct intervention. These levels are viewed as aggregates of target groups.

These levels of intervention constitute a generalized goal hierarchy. The generalized goal hierarchy represents, in operational terms, a set of inter-related aims that A.I.D. regional, Mission, program and project staff generate in concert with host governments and beneficiaries to provide the context for specific planning and programming. At the outset of conceptualizing the regional, country and program or project strategy, it is critical that A.I.D. leadership and evaluators as well make explicit the theory of development on which actions will be based and the causal hypotheses that constitute that theory. Evaluation is responsible for testing those causal hypotheses about development (which are specific to each planning unit, the region, Mission, or program/project) and for testing a generic set of hypotheses, discussed below, that are generalizable to all A.I.D. training.

Phases of the Training Intervention

The training intervention is described in three phases: planning, implementation, and post training. Each phase is important to achieve the kinds of results which will be necessary to bring about impact. Our methodology calls for the testing of hypotheses related to each of the phases of the training intervention. The following is a discussion of each of the phases of the training intervention and the related hypotheses.

- Planning

  The planning of training involves collecting and analyzing information to arrive at decisions regarding needs for training, types of training needed, selection of trainees, design of training programs, duration of training, cost and many other elements of the program. Within an A.I.D. regional bureau or Mission, planning is likely to begin at the sectoral level, so it is here that we will begin the discussion of training preparation.

  Training may be identified as a need by any sector officer including EHRD (Education and Human Resources Development). Potentially, several sectors are represented during planning of training and may also include the regional bureau. ATLAS is an example of a program planned initially at the regional bureau level (Africa) and relying on detailed planning to occur at the Mission level.

LEVELS OF INTERVENTION	PHASES OF THE TRAINING INTERVENTION			CONTEXT VARIABLES	PRECONDITIONS FOR DEVELOPMENT IMPACT	IMPACT INDICATORS
	PLANNING	IMPLEMENTATION	POST TRAINING			
AGGREGATIONS OF DEVELOPMENT IMPACT: AFRICA				• POLITICAL VIABILITY AND STABILITY • SOCIAL VIABILITY AND STABILITY • ECONOMIC CONDITIONS		
AGGREGATIONS OF DEVELOPMENT IMPACT: COUNTRY						
DEVELOPMENT IMPACT: TARGET GROUPS	-representation at planning stage -target group needs, capacities and performance known to trainee	-monitoring and support to trainee reflects unique target group requirement	-work assignment appropriate to unique target group requirements		-income -employment/productivity -advancement,consumption -access to economic/social services -life quality improvements	
PRECONDITIONS FOR DEVELOPMENT IMPACT: SECTORS	-participation of government, business, industry and academia in planning -access to sector planning -donor coordination -adequacy of needed sector resources and services	capacity for managing flow of resources and actions among institutions and donors	-appropriate assignment within sector/subsector -coordination within sector and among sectors -adequacy of sectoral resources and info.		-sectoral/subsectoral needs for trainee knowledge -sectoral/subsectoral capacity -sector/subsectoral performance	
PRECONDITIONS FOR DEVELOPMENT IMPACT: ORGANIZATION/ INSTITUTION	-participation of top management in planning -strategic paln for use of ATLAS training -plan for org. development	-monitors and supports trainee -monitors and trainee research/field work -trains work units and prepares for reentry of trainee	-ensures appropriate work assignment for trainee -provides support services -provides on-job training -provides adequate income and benefits		-demonstrated capacity of institution to utilize trainee knowledge -improvements in institutional productivity and production -improvements in institutional outreach	
PRECONDITIONS FOR DEVELOPMENT IMPACT: INDIVIDUAL TRAINEE	-trainee actively participates in planning career -trainee has mentor -host country, employer, AAI, USAID participate in planning -access to development plans, priorities, human resource plans, labor market analysis	-faculty guidance -contact with host country mentor -research/fieldwork relevant to career plan and to host country dev. plans and priorities	-employment in chosen field -involvement in prof. networks -practical training/ updating/teaching		-personal income -career advancement -contributions to development	

HYPOTHESIS: Development impact increases when planning for training is directly tied to planning for development at the sectoral or multisectoral level. Training is a strategic intervention chosen for its potential to contribute to development outcomes.

Closely related to planning at the sectoral level is planning at the organizational or institutional level. The organizations and institutions that will be responsible for managing and operating development programs and sending and receiving trainees who are a part of development efforts are likely to be actively involved in planning directly with A.I.D. or through mediating structures such as training committees.

HYPOTHESIS: (a) Development impact increases when planning by sectors involves representatives from the organizations and institutions that are critical to performance of the sector(s). (b) Involvement of the top management of such organizations is preferable and is more likely to contribute directly to the efforts of the sector. (c) Organizations and institutions that have human resources development plans and plans for individual trainees are most effective in getting results from training as required by the sector development strategy.

At the micro level of the training intervention is the individual trainee.

HYPOTHESIS: The trainee's contribution to development impact increases when (a) the trainee actively participates in career planning tied to the development initiative in his/her organization or sector, (b) the trainee has a mentor who is part of the institution or sector.

While the target group is defined as the beneficiary group that a development effort is aiming to impact, target groups may also be a part of the planning of training. Women, for example, may be targeted as beneficiaries and representatives of that group may be involved in planning.

HYPOTHESIS: Development impact on a particular target group increases when target group representatives participate in the planning of training.

In this hierarchy of intervention levels, we have treated the country and Africa levels as simple aggregates of the target group level. We have not developed these two levels as active participants in ATLAS.

- Implementation

The implementation of training consists of all of the events that occur between a trainee's departure for training and his or her return to the country of origin. These events may include academic coursework, specialized practical training, internships, extra-curricular activities, independent study, networking, research, thesis preparation or other events. The implementation of training begins with the individual trainee.

HYPOTHESIS: Contributions to development impact increase for the individual trainee when (a) the program of study is made specific to the development plan being implemented by the institution and sector, (b) the trainee is actively monitored and supported towards achieving learning

outcomes that are relevant to the development plan, (c) the trainee engages in relevant learning experiences outside of and in addition to the academic program of study.

The organization or institution in which the trainee will work or otherwise participate after training, can be an active and effective player in the training implementation process.

HYPOTHESIS: Contributions to development impact increase when organizations/institutions are involved in monitoring and supporting trainees during their training.

The sectoral level has an indirect role in training while trainees are in study programs. At the sectoral level it is important that leaders stay aware of changes in plans, target groups or other factors that may affect the nature or content of training that is underway.

HYPOTHESIS: Contributions to development impact increase when sector representatives monitor their sector and specific development efforts and advise the trainee of needed alterations or additions to the training program.

The target group level is most likely to participate in training implementation in one of two ways--as trainees or as members of a sectoral monitoring committee.

HYPOTHESIS: Contributions to development impact increase for specific beneficiaries when target group representatives participate with other sector representatives in program monitoring and feedback.

Once again, the national and Africa levels are not likely to be active participants in the training implementation other than through sector or target groups and aggregated coordinating mechanisms.

- Post-Training

Post-training consists of actions taken by the trainee immediately upon return to his or her country of origin. Post-training includes job placement and professional activities undertaken outside the job. It includes follow-on training or any efforts to continue to refine the skills needed to perform successfully in the selected career. Post-training begins with the actions of the individual trainee.

HYPOTHESIS: Contributions to development impact increase when (a) when the trainee returns to and assumes a job relevant to his or her training and to the country's development needs and (b) when the trainee is involved in follow-on training that reinforces the academic learnings in an applied setting.

The organization or institution in ~~which the trainee~~ is working after return is very important in providing opportunities for the trainee's academic education to be used in the organization/institution.

HYPOTHESIS: Contributions to development impact increase when organizations or institutions employing returned trainees actively support the trainee in performing relevant tasks for which the trainee was educated.

The sector system may or may not be an active player in the post-training phase. Frequently, however, professional associations, interorganizational task forces or other networks become a support system that provides the trainee with reinforcement, additional training, contacts and other resources to perform effectively in the field.

HYPOTHESIS: Contributions to development impact increase when sectoral systems actively support and promote development efforts in which the trainee is involved.

At the post training stage, we expect to see direct involvement of the trainee with the target group in whatever way that involvement is appropriate to the trainee's field.

HYPOTHESIS: Contributions to development increase when the target group or representatives thereof are participants in the professional activities of the trainee either as recipients of services as as partners in the trainee's development work.

The national and Africa levels are not treated as participants in post-training.

Context Variables

Context variables are those factors that exist apart from and independent of the training program that are likely to act as barriers to or facilitators of impact. These context variables can, in fact, dramatically change the potential for impact by altering the environment in which the trainee's skills were expected to be used. Context variables need to be examined at each phase of the evaluation process including at the planning, implementation and post training phases and at the stage of evaluating target group impact. Context variables include political, economic and cultural/social factors that are identified at the design stage of a training program and are tracked for variation in nature and significance throughout the development process. Context variables affect the individual, institutional, sectoral, target group, national and Africa levels of development efforts.

Impact Preconditions

Following the planning, implementation and post training stages of a training program, a trainee, presumably is working and beginning to be part of a stream of inputs and resources that, taken together, are impacting on the target population. While it may be difficult to measure impact and tie that impact directly to an individual trainee or group of trainees, it is possible to assess the capacity and performance of individuals, organizations and sectoral systems in contributing to and benefitting selected target populations. Impact preconditions include both economic or productivity indicators and noneconomic indicators. Several months to several years after the training program ends, these types of measurements are made:

- Individual Level

  ~~Individual~~ capacity. The trainee has the skills, knowledge and attitudes to perform in the job.

  Individual performance. The individual trainee is performing in the service of development. He or she is providing goods or services needed by the organization to reach the target group.

- Organizational or Institutional Level

  Organizational or institutional capacity. The organization has the human and other resources that are sufficient to carry out the development mission related to the target group.

  Organizational or institutional performance. The organization or institution is delivering goods or services to the target group.

- Sectoral or System Level

  Sectoral capacity. The sector(s) for which personnel have been trained has a sufficient combination of institutional resources to impact on the target group. (e.g. Ministry, university, business and NGOs viewed as a system)

  Sectoral performance. The sector system is delivering the services and goods to the target population.

- Target Group Level

  Target group capacity. The intended beneficiaries have the capacity to receive and use the services and goods that the sectoral system delivers.

  Target group performance. The intended beneficiaries are using the services and goods that are delivered by the sectoral system.

National or Africa-wide capacity and performance, for our purposes, are measured as aggregates of target group preconditions.

Impact Indicators

Because impact on target populations is produced by an array of inputs including but not limited to A.I.D. training and other resources, impact indicators are defined at a level beyond the project or program level. Having defined impact as quality of life improvements at the target group level, then training can never be a sufficient condition for producing impact. It can be an important element in the array of inputs and it is the role of research and evaluation to determine how and when training does contribute to development results.

Impact indicators, by our definition, are relatively few. They are not economic indicators. Rather, they are indicators of human development to use the language of the UNDP. They are indicators that are established by each Mission/country and are likely to include:

1. Health status indicators including disease control, nutritional status, mortality;

2. Indicators of physical safety and security including disaster control and availability of shelter;

3. Indicators of freedom from ~~oppression including basic~~ human rights, participation in government, equitable treatment and access to the community's resources.

4. Indicators of sustainability of life systems including limits on environmental degradation and long-term access to healthful natural resources (e.g. clean air and water);

5. Economic indicators such as income, consumption, savings, investment.

Guidelines for Implementation of the Training Impact Evaluation

A comprehensive training impact evaluation is conducted in five parts. The training impact evaluation is longitudinal with a training program and continues after training is completed. The training impact evaluation must involve the EHRD sector and may be guided by that sector, but it also needs to be integrated into the Mission's impact assessment plan. Following is a description of the components of the training impact evaluation plan and the relationship of each component to the overall evaluation scheme of the Mission. The five components of the training impact evaluation are: pre-evaluation study, formative project evaluation, summative project evaluation, program impact evaluation and development impact evaluation.

Pre-Evaluation Study

Pre-evaluation study is preparation for the following four stages of the evaluation process. Pre-evaluation investigations involve the Mission, the host country, and a variety of individuals and institutions that will determine the scope and process for the implementation of a longitudinal training program impact evaluation. The purposes of the pre-evaluation are to:

- Articulate the Mission development impact plan;
- Establish the causal hypotheses on which the development impact plan is based;
- Determine the impact indicators that are being tracked at the Mission level;
- Determine the role of training in the Mission's development strategy;
- Determine the relationship between the impact indicators and the Mission training plan; establish the role of ATLAS in the Mission training plan;
- Establish the role that the Mission development plan and its related training plan play in the country development plan including the relationship to other donor activities;
- Determine the data sources available at the Mission, country or institutional levels that will be used as part of the system for evaluating development impact and the Mission's contribution to development impact.

On the basis of the data collected at the pre-evaluation stage, it should be possible to determine the viability of the training impact evaluation and to determine any gaps or barriers that stand in the way of implementation.

Formative Project Evaluation

Formative project evaluation is, as the name indicates, project level evaluation that occurs prior to and during implementation of a training program. Formative evaluation tests the hypotheses listed above under those two headings. Formative evaluation is highly descriptive and focuses on both the process and the results of planning and implementation of a training activity. The purposes of the formative evaluation are to:

- Revalidate the impact indicators that are being tracked at the Mission level and identify those indicators that are being addressed to some degree by training interventions.

Describe the target groups that are the beneficiaries of the intended impact.

- Describe the sectoral systems that have been organized to deliver the desired impact to the target groups and define the role that A.I.D. is playing in that sectoral system;

  Identify changes in the sectoral capacity or performance that are anticipated as part of the Mission's development efforts.

- Describe the organizations and institutions that comprise the sectoral system and identify those that are slated for training including ATLAS training;

  Identify the changes in the organizational or institutional capacity or performance that are anticipated as part of the Mission's development program.

- Describe the individual trainee population using Mission records;

  For selected trainees, describe the capacity and performance changes expected to result from educational programs.

- Evaluate the training planning process at the individual, institutional/organizational, sectoral and target group levels.

- Evaluate the training implementation process at the individual, institutional/organizational, sectoral and target group levels.

- Assess the vulnerability of the training effort at the planning and implementation stages to context variables that are political, social or economic in nature.

The formative evaluation is introduced at the Mission level with revalidation of the impact indicators and and agreement on the impact preconditions that are driving the training project or program. The formative evaluation examines the individual, organizational/institutional, sectoral and target group levels using several methods and data sources.

At the Mission level, the data collection methods include individual and group interviews and documentary sources. Interviews must include the Director, planning officers and the leadership of the sectors targeted by the Mission and must involve target group representatives. Documentary sources must include Mission-level planning documents, country plans, sectoral planning documents and other guidance on Mission direction. During the formative evaluation, evaluators determine the role of target group representation in the planning and implementation of training.

At the sectoral level the Office of the Chief of Human Resources and his or her counterparts in other sector offices must be included. Documentary sources include project papers, training plans and other documents indicating the role of training in the development process in that Mission and in particular sectors and institutions or organizations. Evaluators must determine the measures of capacity and performance that are reasonable measures of sectoral development and they must describe how those data can be obtained, ideally from existing data sources. During the formative evaluation, the evaluators assess the role of sectoral representatives in the planning and implementation of training.

At the institutional and organizational level it will be important to include leadership of the organizations whose personnel are being trained by A.I.D. With institutional representatives evaluators must determine the measures of capacity and performance that are reasonable measures

of institutional development and they must describe how those data can be obtained, probably through existing data sources. During the formative evaluation, evaluators assess the role of institutions and organizations in the planning and implementation of training.

At the individual level trainees are the primary data source along with documents and records that illustrate the planning and implementation process for the trainee. Those documents should define the capacity and performance changes that are intended to be the result of education for the individual trainee and they should document how the education program and its management are serving to meet those requirements for improved performance and increased capacity.

Summative Project Evaluation

Summative project evaluation is project level evaluation that occurs immediately at the end of a project within 6 to 12 months of the project's completion. Referring to the impact evaluation framework, summative evaluation occurs at the post training stage and tests the hypotheses related to that stage. Summative evaluation is highly descriptive and descriptive data are supported by output measures such as numbers trained, numbers employed, etc. The purposes of the summative evaluation are to:

- Revalidate the indicators of impact on groups that are being targeted by the Mission and revalidate those indicators that are associated with training. Obtain any impact data that are available.
- Revalidate the indicators of impact preconditions at the sectoral level Obtain any impact data that are available.
- Revalidate the indicators of impact preconditions at the organizational/ institutional level. Obtain any impact data that are available.
- Determine the training that has been completed and evaluate post training results at the individual, organizational, sectoral and target group levels. Link these results to an analysis of the potential for impact on preconditions.
- Assess the vulnerability of the training effort at the post training stage to political, social and economic factors.

Summative evaluation occurs at the Mission level, but focuses, primarily on the actions of trainees at the individual level, on the actions of employers at the organizational/institutional level, and on the related actions of organizations and institutions at the sectoral level. The primary data sources are the trainee and the employer with linkages being made to the larger context of post training actions that are occurring along with other event in the sectoral system.

Program Evaluation of Preconditions

Program evaluation of preconditions occurs after the completion of a training project and aims to measure impact preconditions and relate training to the achievement of those preconditions. Program evaluation of preconditions tests the causal hypotheses established at the start of the program and revised continuously throughout its implementation. The evaluation of program preconditions occurs at the individual, organizational/institutional, sectoral and target group levels of program activity and attempts to assess the effects of training along with other inputs in producing preconditions for impact. Such evaluation may be conducted best by developing a case study of one sectoral system, by assessing achievements of each of the institutional members of the system, by evaluating the role of training in achieving institutional outputs and by relating the

performance and capacity of the sectoral system to actions at the target group level. Program evaluation of preconditions must occur at the sectoral or Mission level in most cases, because training is likely to be but one of several interventions that are being used to establish impact preconditions. The purposes of program evaluation are :

- For graduates who are part of the sectoral system, to evaluate individual capacity and performance and establish the role that the trainee is playing in achieving impact preconditions;
- For institutions and organizations, to evaluate the increases in capacity or performance that have occurred as a result of the contributions of graduates along with other inputs;
- For the sectoral system, to evaluate the capacity and performance of the system in in making improvements at the target group level;
- For the target group, to determine the extent of the capacity to receive and use the goods and services of the sectoral system.
- To assess the effects of contextual factors (political, social and economic) on the capacity and performance of the sectoral system and on the capacity and performance of the target group.

Development Impact Evaluation

Impact evaluation is the measurement of changes in the quality of life among target populations. Development impact evaluation extends beyond the trainee and must be conducted within the context of the total development effort.

The Questionnaire

The training impact evaluation questionnaire joins together the hypotheses associated with training impact and the preconditions associated with development impact and results in a preliminary set of questions for each of the levels of intervention identified on the evaluation framework. Furthermore, each level of intervention will be questioned at each of four stages of evaluation that follow pre-evaluation.. In summary, a master questionnaire must include:

- Country or Africa-level Information

  This information is comprised of aggregated data.

- Target Group Level Questionnaire

  Aimed at the Mission and country levels and including target group representatives, this questionnaire is designed to determine target group participation in the planning and implementation of training, the effects of training on target groups at the post training stage, impact preconditions at the target group level, and target group benefits as measured by impact indicators.

- Sector Level Questionnaire

  Aimed at the sector level within the Mission and in relation to host country

institutions, this questionnaire is designed to determine sectoral participation in the planning and implementation of training, the effects of training on the sectoral system at the post training stage and the achievement of impact preconditions at the sectoral level.

- Organization/Institution Level Questionnaire

  Aimed at employers at the institutional/organizational level of intervention, this questionnaire is designed to determine organizational/institutional participation in the planning and implementation of training, the effects of training on the organization or institution immediately after training and the achievement of impact preconditions at the organization/institution level.

- Individual Level

  Aimed at trainees at the individual level of the generalized goal hierarchy, this questionnaire is designed examine the development of an individualized goal hierarchy and to determine trainee participation in the planning and implementation of training, the effects of training on the development-related actions of the the trainee immediately after training and the contributions of the trainee to impact preconditions at the individual level.

At each level the questionnaire includes questions related to contextual factors having the power to affect training impact.

The draft questionnaire outline is attached.

TRAINING IMPACT EVALUATION

Questionnaire Outline

Mission/Country Level Evaluation

Mission/country level evaluation is for these purposes: to establish baseline data during the pre-evaluation stage and to assess development impact.

Pre - Evaluation

The primary data sources for this level of evaluation are documents containing Mission and country development plans. The CPSP, national development plan and other comprehensive planning documents are consulted. Through these documents and through discussions with Mission leadership and country representatives pre-evaluation establishes a baseline and a point of departure for the design of a longitudinal evaluation. Some illustrative questions follow.

1. What are the development impacts towards which the Mission aims? What impact indicators are being tracked by the Mission?

2. What are the causal hypotheses imbedded in the development impact plan?

3. Who are the target beneficiaries of Mission development activities?

4. What are the Mission's strategies for impacting target groups?

5. What is the role of training in the Mission's development strategy? Specifically, what is the role of ATLAS? Establish linkages between the national development plan, the Mission plan and the ATLAS program.

Formative Evaluation

Evaluators revalidate the causal hypotheses that are at the heart of the development plan being implemented by the Mission. They determine how changes in the Mission/country plan or implementation context might be expected to affect the training plan and its implementation.

Summative Evaluation

Evaluators obtain available impact data and work with the Mission to reassess and renegotiate impact indicators and related strategies. Evaluators draw implications for the training program and feedback data to the training implementors.

Program Impact Evaluation

Evaluators obtain available impact data and work with the Mission to reassess and renegotiate impact indicators and related strategies. Evaluators draw implications for the training program and feedback data to the training implementors.

Development Impact Evaluation

While measurement at the Mission level is beyond the direct responsibility of the training impact evaluation, they are, of course, closely related. The Mission development impact evaluation answers these illustrative questions.

1. What are the indications that development impacts are being (have been) achieved?

2. What causal hypotheses (related to Mission programming) are being supported or rejected by the impact data?

3. Who is benefitting and how are groups benefitting from Mission interventions?

4. How effective are the Mission strategies for impacting target groups?

5. What are the effects of training interventions (of ATLAS) on the Mission's development strategies?

Individual Level Evaluation

The individual trainee joins the training process at the planning stage and, therefore, becomes a subject of and a participant in evaluation at the formative evaluation stage. He or she is a subject of the evaluation throughout the development process and for as long as his or her activities can be defined within and attributed to changes at the program impact and development impact stages.

Formative Evaluation

During training program planning and implementation, the trainee is a central feature of development interventions that involve human resources development. From the beginning of the training planning process, we support and assume and active role for the trainee as an instrument of change in the development process. These questions are relevant.

1. How is the trainee involved in planning for his or her training? What are the skills, knowledge and attitudes that the trainee has demonstrated and how are they measured? (What is the individual's baseline for training?)

2. What is the content of and by what process is the trainee developing an individual goal hierarchy?

3. How is the individual training and career plan (i.e. individual goal hierarchy) related to the country's development plan? What is the quality of that plan in terms of its relevance and specificity and how well does the plan identify the trainee's change agent role?

4. How much time, from initial contact to program start-up, is devoted to planning and preparation of the trainee?

5. Are the selection criteria evident to the trainee and does he or she clearly meet the criteria? Is the trainee committed to the training program's development goals? How is the commitment evidenced?

6. Is the trainee aware of support from his or her organization, profession or sector and is there a plan for contact and interaction during the training?

7. Does interaction between the trainee and the organizational;, professional, or sectoral support system occur and does that interaction influence the content and quality (relevance) of training?

8. Is the training institution (university faculty and counselors) aware of the trainee's development aims and does the institution actively support the trainee's learning along lines consistent with the trainee's development plan (individualized goal hierarchy)?

9. Does the trainee engage in extracurricular activities that support his or her training plan?

10. Does the trainee undertake specific projects during the academic program that are directly related to the training plan and to the role he or she will play in the development of the country (Papers, thesis, research, dissertation, conference participation, etc.)?

11. Is the trainee actively involved in monitoring his or her own program in relation to events back home so that program adjustments can be made if needed? How does this occur?

Summative Evaluation

Summative evaluation addresses the trainee's post-training actions. These questions are relevant.

1. What skills, knowledge and attitudes has the trainee acquired as a result of training?

2. What employment does the trainee obtain? Was the job the one for which the trainee planned? If not, is the job directly related to the trainee's preparation and to the national development plan to which the individual training plan related?

3. Is the trainee using his or her skills, knowledge and related experience acquired during training?

4. In what non-employment activities related to the individual training plan and to development goals does the trainee engage?

5. In what other ways is the trainee contributing to the sector or target group?

Program Impact Evaluation

At this stage of the assessment the individual trainee must be viewed in the context of a development program and in interaction with other inputs.

1. What is the contribution of the individual trainee to the institution/organization capacity to deliver goods and services to the sector system or target group?

2. According to measures of individual performance, how is the trainee contributing to the delivery of goods and services to the target group or sector system?

Development Impact Evaluation

For some individuals, it may be possible to measure their impact on target group beneficiaries. While these are unusual cases, they can and should be captured.

Organizational Level Evaluation

At this stage of the evaluation, we will look first at the organization or institution that is the

employer of the trainee. Employing organizations/institutions are likely to be involved in training from the earliest planning stages. These questions should be posed to employers.

Formative Evaluation

1. How are employing institutions and organizations involved in planning for training? Are they clear about their institutional or organizational role in contributing to development impact?

2. How are employers involved in setting and using selection criteria? How does the selection process ensure adherence to the criteria?

3. How is the employer involved in career planning for trainees? How is the employer involved in creating the individualized goal hierarchies for the trainees?

4. In what ways, if at all, is the employer involved in the design or selection of the training program?

5. What support system has the employer put in place to ensure that the trainee is actively informed and supported during training?

6. Does the support system (e.g. mentor) actively support the trainee during training? How? How is the trainee's program modified as a result of the interaction?

7. Does the organization engage in planning and preparation for the return of the trainee and is that preparation directly related to the role the individual will play in making contributions to development?

Summative Evaluation

1. Does the employing organization assign the returning trainee to a job relevant to the development mission for which he or she was trained?

2. Does the employing organization support the trainee in using his or her newly acquired and relevant skills and knowledge to contribute to the organization's role in development?

Program Impact Evaluation

1. As a result of the training of one or more ATLAS participants, what capacities have the organization acquired that have the potential to impact the sector system or target group?

2. As a result of the training of one or more ATLAS participants, what is the organization doing to impact the sector system or the target group?

Development Impact Evaluation

1. What role does the organization play in achieving impacts at the sectoral or beneficiary group level?

2. What development impacts (if any) can be attributed to organization performance or to improvements in sectoral performance?

Sectoral Level Evaluation

The sector level is comprised of a system of organizations and institutions that together produce impact. The contributions of an individual trainee may or may not be mediated by a sectoral system, but sectoral representation is nearly always seen in an A.I.D. development effort from the earliest stages of planning.

Formative Evaluation

Project papers and other sector level planning documents are useful references for this exercise.

1. What development impact goals of the Mission are the responsibility of the sector? What are the contributions to target groups that are the responsibility of this sector? What individuals, organizations and institutions comprise the sector system?

2.. Describe the planning process undertaken by the sector and describe how training in general and ATLAS in particular are part of the sectoral strategy.

3. Did sector representatives participate in establishing selection criteria or in selecting participants? How?

4. Did sector representatives have input into the design or designation of training programs?

5. Are sector representatives involved in monitoring training or in providing input to trainees for purposes of strengthening the relevance of their training programs?

Summative Evaluation

At the sector system level, contact with and influence on the trainee can occur through networks such as professional associations or alumni groups.

1. Upon the return of the trainee are professional associations available to and accessed by the trainee to reinforce his or her training and the development mission he or she is intended to fulfill?

2. Does the sectoral system function to support the appropriate placement and use of the skills of the returning trainee?

Program Impact Evaluation

1. As a result of ATLAS training, what capacities have been acquired by the sector that were unavailable in the sector prior to ATLAS?

2. As a result of ATLAS training what goods or services are being delivered to target populations and what is the nature and quality of the goods and services delivered?

Development Impact Evaluation

1. What role does the sectoral system play in achieving impact at the target group level?

2. What target group impacts can be attributed to sectoral level performance?

Target Group Level Evaluation

Target group level evaluation occurs at the formative, summative, program impact and development impact stages of the evaluation cycle. Data sources are representatives of the Mission's intended beneficiaries and/or organizations that represent the beneficiaries.

Formative Evaluation

1. Describe the involvement of the beneficiary group in planning for training. How extensive was the involvement and what roles did beneficiary group representatives play in the planning process?

2. Describe how beneficiary group representatives participate in training? How do selection criteria support the involvement of target group representatives?

3. Describe how beneficiary groups are involved in training implementation as resources to trainees. Describe how changes at the target group level are communicated to trainees during training.

Summative Evaluation

1. Describe how beneficiaries or their representatives are involved in trainees' post-training employment.

Program Impact Evaluation

1. Describe the capacity of the target population to obtain and use the goods or services that are being provided by the system of which the trainee is a part.

2. Describe if/how the target population is using those goods and services.

Development Impact Evaluation

1. Measure the impact of Mission interventions on target groups.

2. Relate development impact to training interventions and the achievement of established preconditions for impact.