

METHODOLOGIES FOR ASSESSING THE IMPACT OF  
AGRICULTURAL AND RURAL DEVELOPMENT PROJECTS:  
A DIALOGUE

PROGRAM DESIGN AND EVALUATION METHODOLOGY REPORT NO. 11

by

Krishna Kumar, Senior Analyst  
(Center for Development Information and Evaluation, A.I.D.)

Agency for International Development

June 1989

The views and interpretations expressed in this paper are those of the author and should not be attributed to the Agency for International Development.

## TABLE\_OF\_CONTENTS

	Page
Foreword .....	v
Acknowledgments .....	vi
Glossary .....	vii
1. Major Findings and Conclusions of the Workshop .....	1
1.1 Conceptual and Methodological Issues .....	1
1.2 Assessing Impact on Agricultural Production ...	2
1.3 Assessing Impact on Household Income .....	3
1.4 Assessing Impact on Food Consumption and Nutrition .....	4
1.5 Assessing Impact on Natural Resources and the Environment .....	5
1.6 Organizational Issues .....	6
2. Background Information on Methodologies for Assessing Impacts .....	6
2.1 Conceptualizing Impacts .....	7
2.2 Core Impact Areas .....	10
2.3 Establishing Causal Relationships .....	12
2.4 A Realistic Approach to Causality .....	14
2.4.1 Determining Change .....	15
2.4.2 Relating Observed Changes to the Intervention .....	17
2.5 Rapid, Low-Cost Data Collection Methods .....	19
2.5.1 Indicator-Focused Sample Surveys .....	19
2.5.2 Informal Surveys .....	20
2.5.3 Community/Village Interviews .....	21
2.5.4 Focus Group Discussions .....	22
2.5.5 Interviews With Key Informants .....	23
2.5.6 Direct Observation .....	23
Appendixes	
A. Summary of Proceedings	
B. List of Participants	
Bibliography	

## FOREWORD

In an environment of shrinking resources and some skepticism about the effectiveness of aid, the U.S. Congress, PVOs, and the concerned public are pressing A.I.D. for information on the overall impacts of its development assistance. They are no longer satisfied with the data on financial disbursements and physical implementation of projects and programs, but want to know what difference, if any, these interventions have made in the lives of people and the development of the countries involved. To put it in evaluation jargon, they are asking for "impact assessments."

Impact assessments of development initiatives pose major conceptual and methodological issues that are yet to be satisfactorily resolved. The questions arise: What should be the precise focus of impact assessments? What specific impacts should be examined? Should we study both the intended and unintended impacts? What are cost-effective strategies for gathering and analyzing impact data? What type of methodological rigor can be realistically expected?

To discuss these issues in an informal setting, the Center for Development Information and Evaluation organized a two-day seminar entitled Methodologies for Assessing the Impact of Agricultural and Rural Development Assistance in January 1988. The seminar, which was attended by experts from international donor agencies and research institutions, came to a set of interesting findings and conclusions that are of wider interest to the development community.

This report includes the background paper, a summary of discussions held, and the major findings and conclusions of the seminar. I believe that the report will be useful to A.I.D. staff, contractors, and host country officials, who are now grappling with the problems of impact assessments, particularly in agricultural and rural development sectors.

Janet Ballantyne  
Associate Assistant Administrator  
Bureau for Program and Policy  
Coordination  
Center for Development Information  
and Evaluation  
Agency for International Development  
June 1989



## ACKNOWLEDGMENTS

The author wishes to acknowledge his gratitude to the various participants in the seminar who commented on the background paper and gave their views and suggestions, which are reported in this paper. He is grateful to W. Haven North and Annette Binnendijk for their comments on its earlier version. Finally, he is thankful to Allen G. Turner and Ernest Carter for their assistance in preparing Appendix A.

## GLOSSARY

### List\_of\_Terms

The following standard definitions of some of the terms used in this paper are taken from "Methods and Procedures in A.I.D. Evaluation" (OECD 1986).

- baseline study      The analysis describing the situation prior to the provision of aid, which is used to determine the results and accomplishments of a development activity and which serves as an important reference for the ex post evaluation.
- evaluation          An examination, as systematic and objective as possible, of an ongoing or completed project or program, its design, implementation, and results, with the aim of determining its efficiency, effectiveness, impact, sustainability, and the relevance of its objectives. The purpose of an evaluation is to guide decision-makers.
- ex post evaluation      An evaluation of an intervention after it has been completed. Its purpose is to study how well the aid served its purposes and to draw conclusions for similar interventions in the future.
- impact              A term indicating whether the project or program has had an effect on its surroundings in terms of technical, economic, sociocultural, institutional, and environmental factors.
- indicator          A measure used to demonstrate a change or another result of a project or program.
- input                The means by which aid is provided; the set of means (sources and activities) to be mobilized to produce the output.

GLOSSARY (cont.)

- monitoring      A management function that uses a methodical collection of data to determine whether the material and financial resources are sufficient, whether the people in charge have the necessary technical and personal qualifications, whether the activities conform to work plans, and whether the work plan has been achieved and has produced the intended results.
- outputs      The results of the aid, that is, the set of concrete results to be produced through sound management of the inputs; intermediate results necessary to achieve the purpose; or goods and services produced or directly controlled by program personnel.
- sustainability      The extent to which the objectives of an aid activity will continue to be met after project assistance has terminated; the extent to which the groups affected by the aid want to and can take charge of activities in order to continue to achieve their results.
- target group      A group of people intended to be affected positively by an aid activity. Beneficiaries are those who are affected positively by an aid activity.

List\_of\_Acronyms

- A.I.D.      U.S. Agency for International Development
- CDIE      Center for Development Information and Evaluation (of A.I.D.)
- IFAD      International Fund for Agricultural Development
- PPC      Bureau for Program and Policy Coordination (of A.I.D.)

## 1. MAJOR\_FINDINGS\_AND\_CONCLUSIONS\_OF\_THE\_WORKSHOP

On January 21, 1988, the Agency for International Development (A.I.D.) sponsored a 1-day workshop to discuss methodologies for assessing the impact of agricultural and rural development assistance. The workshop's 19 participants--representing A.I.D., academia, and international organizations--examined concepts and general methodological issues associated with impact assessment and methods for assessing impacts on agricultural production, household income, food consumption and nutrition, and natural resources and the environment. Participants also discussed organizational issues associated with the need for and management of impact data. (See Appendixes A and B for the summary of proceedings and list of participants for the workshop).

Below are the major findings and conclusions of this workshop, presented by topic area.

### 1.1 Conceptual\_and\_Methodological\_Issues

It is unrealistic to expect evaluators to try to compute the net impacts of agricultural and rural development assistance. The two research strategies for computing net impacts, quasi-experimental designs and the use of rigorous statistical controls, have not proved practical in agricultural and rural development projects. Both require extensive data collection over extended time periods, which is generally not feasible in the setting of developing countries because of various institutional and technical constraints. Moreover, the cost of data collection tends to be exceedingly high and cannot be justified given the Agency for International Development's (A.I.D.) budgetary constraints.

Therefore, instead of trying to measure the net impacts of agricultural interventions, A.I.D. evaluators should seek to determine only whether project activities were a contributing factor to the observed change. This requires answering the following two questions: (1) Have anticipated or unanticipated changes occurred in core impact areas? (2) Can these changes be plausibly related to the intervention itself?

Whenever possible, the first question should be answered on the basis of quantitative data, whereas the second should be answered on the basis of qualitative studies and theoretical evidence. Together, the answers to these questions are sufficient to determine anticipated and unanticipated impacts of an intervention.



Several rapid, low-cost data collection approaches exist that can provide useful information for project monitoring and evaluation. These include key informant interviews, community meetings, focus group discussions, field observation, informal surveys, and small sample surveys. In many cases, such approaches can substitute for more expensive, rigorous methods to save resources and time. Efforts should be made to determine the validity and reliability of the findings reached by rapid, low-cost methods by comparing them with those gathered through more rigorous methods.

## 1.2 Assessing\_Impact\_on\_Agricultural\_Production

Changes in agricultural yields are frequently used to assess the effectiveness of project interventions. However, recent studies show that the widely used crop-cutting method for estimating crop yields can result in serious measurement errors ranging from 10 to 30 percent. Moreover, the method is highly resource intensive, as it involves keeping a large number of enumerators in the field. Efforts should be made to find low-cost alternatives to this method that could then be applied in different agricultural systems.

Farmers' production estimates at harvest time can be quite accurate, or at least as accurate as those based on the crop-cutting method. Reference was made in the workshop to a five-nation study conducted by the World Bank that found farmers' estimates to be within a range of 10 percent measurement error and thus to be a relatively accurate measurement of total production (Verma, Merchant, and Scott 1988). Such estimates cannot be used for measuring yields, however, particularly when farmers have small fields or fragmented holdings. More studies are needed comparing farmers' subjective estimates with objective measurements in a variety of projects before farmers' estimates can be treated as a substitute for more objective methods.

Agricultural yields can also be estimated through the use of plant-process models. Such models can bridge the gulf between yields on research stations and yields on farmers' plots by incorporating a variety of basic data on variables such as soil, rainfall, inputs, and cultural practices. The reliability of the findings of this model depends on the quality of available basic data.

The emphasis on target crops, which is common in most agricultural projects, generally results in not enough attention being given to other agricultural products, such as nontarget crops, livestock, fisheries, and edible wildlife. The problem



with this approach is that while the production of target crops may increase, the project can adversely affect the production of other agricultural products. Therefore, evaluators should go beyond examining the impacts of a project on targeted crops and try to answer the following questions: What crops or other productive activities suffered as a result of an increased emphasis on the target crop? How were these products used? Who was responsible for their production and marketing? What was the opportunity cost of the new crop?

Finally, the impact of an intervention on agricultural production can be affected by gender stratification. Female farmers play a very important role in agricultural production worldwide, so evaluators need to focus on project impact on agricultural production in farms managed by both male and female farmers.

### 1.3 Assessing\_Impact\_on\_Household\_Income

Although accurate measurement of household income for non-wage earning groups is difficult, it becomes almost impossible for rural households in developing countries for several reasons. Rural households have access to nonpurchased goods (i.e., homemade items and plants and animals gathered or hunted by family members) and unpaid family labor. The seasonality of agricultural production and therefore of farm income and fluctuations in the prices of the goods produced create additional problems. Because of these problems and the difficulty of getting truthful answers from reluctant respondents, income survey data tend to be unreliable. If gathered, such data should be supplemented with information collected from other sources.

Evaluators can assess general trends in household income without expending considerable resources by conducting small sample surveys for extended periods of time. Well-designed surveys with a sample size ranging between 80 and 200 cases can provide sufficient information for making valid statistical inferences. The data from such surveys can be aggregated at program and regional levels. Such surveys should focus on household expenditures and a few standard-of-living (assets) indicators. These indicators can be easily identified by evaluating the socioeconomic setting of the project. Although the choice of specific indicators will vary by region, they should address items such as the nature of housing, household furnishings, sanitation, and nutritional status.

In assessing the impact of development projects on income, attention should also be given to the intrahousehold dimension. Issues of who earns and controls income are important for mea-

asuring overall project impacts and should be carefully examined.

It is easier to assess the income effects of a program (which comprises several projects extending over longer time periods) than of a single project. In the case of a program, time-series data on standard-of-living indicators for a small sample of households, along with macro-level data routinely gathered by governments, are usually sufficient to reveal underlying trends. More intensive data collection activities are required for assessing the impact at the project level.

#### 1.4 Assessing\_Impact\_on\_Food\_Consumption\_and\_Nutrition

Increased production alone cannot meet food consumption objectives. Several factors, such as seasonality, crop mix, income, role of women, crop labor requirements, market, and government policies, affect the linkages between production and consumption. Therefore production itself is not always a good indicator of food consumption, much less an indicator of nutritional status. A variety of indicators (e.g., food availability, per capita calorie consumption, food expenditure, amount and type of food consumed, and nutritional status) can be used for assessing the impact of a project, depending on the nature and scope of the project.

Anthropometric indicators for measuring nutritional status are useful in assessing project impact. They are simple to use, and the needed data can be gathered with modest resources. Enumerators can be trained within a short period, and a nutritional module can be added to surveys undertaken for other purposes, thereby reducing overall costs. Moreover, evidence indicates that even short-term changes in food availability and income are reflected in anthropometric data on children. The obvious limitation of such indicators is that they measure change in nutritional status but they do not explain the process by which it occurs.

Questions about the frequency of meals consumed during different agricultural seasons, the number of nights when the respondent goes to bed hungry, and the respondent's knowledge of the number of people who go to bed hungry can also give a reasonable indication of the extent of food consumption and hunger among target populations. However, the final selection of indicators, whether indirect or proxy, will require an in-depth understanding of local food consumption patterns. Moreover, findings of surveys that use such questions should be complemented with information gathered from other sources.

## 1.5 Assessing\_Impact\_on\_Natural\_Resources\_and\_the\_Environment

Some of the problems of assessing the effects of development projects on the natural resource base are different from those of measuring impacts on production, income, or food consumption. The unit of account of the project is generally not congruent with the unit of account in the ecosystem. The nature of natural resources dictates a need for a broad, integrated approach; comprehensive baseline information is needed at various levels to make valid generalizations about the impacts of agricultural projects and programs. And evaluations must be conducted over longer time horizons to adequately capture changes.

The concept of natural resource systems can provide a basis for developing suitable models for assessing site-specific and regional impacts. Successful models have already been developed to monitor some aspects of the natural resource base, including erosion rates, water quality, rangeland conditions, and maximum sustainable yields. Still, it is important to recognize that agreement on model building is lacking within the scientific community and donor agencies; a more concerted effort is needed to achieve consensus on these issues. Accurate models should be constructed with reference to the properties of different ecosystems and the factors that affect them. Such models should focus not only on physical and biological dimensions of the problem, but on social and economic aspects as well.

Although considerable information about the natural resource base already exists, there are several reasons why it has not been fully utilized for designing and evaluating agricultural projects. One is that the information is not presented in a form that is readily understandable and useful to policymakers. The information is generally written in technical language. Moreover, because no attempt is made to explain and quantify the links between natural resources and productivity, policymakers, who have to examine the issues in terms of costs and benefits of different policy options, are unable to utilize it. Another factor is the lack of agreement on assessment standards for environmental impacts. International donor agencies are often not in a position to impose their own standards on developing countries. Faced with the urgent need for increased agricultural production, host countries tend to overlook the harmful effects of certain project activities on their natural resource base, which emerge only after a long time. Finally, evaluators have little incentive to use the information that is available because they feel that they can do little about environmental issues anyway.

## 1.6 Organizational\_Issues

Stakeholders in projects or programming are or should be interested in the findings of impact evaluations. Stakeholders include project management staff, host governments, A.I.D./Washington, and the U.S. Congress. The information needs of these stakeholders are not always identical. While impact assessments should be designed to meet the needs of all stakeholders, priority should be given to the needs of project management and host governments. This requires that host governments act as equal partners in the implementation and management of impact evaluations.

Although project managers should have access to impact data, the responsibility for gathering the data should not usually be entrusted to them for several reasons. Managers remain interested parties and thus cannot be expected to be totally objective. Moreover, they are fully occupied with implementation problems, and the management of data collection efforts adds to their already heavy burdens. Above all, many long-term impacts are not visible during the life span of most projects. Hence, an organization that will outlive the project is needed to collect and analyze data. The ideal organizational unit to undertake this responsibility is the line agency in which the project is located.

Efforts should be made to further develop and strengthen institutional capabilities of host country agencies to conduct impact assessments of agricultural projects and programs. This will require providing them with technical assistance and monetary resources on a long-term basis.

## 2. BACKGROUND\_INFORMATION\_ON METHODOLOGIES\_FOR\_ASSESSING\_IMPACTS<sup>1</sup>

This section provides a framework for a discussion of the methodologies for assessing the impact of agricultural projects. It seeks to clarify the concept of impact, examines the question of causality, and outlines practical strategies for impact assessments. It also lists various data collection options that can provide relevant information with modest investments of time and resources.

---

<sup>1</sup>The information contained in this section was provided to workshop participants as background material to stimulate their

discussion of impact assessment issues.

The information presented is designed to raise issues for discussion. Even when certain positions have been taken, the purpose is to stimulate thinking on the subject rather than to provide definitive answers. Obviously, if there were definitive answers to the issues raised, there would be little justification for discussing them in the workshop.

Two limitations of the material presented here should be mentioned. First, it focuses primarily on projects that are designed to increase agricultural production and incomes of farmers and excludes interventions whose stated objectives are different (e.g., improvements in nutrition or prevention of environmental degradation). Second, although the issues discussed have direct relevance to the assessment of the impacts of policy reform and structural adjustment assistance, the paper does not directly focus on them.

## 2.1 Conceptualizing Impacts

Broadly speaking, three categories of outcomes of agricultural interventions can be distinguished.

The first and probably the most widely understood category includes the outputs of the activities undertaken by a project or program. For example, the 250 input-supply depots supported by a project distribute 1,000 tons of fertilizer every agricultural season, or the 2,000 electric water pumps installed under an irrigation scheme supply water to 5,000 small farms. Such outputs are examined for monitoring and mid-term evaluations.

The second category consists of the effects of such outputs on production. The outputs are not an end in themselves. Rather, they are expected to increase the production and productivity of the existing farming systems. Thus fertilizer or water is supplied so that farmers can increase their yields or grow new crops. Such effects are known as *purpose-level effects* in the terminology of A.I.D.'s Logical Framework.

The last category of outcomes consists of the impacts of an agricultural project on individuals, institutions, society, and the environment. These impacts are also called *goal-level effects*. In most, although not all, cases these impacts are generated primarily as a result of the changes in production systems initiated by an intervention. Undoubtedly, at the individual level the most important and obvious effect is on the economic welfare of farmers as measured by changes in their income, standard of living, or nutrition. Examples of institutional-level impacts are the expansion of free markets or the

growth of participatory institutions that articulate the

interests of the farming populations. Examples of impacts on society are changes in the social stratification system or in the values and beliefs of the people. Environmental impacts are the impacts of projects or programs on soil, water, forests, and other natural resources.

In this discussion, the word "impact" is used to refer to the last two categories of outcomes. A middle position has been taken by excluding outputs but including their effects on agricultural production.

A few observations should be made about the nature of the impact of agricultural projects/programs that bears on the conduct of impact assessments. First, agricultural development projects tend to have impacts that were not envisaged at the design stage. Despite their best efforts, designers of development initiatives are not always able to anticipate their major outcomes, because of a paucity of relevant information or faulty judgment. This is particularly the case with innovative projects (i.e., the underlying intervention models are untried) and projects that are being introduced into different socioeconomic settings than in the past. Since unanticipated effects can be as important as anticipated ones, both are examined in impact assessments.

Second, impacts can be both positive and negative. Almost all projects or programs have some outcomes that are viewed as undesirable by a segment of the population. Because agricultural interventions are designed to change the status quo, they do not affect all groups in the same way. Some groups benefit more, others benefit less, and still others lose. Moreover, agricultural interventions affect the physical environment. To provide a balanced picture, an impact assessment should describe both positive and negative results of an intervention.

Third, the relationship between an intervention and its presumed outcomes tends to become more and more distant as one moves from outputs to purpose- and goal-level impacts. Consequently, the attribution of causality becomes progressively more problematic.

An example can illustrate this point. Suppose an agricultural services project supplies fertilizer to smallholders. The availability of X tons of fertilizer to farmers is thus the output of the project, and the causal relationship between the project and the availability of fertilizer is obvious and indisputable. However, when evaluators try to examine the effects of the project on yields, the situation changes. The availability of fertilizer does not mean that farmers are using it, much less that fertilizer use has contributed to greater yields. Yields

are affected by myriad factors in addition to fertilizer, such

as climate, prices, irrigation facilities, farmers' risk aversion, and the land tenure system. Project planners examine all such factors and make assumptions about them. Therefore, positing a causal relationship between the project and changes in yield becomes problematic. And the situation becomes still more uncertain when evaluators attempt to measure the project's impact on farmers' incomes, because several variables that can affect farmers' income will also have to be considered. For example, commodity prices, costs of inputs, procurement policies of the host government, and taxes may affect the level of income gains resulting from increased yields. Thus the list of factors that must be examined in assessing impact becomes larger as one moves from output to goal-level impacts.

Finally, large-scale agricultural interventions, particularly the successful ones, tend to have far-reaching consequences for the socioeconomic systems of the region and even the country. Because the various elements of a society--economic, social, political, and cultural--are interrelated, interventions have "multiplier effects" that are not easy to conceptualize or operationalize in a rigorous fashion.

The widespread adoption of a high-yielding variety of wheat by farmers in Asia--a phenomenon known as the Green Revolution--provides a good example. The effects of the adoption process did not end with the increased yields of wheat and greater incomes for innovative farmers. Adoption also contributed to a basic transformation of existing agrarian systems in the affected regions, including changes in cropping systems, greater demand for consumer goods and services, growth of agroindustrial enterprises, labor shortages during peak seasons, in-migration, new and differentiated roles for women, expansion of education, increased economic disparities, and the growing influence of the rural elites in political systems.<sup>2</sup>

Thus, the obvious implication is that impact assessments cannot examine all the changes that can be associated with a major intervention. Rather, impact assessments should focus on a few, carefully selected impact areas that are of critical importance from the point of view of the main objectives of an intervention.

---

<sup>2</sup>It is interesting to note that these impacts of the Green Revolution were identified largely by academic scholars pursuing independent studies of economic and social change in the regions. Early evaluations, which focused on performance, generally failed to systematically examine these impacts.



## 2.2 Core\_Impact\_Areas

The substantive focus of impact assessments is bound to differ according to the nature, objectives, and performance of the project or program; the availability of data and information; and the disciplinary background of evaluators. An analysis of impact evaluations conducted by A.I.D. indicates that they have had remarkable differences in focus.<sup>3</sup> Even evaluations of projects or programs conducted in the same subsector (e.g., irrigation or integrated rural development) did not always focus on the same impact areas.

While flexibility in the selection of substantive focus is necessary to capture the uniqueness of each intervention, there is an imperative need to focus on a set of core impact areas and to use a common set of indicators in order to build a body of empirically grounded knowledge. By focusing impact assessments on core areas, the quality of comparative analysis and the validity of the findings can be improved. Moreover, this will lead to more empirically grounded "synthesis reviews" of assessment findings, conclusions, and recommendations in different subsectors.

The following are some of the core impact areas stressed by international donor agencies.

1. **Production\_and\_productivity.** At the purpose level, the most important impact of an intervention is on agricultural production and productivity. Although the two are not the same, they are closely related in subsistence agriculture, where major increases in production are not possible without substantial gains in productivity. Thus, depending on the nature of the project, the effects on yields, livestock production, fisheries, and forest products can be studied.

Developmental initiatives also indirectly affect production. For example, increased agricultural production stimulates the production of other goods and services through forward and

---

<sup>3</sup>A.I.D.'s Center for Development Information and Evaluation (CDIE) has conducted impact evaluations of a wide range of agricultural and rural development projects and has prepared synthesis reviews. Some of the subsectors covered are agricultural research, agricultural services, irrigation, rural electrification, integrated rural development, and institutions of higher agricultural education.



backward linkages. Therefore, both the direct and indirect effects of an intervention should be examined. In some instances, this examination will require an assessment of the impact at a regional level.

2. Income/standard\_of\_living. The impact of an intervention on income is also an important area to be considered. After all, the goal of all agricultural interventions is to increase the income of the rural populace and thereby improve their standard of living. When direct measurement of income is not feasible, standard-of-living indicators can be used to measure the effects on income. Often, the distributional aspect of income benefits might also be studied to determine whether the income benefits are also reaching vulnerable groups--women farmers, landless laborers, and ethnic minorities.

3. Food\_consumption. Although successful agricultural projects and programs tend to have positive effects on food consumption, some groups may be adversely affected by an intervention. An overemphasis on cash crops, particularly for export, can lead to the neglect of traditional food crops.<sup>4</sup> Income earned from the sale of cash crops may not be spent on food for the family. Land reclamation projects have occasionally deprived local tribal groups of edible wildlife that existed on the land before the land was cleared for agriculture. It is therefore desirable that the impacts on food consumption also be examined.

4. Environment. Increasing attention is being given to the importance of assessing the environmental effects of agricultural interventions. Even well-conceived interventions can adversely affect soil quality, water resources, forests, biological diversity, and human settlements, thereby undermining the prospects for long-term development. Conversely, many agricultural interventions can contribute to an improvement in the natural resource base while increasing production and productivity. Agroforestry projects supported by A.I.D. are good examples.

For assessing the impact on these four core areas, three general criteria should be kept in mind. The first is coverage. For example, what proportion of target farmers increased their incomes by raising their production and productivity?

---

4A.I.D.'s Office of Policy Development and Program Review recently commissioned several studies for the impact of agricultural exports on food production, consumption, and nutrition. These studies, directed by Dr. Joachim von Braum, are being conducted by the International Food Policy Research

Institute.

Obviously, a project can hardly be classified as a success if only a small proportion of the target populations benefited by it. The second criterion is sustainability, which means that the benefits produced by an intervention should continue after donor assistance ends. If the benefits are not sustainable, they will not make a meaningful difference in the conditions of farmers. In the past, evaluators have often ignored the issue of sustainability. Finally, the criterion of economic efficiency is very significant. Simply put, it means that the project should have a positive rate of return on the investment.

### 2.3 Establishing\_Causal\_Relationships

The critical methodological issue in impact assessment is that of establishing the causal relationship between an intervention and the observed changes. This is the question every impact assessment grapples with, although with limited if any success. Evaluators, despite their best efforts, often do not succeed in proving the elusive causal relationship with the methodological rigor normally expected in social and economic research. And the reasons are not difficult to discern.<sup>5</sup>

To establish a cause and effect relationship in a rigorous manner, evaluators must be able to isolate the effects of an intervention from those of confounding factors in order to measure "net impacts" rather than "gross impacts." Two research strategies can be followed for this purpose: quasi-experimental designs or the use of statistical controls.

Quasi-experimental designs generally require that a control group be selected by random sampling. When random sampling is not possible, as is the case with most aid projects and programs, evaluators can identify/select a matching group on the basis of common attributes and environmental conditions. After the data have been gathered for several time periods for both groups, the changes for both groups should be compared and net impact measured by subtracting the changes in the control group from those in the project target group. For example, if an impact assessment of an agricultural services project finds that the annual rate of growth in cereal production was 5 percent in the project area and only 2 percent in the control region, evaluators would conclude that the project had contributed a 3-percent annual increase in cereal production.

---

<sup>5</sup>For a detailed discussion of these issues see Casley and Kumar (1987, especially chapters 7-9).



Quasi-experimental designs are extremely difficult to use in impact assessment for several reasons. The identification or construction of the control group for an intervention is generally difficult and often virtually impossible. Moreover, even if a valid control group is selected, it cannot be assumed that no developmental initiatives affecting this group will be undertaken for the duration of the study simply because evaluators are trying to measure the impacts of a project. Rather, there is a high probability that because the control group did not receive the needed assistance earlier, it may receive it during the study period. If this happens, the whole effort will be wasted.

In addition, quasi-experimental designs require extensive data. The collection, coding, and analysis of large longitudinal data sets in a developing country setting pose innumerable technical and managerial problems that should not be underestimated. Moreover, such data tend to have a high incidence of nonsampling errors that undermine the validity of findings and conclusions. Finally, reliance on such designs can result in misleading conclusions in many cases. For example, if the magnitude of change in the project is not large, it might not be discernible at all.

The second strategy for establishing causal relationships--using statistical controls--is also not very practical. This strategy involves the collection of data for relevant independent, intervening, and dependent variables for several points in time. The net impacts are computed by using sophisticated statistical techniques such as multivariate regression, analysis of variance, or factor analysis. Thus this strategy also requires extensive data collection.

One major drawback of this strategy is the time-series data needed to meet the statistical requirements. In many instances, the period for which data are needed is so long that the method is not feasible. This can be illustrated by a simple example. Consider a project designed to increase cereal production by 6 percent annually in a rain-fed agricultural system with an estimated variability in annual production of about 15 percent. In this case, to determine at an 85-percent confidence level that cereal production has increased 6 percent annually, production data for 8 years would be needed. And for a higher level of confidence, say 95 percent, the number of years for which data are needed would increase to 11 (see Casley and Kumar 1987, 118). In any case, the impact assessment will require time-series data covering a longer period than the usual 6- or 7-year life of an agricultural project. The truth is that the time-series data for undertaking a rigorous statistical analysis are usually not available.

The two strategies discussed above are based on the premise that impact assessments were planned and data-gathering efforts were initiated at the inception of a project, which is an unrealistic assumption. Experience has shown that only a few projects or programs provide for systematic data collection for impact assessment. Even when plans are made, they are often not implemented because of the lack of interest, if not the outright hostility, of management; the difficulty of finding people or institutions with the technical skills needed to gather and analyze the data; and the paucity of resources. And in the rare instance when such plans are implemented, the data generated tend to have serious errors that undermine their utility for impact assessment.<sup>6</sup> Whatever the reasons, the majority of impact assessments have been undertaken (and are likely to continue to be undertaken) in the absence of carefully collected time-series data, which means that the rigorous strategies discussed above are not very realistic or practical for most impact assessment efforts.

Finally, the problem of resource constraints should not be overlooked. These two strategies for establishing causal relationships involve a considerable investment of resources. Thus the value of using these strategies is questionable except for a few highly innovative pilot projects whose outcomes are uncertain<sup>7</sup> or for large interventions for which the proportional cost of research and data collection is likely to be low.

Under typical project conditions, expectations should be realistic. Instead of aiming at establishing precise causal relationships by computing net impacts, evaluators should be content with less rigorous but more feasible strategies that can provide a good indication of the impact of an intervention.

#### 2.4 A Realistic Approach to Causality

In assessing the impact of agricultural interventions, evaluators should examine the following two questions individually:

---

<sup>6</sup>This widely held view was confirmed by a recent analysis of A.I.D. experience on this subject (Norton and Benoliel 1987).

<sup>7</sup>The ideal course for conducting rigorous evaluations of innovative projects is to entrust the responsibility to reputed research institutions at the beginning of an intervention.



- Have anticipated or unanticipated changes occurred in the impact areas under examination?
- Can these changes be plausibly related to the intervention?

While these questions are interrelated, answering them may require different types of data and evidence.

#### 2.4.1 Determining\_Change

To determine whether changes have occurred in relevant impact areas, data are required on pre- and post-intervention status and usually on the intervening period. While such data should give an indication of the nature and magnitude of change, the data need not be precise or elaborate.

Several simple methodological strategies have been used successfully to assess changes in impact areas and even in their magnitude. Three of these strategies are mentioned below to illustrate their use.

One strategy is to gather time-series data on a few key impact indicators that were carefully selected at the beginning of the project or program. The main criteria for selecting these indicators are that they be sensitive to project-induced changes and that they can be measured without the need for elaborate, complex data collection efforts. Just because an indicator seems appropriate (e.g., income or calorie consumption) does not mean that the necessary data will be easy to collect.

The data for impact indicators are usually gathered through baseline and follow-up surveys. A good example of such surveys is provided by Malawi, where systematic data on the outputs and impacts of area development projects are gathered throughout the country using identical questionnaires. As a result, excellent time-series data are available for Malawi that can provide information and insights about the impacts of area development initiatives. Unfortunately, however, because of technical and institutional constraints, data analysis has been inadequate.

This strategy requires that the data-gathering process commence early in order to discern changes, which, as indicated earlier, is not a realistic expectation. Another problem is that indicators measure changes only in impact areas that were anticipated at the design stage. It is quite possible, however, for many critical areas to be overlooked at this stage and thus not be covered by the information system. For example, in early

dairy projects, designers did not anticipate that commercialization might have adverse effects on the nutritional status of

farm families by tempting farmers to sell their total milk production, thereby depriving their families of much-needed protein. Had impact assessments been conducted primarily on the basis of previously selected indicators, this consequence might have been overlooked. Thus indicator data alone are not sufficient for impact assessments.

A second strategy is to conduct what may be called "retrospective surveys," particularly of the target populations. In such surveys, respondents are asked to compare pre- and post-intervention conditions with reference to relevant variables. For the impact assessment of an integrated rural development project, for example, farmers could be asked about their yields, major crops, livestock, income, expenditures, possessions, housing conditions, and the like for the pre-project stage as well as the present. The reported differences between present and past conditions, if any, provide evidence for changes that have occurred in the condition of target populations or geographical areas since the implementation of the project.

In some instances, retrospective surveys can be designed to seek the views of experts and other knowledgeable individuals rather than of target populations. For example, to assess the impact of a project on agricultural production, surveys can be conducted of extension workers, village traders, staff of the agricultural procurement agency, village officials, farm economists, agronomists, and even a few well-informed farmers. The responses of such informants tend to be more knowledgeable than those of the target populations, and provide a more balanced perspective.

The limitations of this strategy are quite obvious. The long recall period contributes to the low reliability of data. Farmers forget and therefore give inaccurate information. Moreover, they have a tendency to minimize the improvements that have occurred over time. Despite these shortcomings, the findings of such surveys can provide a reasonable indication of change. Moreover, retrospective surveys illuminate the perspectives and perceptions of farmers and thus give an insight into the subjective dimension of change.

A third strategy is to conduct cross-sectional surveys that generate comparative data on farmers, farm households, or the area covered by the project or program as well as on those not covered by it. For example, in an agricultural credit project, yields, incomes, or living conditions of the farmers who received credit from the project can be compared with those of farmers who did not receive credit. Or economic and social conditions in a region with an area development project can be compared with conditions in a similar region without such a project. For



analytical purposes, the present conditions of the control group are used as a proxy for the pre-project conditions of the treatment group. Unlike quasi-experimental designs, this strategy involves data collection at only one point in time--at the time of impact assessment. Thus data collection efforts are relatively simple and less elaborate.

One limitation of quasi-experimental designs afflicts this strategy as well. As pointed out earlier, the identification and construction of control groups is exceedingly difficult in development settings. Groups that look comparable might not be. For example, the control group farmers who did not receive loans from the project might have had easy access to credit from other sources and therefore did not need the credit offered through the project. In this case, using these farmers as a control group might obscure the benefits to the farmers who received credit from the project. Extreme care is therefore necessary in designing such surveys and in drawing inferences from them.

In any case, the three strategies just described should be used in conjunction with one another and with other sources of information. Often, secondary data are available that can shed some light on changes in the impact areas under examination. For example, in countries in which governments procure major agricultural crops, procurement statistics give a fair indication of production trends, provided the share of crops kept for domestic consumption and sold in informal markets remains constant. In other cases, studies conducted for altogether different purposes can be used to establish benchmarks for comparison. In an impact evaluation conducted by A.I.D. in Egypt, evaluators used old photographs of the communities in estimating the changes that had occurred since the intervention (Johnson et al. 1983). And when quantitative data cannot be generated, qualitative methods can be used to identify the nature of change.

The essential point is that it is often possible to obtain a reasonable indication of change--and even its magnitude-- without engaging in large-scale data collection. Once the nature (and possibly the magnitude) of change has been determined, the next step is to examine its relationship to the relevant intervention.

#### 2.4.2 Relating\_Observed\_Changes\_to\_the\_Intervention

The second question (Can the change be related to the intervention?) is not difficult to answer if the objective is to determine whether the intervention has contributed to the change



rather than to measure its net impact. A two-pronged strategy that relies on experience (theoretical or empirical) and qualitative studies can be used for this purpose.

It should be recognized that the intervention models underlying agricultural projects are usually based on both theoretical and empirical evidence. Therefore, in many cases an evaluator can be reasonably confident after examining the outputs of an intervention, that specific project activities have also contributed to the observed change. For example, it is known that the adoption of high-yielding varieties of maize seed, together with the use of the required inputs, results in increased production. Hence, if the data show that maize production has increased in the project areas, it can be confidently assumed that the project was responsible for at least some of the increase.

In this case, evaluators can go a step further. If it is found that the price of maize and the relative costs of production have remained constant, they can also infer that the increased production was responsible for higher incomes for farmers who are using the high-yielding variety of seed. In fact, by using appropriate models, incremental gains in income for various types of farm households can be predicted with reasonable accuracy.

The reliance on theoretical or empirical evidence is not sufficient, however. In many instances, wrong inferences can be drawn because the underlying assumptions are invalid or because other factors are responsible for the observed change. For example, in the case described above, the procurement prices for major crops might have been raised by the government under pressure from international agencies, and it might have been the higher prices that motivated farmers to increase their maize production. It is therefore important to undertake qualitative studies that focus on key assumptions or alternative explanations.<sup>8</sup> These studies will further improve the validity of the findings and conclusions.

This can be illustrated by a simple example. Suppose the data show that maize production and farm incomes have risen in the project area after the introduction of the new variety of maize. The issue to be resolved is whether the project has been a major contributing factor to these increases. Two steps might be taken to explore this question. First, evaluators would review experience with similar projects in the country or

---

<sup>8</sup>Such studies are akin to the "Follow-up Diagnostic Studies" described in Casley and Kumar (1987, Chapter 6).



elsewhere. In the case discussed above, such a review may indicate a strong relationship between the adoption of the high-yielding variety of maize and increased production and income. These findings provide strong evidence that the project was a major contributing factor to the increases. The second step is to examine the validity of the assumptions underlying the posited relationship. This can be done through qualitative interviews with key informants, focus group discussions, group interviews, or informal surveys. Evaluators can, for example, find out from key informants whether there were other factors that might explain the increased production and income of farmers. If such factors are identified, their relative importance will have to be considered in making necessary inferences and drawing conclusions.

If the two-phased approach outlined above is followed in impact assessments, the issue of causality can be satisfactorily resolved. Although the resulting findings might not be as rigorously derived as methodological purists would like, the findings would be valid and reliable and serve the needs of the development community. In any case, the findings would be more valid than those generally produced through the impact assessments conducted by international donor agencies or host countries.

## 2.5 Rapid, Low-Cost Data Collection Methods

There are several rapid, low-cost data collection methods for gathering information that can be used to support the impact assessment strategies described in Section 2.4.9. Although these methods do not involve rigorous procedures for data collection and analysis, they can deliver relatively accurate information on a wide range of subjects. These methods are mentioned here not necessarily as alternatives to more rigorous methods (e.g., comprehensive sample surveys, agricultural censuses, field experiments, or even participant observation) but as complementary to them in many instances.

### 2.5.1 Indicator-Focused Sample Surveys

Indicator-focused sample surveys differ from conventional sample surveys in three respects. First, they focus on a few carefully selected impact indicators. Consequently, the number

---

<sup>9</sup>These methods are described more fully in Kumar (1987b).



of questions in the questionnaire does not exceed 15 to 20. In this respect, they are akin to public opinion polls. Second, although probability sampling is used, the sample size is kept to a bare minimum by reducing the number of control variables. Third, enumerators are permitted greater flexibility in conducting interviews. They are allowed to ask questions that are not included in the questionnaire if they believe the questions would elicit useful, additional information or help to clarify an issue.

Such surveys tend to focus on two types of indicators. The first are community/village-level indicators that describe local conditions. The unit of analysis for such indicators is the community/village rather than the farm household, which considerably reduces the cost of data collection. Simple indicators, such as the percentage of houses in a village with concrete/tin roofs or with access (within 2 miles) to an elementary school, or the number of tractors in the village, can be constructed. These can give an indication of the economic status of the community/village and enable any change to be discerned over time. The second category of indicators includes micro-level indicators derived from household data. Examples are volume or monetary value of major crops, dairy products, or fisheries; per capita household income or expenditures; type of dwellings; and household possessions. Such indicators provide more precise information than do community/village-level indicators.

Indicator-focused surveys are particularly useful when quantitative evidence about the magnitude of change is required. With careful preparation, they can generate timely information with a modest investment of resources. Their only limitation is that the data generated cannot be subjected to complex statistical treatment because of the small number of variables covered and the small sample size.

### 2.5.2 Informal Surveys

Farming systems specialists have developed informal surveys to increase the understanding of key characteristics of farming systems in targeted areas. Several variants of such surveys have been used. In one variant known as "Sondeo," a team of 14 scientists is sent into the field for a week. Each day, team members break into seven pairs and conduct field observations and interviews. Each evening, they compare their findings and explore the implications. Another version of the informal agricultural survey relies on a multidisciplinary team to jointly conduct a field survey on the basis of an interview guide. This procedure permits the free flow of ideas between the team and respondents. The interviews are conducted in informal settings



in individual and group sessions. Still another variant is the "exploratory survey," which uses a relatively more structured and comprehensive questionnaire. The common elements among these informal surveys are their use of informal procedures for gathering data and their combination of field observation and interviews, which provides a fuller picture of the situation.

The methodology of informal surveys can be used both for discerning changes in impact areas and for exploring the relationship of the changes to the aid intervention. Since informal surveys also involve direct observation, the data they generate are more accurate than those generated by structured sample surveys. That they are largely qualitative does not mean that they cannot be used for comparative analysis. In fact, because of their qualitative nature, they often permit more insightful comparisons over time or across different targeted zones than do structured surveys. Some of the data gathered through informal surveys can be quantified and used for refining and improving economic models designed to predict the outcomes of investments in agricultural initiatives.

One limitation of informal surveys is that they are susceptible to investigator biases, especially those of investigators who have a strong position. A multidisciplinary team composition can reduce such biases and increase the validity of the findings.

### 2.5.3 Community/Village Interviews

Community/village interviews can be a valuable source of information and ideas for impact assessments.<sup>10</sup> They take the form of public meetings open to all the members of a village or community. To get reliable data, the meetings are conducted on the basis of an interview guide that lists the questions to be asked. An advantage of a carefully prepared guide is that relevant questions can be phrased in simple language that participants can understand and respond to.

Community/village interviews can even generate quantitative data, which are usually of two types: (1) community-level statistics, best gathered by using a predesigned form to enter information provided by the various participants in a meeting, or (2) quantifiable data about the behaviors, perceptions, and assessments of the participants, which is gathered by tallying

---

<sup>10</sup>Community interviews and focus group discussions are described in detail in Kumar (1987a).



the "yes" and "no" responses to questions on specific topics. The second type of data could be biased if the respondents are not representative of the target populations about which generalizations are to be made.

To avoid bias, it is important that the communities/villages selected be representative of the total population. The technique of quota or expert sampling can be used to help ensure a representative selection of communities/villages for impact assessment.

Such interviews are very useful in impact assessments, especially for gaining a beneficiary perspective. One of their merits is that the participants in a meeting tend to correct each other, thereby improving the quality and accuracy of information. Moreover, community/village interviews are relatively inexpensive to conduct. Experience in developing countries indicates that meetings called for this purpose are usually well attended if they are publicized in advance. On the negative side, these meetings are easily manipulated by government officials, project and program staff, and local leaders, and evaluators should not rule out this possibility.

#### 2.5.4 Focus\_Group\_Discussions

Focus group discussions, which have proved extremely useful in the design and evaluation of health and family planning projects, can also be used for impact assessments of agricultural interventions.

In focus groups, participants discuss ideas, issues, insights, and experiences among themselves. A group is limited to 6 to 10 carefully selected individuals who share a common background. The moderator keeps the discussion focused by subtle probing techniques and prevents a few participants from dominating the discussions. The value of focus groups lies in the fact that in a small group situation, many intricate issues can be explored in depth. The group situation often reduces individual inhibitions and enables participants to freely express themselves.

Focus groups can be used to identify project or program impact areas, examine the relationship between an intervention and the observed changes in the impact areas, and explore various explanations for the observed change. For example, for an assessment of the impact of an agroforestry project, several focus groups can be organized to discuss how the project has affected, if at all, soil, water management, yields, and farmer

incomes. Separate focus group discussions will have to be held with farmers and experts familiar with local conditions.

As a method of social research, focus group discussions are quite economical and in many cases can provide an in-depth understanding that might not otherwise be gained. The participants in a group stimulate each other, shining new light on the subject. Although much depends on the skills of the moderator, often the sum is greater than the individual contributions.

#### 2.5.5 Interviews\_With\_Key\_Informants

Simply stated, key informant interviews involve interviewing a selected group of individuals who are likely to provide the needed information, ideas, and insights on a particular topic. Such interviews are conducted using a guide that lists the main topics to be covered in a session. The interviewer frames the question in the course of discussions and uses probing techniques to focus the discussion.

Key informants should be carefully selected to reflect different types of expertise, background, and experience. The best course is to identify, according to the nature of the project, appropriate sources from which the key informants can be drawn and then to select a few from each group. Thus, for example, if evaluators are examining the impact of an agroforestry project, they would prepare a list of groups that are likely to be most knowledgeable on the subject--project farmers, soil scientists, agronomists, horticulturalists, economists, and government officials--and select and interview a few from each category. Other people identified during the interviews who may possess relevant information and ideas may also be interviewed.

Although key informant interviews can be used for assessing the magnitude of change, they are most effective for examining the relationship between agricultural projects and observed changes. The key informants can explain how and why the specific intervention has or has not affected the impact areas on which an impact assessment is focusing. They can, for example, tell whether maize production and farmer incomes have increased as a result of the wider availability to local traders of a hybrid variety of maize seed promoted by the project. Or they can explain why the increased maize production has not resulted in greater incomes to the farmers in the project. In many instances, key informant interviews can be conducted to gain an idea of the nature and extent of changes in impact areas.

#### 2.5.6 Direct\_Observation

Direct observation involves the careful gathering of data on the basis of well-designed instruments that can range from observation record forms to the special tools required for test-

ing the quality of soil or water. In many instances, it also involves interviewing people in the field. Data gathering through direct observation is systematic, not casual or informal.

Direct observation should be conducted by a team of experts rather than by a single individual. A team approach provides a more comprehensive picture and helps prevent individual biases from distorting the findings.

Properly kept records of the direct observations conducted at different points in time during and after a project or program can indicate the nature and magnitude of change. Moreover, visits to project and nonproject sites can give evaluators a sense of the changes that might be attributable to specific interventions.

In conclusion, it should be stressed that realism rather than an obsessive concern for methodological rigor should inform impact assessments. Evaluators should tap all possible sources of information and ideas and analyze them with objectivity.

## APPENDIX\_A

### SUMMARY\_OF\_PROCEEDINGS

#### 1. INAUGURAL\_SESSION

In opening remarks at the January 1988 workshop, Richard Bissell, Assistant Administrator of the Agency for International Development's (A.I.D.) Bureau for Program and Policy Coordination (PPC), noted that A.I.D.'s programs in agriculture, rural development, and nutrition represent the core of its development efforts. And yet, the Agency has not succeeded in succinctly describing the objectives of these programs to its own staff and outside constituencies. However, the recently drafted focus statement, which defines the goals of these programs in terms of achieving growth in income, increased food consumption, and the preservation of the natural resource base, represents a significant advance in articulating overall Agency objectives.

Bissell observed that the Agency largely relies on Country Development Strategy Statements, Action Plans, project financial data, and a few studies to describe its achievements. These sources, although useful, do not provide a full or accurate picture of the effects of development assistance; they describe project performance and not impacts. There is growing pressure from the U.S. Congress, advocacy groups, and A.I.D. constituencies to demonstrate ultimate impacts--achievements that the Agency can demonstrate in improving the living conditions of people and contributing to the development of poor countries.

Bissell stressed two areas that should be specifically examined with respect to impact assessments. First is the issue of sustainability, which refers to the continuation of development activities originally supported by the project once donor assistance ends. In the past, impact evaluations have overlooked this element despite overwhelming evidence that many projects could not be sustained over time and thus produced limited impacts. Second is the issue of environmental impact. Assessment of impacts on the environment has become an increasingly important issue as both donor and host agencies have come to recognize that preservation of the natural resource base is essential for long-term development. The achievement of short-term gains in production at the expense of the environment has increasingly come to be viewed as poor economics.

W. Haven North, Associate Assistant Administrator of PPC's Center for Development Information and Evaluation (CDIE), spoke next, mentioning that although PPC/CDIE has conducted impact

evaluations for some time and in the process has produced a

large number of insightful reports, it is still struggling to find a better way to assess and describe impacts in the agricultural and rural development sectors. Often, evaluation teams sent abroad return feeling that they were unable to assess the precise effects of interventions, much less plausibly attribute them to the Agency's efforts. Complicating efforts to assess impacts and attribute outcomes to specific Agency interventions are such factors as lack of consensus on what impacts should be examined, absence of time-series data, differences in the disciplinary orientations of evaluators, inability to gather current information, and, perhaps, inadequate analysis.

North posed three major questions for the participants to examine in depth. First, what should be the central focus of impact evaluations? Since it is not reasonable to try to examine all possible effects and consequences of an intervention, it is necessary to sharpen the focus of impact evaluations. North then listed four possible areas--agricultural production, income, food consumption and nutrition, and the environment--that might constitute the core focus of impact evaluations. Second, what kind of evidence is required to determine whether A.I.D.'s efforts have produced the desired results? While it is not possible to establish precise causal relationships between interventions and observed changes, evaluators still need to generate credible data and information about these changes. Simple anecdotal evidence is not sufficient. Finally, are there cost-effective alternatives to large-scale, data collection efforts? Given that the budget constraints experienced today are likely to continue, the Agency cannot afford to invest heavily in data collection and analysis for impact evaluations.

## 2. CONCEPTS\_AND\_METHODOLOGIES

### 2.1 Introduction

Krishna Kumar of PPC/CDIE initiated the discussion on several methodological issues relevant to assessing the impacts of agricultural projects. He made three important points.

First, evaluators' tendency to focus on different impact areas, depending on the nature of the project, the data available, and their own disciplinary orientation, has come in the way of synthesizing the findings of various evaluations. Thus, a coherent body of knowledge about the impacts of agricultural interventions in different socioeconomic contexts is lacking. While some differences in focus among different evaluations are unavoidable, and even necessary, assessments should focus on a

set of standardized core areas common to all evaluations. Kumar

suggested that agricultural production, household income, food consumption and nutrition, and the natural resource base can provide a useful framework for project impact assessment.

Second, Kumar suggested that it is unrealistic to expect evaluators to try to compute the net impacts of agricultural and rural development assistance. The two research strategies for computing net impacts, quasi-experimental designs and the use of statistical controls, have proved impractical for agricultural and rural development projects. Both require extensive data collection over extended time periods, which is generally not feasible in rural settings due to institutional and technical constraints. Moreover, the cost of such data collection tends to be exceedingly high and is usually beyond the capability of most projects.

In Kumar's view, evaluators should not try to measure the net impacts of agricultural interventions, but instead should be content to simply determine whether project activities have contributed to the observed changes. This approach requires answering the following two questions: (1) Have anticipated or unanticipated changes occurred in the impact areas under examination? (2) Can these changes be plausibly related to the intervention? Whenever possible, the first question should be answered on the basis of quantitative data, whereas the second should be answered on the basis of qualitative studies and theoretical evidence. Together, the answers to these questions are sufficient to examine the effects of an intervention, whether anticipated or unanticipated.

Finally, Kumar mentioned that several rapid, low-cost data collection approaches can provide useful information for project monitoring and evaluation. These include key informant interviews, community meetings, focus group discussions, field observation, informal surveys, and small sample surveys. In many cases, these approaches can substitute for more expensive, rigorous methods to save resources and time. Efforts should be made to determine the validity and reliability of findings reached by low-cost methods by comparing them with those arrived at through more thorough and rigorous methods.

## 2.2 Discussion

Participants generally agreed that assessments cannot focus on all impact areas. This is especially true for major agricultural and rural development interventions, which tend to produce far-reaching economic, social, and cultural effects. Evaluators need to sharply focus impact evaluations on a few core areas.

There was general consensus among participants that the four

areas that had been identified--income, agricultural production, food availability and consumption, and the natural resource base--provided a common focus.

Participants identified and elaborated on several problems that confront evaluators. First, many intervention impacts, particularly in the income and natural resource areas, do not become visible during a project's life of 5 to 7 years. Second, many project impacts are not confined to the boundaries of a particular target area or sector. By limiting the scope of an investigation to a geographical area or sector, evaluations will overlook other important and far-reaching effects outside the project area or in other sectors. Third, if precise impact indicators are identified during the design phase of a project, unanticipated effects may remain unanalyzed. This is especially a problem in situations where baseline data are used to assess future change. Fourth, in many cases, several projects are operating in the same area or sector. In such cases, the effects of one project cannot be separated from the effects of others for evaluation purposes. For example, one participant described research in Zambia and Zimbabwe, where social and economic changes had been closely monitored over several decades. Although it was possible from among a multitude of donors to identify projects that achieved greater progress than others, the changes recorded over this period could not be attributed to individual projects. Finally, because projects operate within dynamic political and economic settings, many factors external to the project's setting also influence outcome.

A cross-cutting issue stressed by several participants was the problem of impact sustainability. In addition to assessing whether interventions have produced their desired impacts, evaluations should also examine whether impacts are sustainable in the long run. Too often, well-designed projects have generated positive impacts, but these impacts could not be sustained after external funding ended.

Participants agreed that, in general, evaluators must combine informal, qualitative studies with quantitative measures to infer causal relationships. To establish whether change has occurred, evaluators should rely on quantitative measures. To establish linkages between interventions and observed change, qualitative studies are needed. Furthermore, one must understand the processes underlying observed changes. Longitudinal studies are not the only source of useful information on this topic; short-term studies may also be used to expand on the information regarding a certain trend.

In many cases, evaluators can draw inferences from examples of linkages that have been well demonstrated in similar situa-

tions elsewhere. However, one should be careful not to assume

too much. For example, in Asia, the impact of fertilizer on yields is well documented, and in many cases can be quantified; in Africa, such relationships are less clear.

Several participants stressed that quantitative data are necessary for national planners and decision-makers who need to compare the impacts of different projects and make decisions about their funding. Although all participants agreed with the need to compare impacts of projects implemented in different social and economic settings, they were not sure that the same set of indicators could be uniformly applied to all or even a majority of such projects. It was also mentioned that the need to quantify findings often forces evaluators to measure project impacts in monetary terms. While this ensures comparability, it can deprive evaluations of insightful information about the nature of impacts and the factors that affect them. Moreover, some of the assumptions underlying such analyses remain problematic. Some participants suggested that the overall impacts of different projects can be compared even when quantitative information is not available.

Participants agreed that there are valid low-cost ways of gathering comparative impact data that are useful to project management, agency planners, and funding decision-makers. They noted several characteristics of low-cost methods that are sometimes overlooked. For example, low-cost approaches can include quantitative measures. Statistically sound quantitative assessments can be made using a modest sample size of 100 or 200 individuals, as long as there is no need for disaggregation by sex, age, or other factors. In this connection it was suggested that one cost-effective way of collecting useful impact data was through adding relevant modules to an ongoing survey.

It was observed that what was important in many cases was to obtain information quickly. For example, managers need to know quickly whether an intervention is on track or is having a negative impact. In such cases, low-cost methods can prove extremely fruitful.

Participants pointed out that sometimes generalizations can be made on the basis of a few cases. For example, the World Bank analyzed its training and extension methodology in one state in India and generalized these results to evaluate the training and extension system in other states. To assess the utility and validity of a wide range of low-cost impact measures, donors must evaluate in-depth at least some of their thousands of projects by establishing rigorous experimental situations.

As noted earlier, both quantitative and qualitative information can be useful in impact evaluation. One participant

observed that the use of quantitative data alone may not be

desirable. For example, the International Potato Institute found that its evaluations, which included quotes from many knowledgeable trainees and observers, were viewed as credible, because people knew that quantitative data could be manipulated so easily. Thus, in many instances, anecdotal evidence and the views and comments of experts more effectively convey the message than does quantitative data.

### 3. ASSESSING\_IMPACT\_ON\_AGRICULTURAL\_PRODUCTION

#### 3.1 Introduction

Dennis Casley of the Monitoring and Evaluation Division of the World Bank opened this session by reviewing the problems encountered using traditional approaches for measuring the impact of agricultural projects on crop production.

He noted that a variety of problems face evaluators in estimating crop production at the project level. Even under relatively stable conditions, such as irrigated agriculture in Asia, production estimates can show an annual variation of 15 to 20 percent. This level of variation is itself much greater than the 5-percent increase in production that projects are generally expected to achieve. Thus, if an appropriate probability sample is used to measure the effects of a project, it might require collecting time-series data for 10 or more years. This is undoubtedly not feasible in most cases.

Casley pointed out several problems with the crop-cutting approach, a traditional crop estimation method widely used in Asia and Africa, which involves weighing samples of dried crops taken from randomly selected subplots. These subplots are selected from identified fields in accordance with specified procedures. The crop-cutting method is very expensive and involves the participation of many enumerators in the field. Moreover, it generally results in biases or measurement errors of 10-15 percent and sometimes as high as 30-35 percent. Thus, despite its wide acceptance, this method does not provide reliable enough results.

As an alternative, Casley advocated the use of farmers' production estimates, which can be gathered at a fraction of the cost of the crop-cutting method. For example, one can ask farmers about their actual or even expected production. Rigorous checking of farmers' estimates has shown that such estimates usually result in measurement errors of 10 percent or less, which can be minimized if evaluators can identify farmers who are

skilled in crop estimating. Most farmers will not deliber-

ately mislead investigators; in cases where this does happen, it is often due to political factors and generally results in readily detectable clusters of anomalous data. Data collected from farmers' production estimates should be supplemented with case studies on selected farmers. In these instances, the whole field should be measured instead of taking a limited sample of cuttings.

Casley emphasized two additional points. First, it is important to go beyond assessments of impacts on target crops to include non-target crops, livestock, fisheries, and edible wildlife. This is important because interventions designed to increase target crops may adversely affect the production of other foods important to improved diets. Second, it is important not to exclude women producers who may be affected differently from their male counterparts.

### 3.2 Discussion

Participants generally agreed on the usefulness of farmers' estimates of their production. It was also generally agreed that in many countries officials rely largely on estimates from commercial as well as small farmers, an approach that has produced relatively accurate production estimates. A case study of cocoa production in Ghana was mentioned as supporting this observation.

Participants noted that although farmers' estimates are useful in measuring total quantity produced, they are not as reliable in assessing yields per unit of land. This is especially true when farmers cultivate small, irregularly shaped fields or fragmented holdings. For example, in estimating size of holdings, Sudanese farmers who hired laborers and farmers who worked their own land did not provide equally accurate figures. Other factors can also distort estimates.

It was suggested that yield increases could be estimated using plant-process models that incorporate a variety of data. Such models can bridge the gulf between farmers' yields and research station yields by incorporating data on variables such as soil, rainfall, inputs, and cultural practices. The results can be cross-checked by using information from research station plots and estimates from farmers. Again, the point was made that accurate results require that substantial prior information be combined with reasonably complete and reliable current data to establish standards, to indicate the direction of causality, and to provide basic monitoring. In the United States, attempts are underway to calibrate a system that links plant process, root

zone, groundwater, surface water, and other variables to

cropping practices. These types of models could be adapted to other countries fairly easily, because they are driven by basic information on soil type, weather, and crops and residue associated with cultivation practices.

Some participants pointed out that the emphasis on target crops common to most agricultural projects causes problems because it generally ignores impacts on other forms of agricultural production, such as non-target crops, livestock, forests, and fisheries. For example, it is important to take into consideration the availability and access to wild edible plants and animals, which can contribute significantly to local diets. This highlights the need to assess impacts on natural resource conservation.

Moreover, farmers' perspectives on the importance of target crops may vary. The farm household may be affected in a variety of ways by a project's emphasis on a few "key" target crops. Such impacts may be negative or positive depending on the individuals concerned, whether they are women, children, landlords, or tenants. To fully assess project impacts, evaluators must probe beyond issues directly related to the target crop. They should try to answer such questions as: What crops or other productive activities were pushed aside by an increased emphasis on the target crop? What were the goods ignored by the project used for? Who was responsible for producing and marketing these goods? What resources were no longer available for producing other goods?

Some participants brought up the issue of project impacts on livestock. They suggested that although some interventions (e.g., vaccinations) are easy to describe, their impacts on production are difficult to assess. Conventional measures like herd off-take (the rate at which animals are removed for sale or slaughter) are not very useful because of the multitude of intervening factors, such as financial stress, forage availability, market demand and associated price changes, and so forth. Off-take may be best applied in conjunction with the concept of critical herd size. If a given farm household or production unit's herd is above the critical size, then off-take can be used to measure cash income. If, on the other hand, it is below critical size, then off-take may be an estimate of financial stress--a need for cash beyond the unit's capability to meet that need.

Similarly, the emphasis on production impacts should not distract evaluators from looking at the effect of project interventions on other related factors such as timing, storage, post-harvest loss, quality, perishability, and marketing.



Finally, participants thought that it was important to evaluate agricultural production impacts on farms managed by both male and female farmers. For example, women farmers in Africa play very important and often leading roles in household agricultural production. In many cases, men and women respond differently to survey questions concerning the farm production unit.

#### 4. ASSESSING IMPACT ON HOUSEHOLD INCOME

##### 4.1 Introduction

Thayer Scudder of the California Institute of Technology began the discussion on assessing impact on household income by elaborating on several points. First, he mentioned that it is generally difficult to get accurate data using surveys on household income. Evaluators find it difficult to accurately measure the value of nonpurchased goods, unpaid family labor, and income flow variations through time. Respondents often find it difficult to answer questions on income truthfully. Scudder stressed that assessments that focus on income alone have proven insufficient and must be supplemented with and linked to information on the standard of living (assets) and household expenditures. Evaluators must assess what income is used for and what the developmental implications of that use are.

Second, Scudder pointed out that small, low-cost surveys can capture good information on the standard of living and expenditures of the target population. Information on standard of living is especially easy to collect because assets are visible. Identifying trends over longer time periods can be accomplished without employing extensive resources by reducing the number of items studied and by conducting surveys of limited intensity. A surprisingly small sample that focuses on a limited number of items is all that is required to adequately monitor and evaluate project impact on household income.

Only two elements are required for an appropriate longitudinal study that can help evaluators understand impact trends: (1) periodic assessments of income, standard of living, and expenditures and (2) establishment of procedures that will sustain the monitoring efforts. Citing the Mahaweli project in Sri Lanka, Scudder mentioned that periodic surveys that sample a total of 30 households have provided assessments that are 80 percent accurate over a project area encompassing one million people. Such surveys have required only 3 weeks of fieldwork with reports released just 3 weeks after the completion of

fieldwork. These surveys covered items such as quality of housing, types of household furnishings, levels of sanitation, and nutritional status.

Third, Scudder mentioned that standardized questionnaires focusing on a set of standard-of-living indicators can be used for a variety of countries or project settings without requiring extensive changes. Despite vast cultural and economic differences among rural populations, the same indicators are frequently applicable in different countries. The underlying reason is that throughout the world, people who are moving from subsistence standards of living to a position where they enjoy modest surpluses tend to make similar choices on how to spend their cash. They generally invest in improved housing, new furnishings for their house, education for their children, and other nonfarm activities.

Finally, Scudder mentioned that it is easier to assess the income effects of a program than those of a single project. In the case of a program, time-series data for a small sample of households on standard-of-living impacts, coupled with macro-level data routinely collected by governments, are usually sufficient to reveal underlying trends. More intensive data are required to assess project-level impact.

## 4.2 Discussion

Participants began by discussing definitions of the terms "household" and "income." Some participants thought that certain difficulties arise in identifying the household unit. Usually, one can assume a one-to-one relationship between a household and a land holding. However, the point was made that in some instances it is easier to define the farm production unit than the household. Furthermore, some participants suggested that it is not as simple to define a household with respect to consumption as it is with respect to production. Still, most participants agreed that it is possible to define and use the concept of household if it is agreed that a reasonable amount of ambiguity is acceptable. The underlying notion of a group of people living together, sharing food, and pooling their income for the common good is generally sufficient for research purposes. When assessing impacts at the project level, evaluators should not become overly concerned with definitions; once a unit is defined the definition need only be used consistently to yield valid information. (The difficulty comes at the national level, when comparisons of data are made that are based on different definitions of "households.")



Participants discussed alternative definitions and interpretations of "income." It was suggested that the income objective of development projects might best be defined as "moving people beyond subsistence by increasing their purchasing power." Participants stressed that the concept of total income has both monetary and non-monetary components; often, there is a mistaken tendency to look only at cash purchasing power.

Several participants stressed that approaches focusing solely on production and income are limited and can be misleading. One cannot always assume a direct link between increased production and improved living standards. For example, in the case of the Mahaweli project in Sri Lanka, both rice production and farmer incomes rose; however, farmers had to use most of their incremental income to service new debt. Thus, their standard of living did not improve.

Participants also noted that cash expenditures on productive or consumer goods may not always be a good proxy indicator for changes in income in a project area because changes in such expenditures may be the result of external factors. For example, in the Mahaweli project, many tractors and other equipment were purchased. However, this equipment was purchased by settlers who had worked as contractors during the project's construction phase. Thus, increased cash expenditures were not a signal for increased income resulting from greater cash crop farming.

It was argued that standard-of-living data may not be useful for comparative assessment. In this context, the World Bank's more intensive Living Standards Measurement project, as carried out in Peru and Cote d'Ivoire, provides a useful alternative. Although a very long questionnaire was used, data were collected and analyzed quickly, and reports were prepared in a matter of weeks. This approach provides a means of rapid assessment, albeit at considerable expense.

Participants identified several key points in carrying out household surveys. When an income survey is undertaken, evaluators must look at each source of income and each earner. They may begin with data on the household head, but should go on to develop a list of household members that includes every member's age, sex, and primary activity. To assess income effects, evaluators should examine such factors as the flow of income, the value of farm products produced, the value of each household member's time, implicit wage rates, and so forth. Each question must be thoroughly probed. It is extremely important to understand intrahousehold dynamics, especially as they relate to gender issues. Rather than treat the household as a monolithic unit, special attention should be focused on the often differing roles men and women play in earning and controlling income.



Participants stressed the point that project managers do not need precise quantitative data about income. Often they can make realistic assessments of change by using qualitative information gathered through informal data collection techniques. These managers live in the area and are familiar with observable changes in living conditions. However, program managers need data and information that can be aggregated and are comparable across different cases. Therefore, the selection of indicators and data sources is likely to vary depending on the needs of different users.

## 5. ASSESSING\_IMPACT\_ON\_FOOD\_CONSUMPTION\_AND\_NUTRITION

### 5.1 Introduction

Tim Frankenburger of the Office of Arid Land Studies of the University of Arizona began by discussing the extent to which food consumption and nutrition are related to other impact areas and noted several broad issue areas that clarify food production and consumption linkages. Frankenburger stressed that increasing production will not ensure that consumption objectives are met. He observed that in Malawi, for example, a country that has successfully increased agricultural production, the incidence of child malnutrition is far higher than it is in many other African countries that have achieved less dramatic results in agricultural production. He also questioned whether increased food supplies and increased income are necessarily compatible project objectives.

Frankenburger highlighted several key issue areas that are important to consider in terms of the degree to which food production and consumption can be linked: seasonality (impact on malnutrition, human energy, and labor demands); crop mix (impact of minor crops and declines in crop diversity); income (impact on payments in cash and kind); role of women (impact of changes in cooking patterns and time devoted to breast-feeding or child care); crop labor requirements (impact on food distribution within a household); market prices (impact of levels and stability); and the role of government (impact of exchange rate, and import and export policies). In all cases, examples may be cited in which increased income and greater food production will not necessarily result in greater food consumption and improved nutritional levels. Indeed, it is possible to identify situations that suggest inverse relationships.

Frankenburger noted that the choice of whether to use direct or proxy indicators will be determined in part by the end use of

the data. The indicator chosen will define, or be

defined by, the concept of food consumption chosen. Proxy indicators for food consumption include food availability, food expenditures, amount of food eaten, the proportional mix of different food groups consumed, or nutritional status as determined by height and weight measurements. The final selection of indicators, whether direct or proxy, will require a thorough knowledge of local cultures.

Finally, Frankenburger described recent work in Mauritania that found the following indicators useful in assessing project impacts on food availability, amount of food consumed, and nutritional status: food in storage right before harvest, which provides insight into the degree of household vulnerability; near-liquid assets in the form of, for example, small livestock that can easily be sold; a subsistence potential ratio, which is derived by converting production potential into calories required; and a dietary complexity ratio, which measures the intake frequency of particular food items over a given time period.

## 5.2 Discussion

Participants acknowledged that assessments that accurately measure impacts on individual well-being must go beyond the concept of food production to include impacts on food consumption and nutrition. Participants reviewed several indicators and discussed their advantages and limitations. Several low-cost field techniques for gathering data on food consumption were discussed. One participant noted that when measuring food consumption at the household level, food volume (not weight) is an easy to use and reliable measure. It was suggested that frequency questions would also provide useful results; for example: How many meals do you eat per day? How many people went hungry in the village this week? Also, given adequate knowledge about a specific cultural context, one can draw useful conclusions from the frequency of use of specific foods added to the diet.

Some participants voiced strong reservations about the utility of these indicators for measuring food consumption. One participant thought that impact evaluations can make little use of frequency indicators for the consumption of particular foods and that such data cannot be satisfyingly correlated with specific project interventions; thus, management decisions could not be based on such findings. Another participant mentioned that he did not find food consumption information useful in measuring the impact of agricultural and rural development projects. Rather, he relied on income data, despite difficulties in col-

lecting it. Nonetheless, some participants strongly maintained that the most significant assessment of project impact on individual well-being were based on food consumption data.

Discussion then shifted to the usefulness of income and food expenditure data in evaluating impact on food consumption and nutrition. If household income is used as a proxy for food consumption, for example, it is important to identify who receives and uses that income. The results of a survey in Guatemala demonstrated that the proportion of income spent on food for the family was closely associated with the amount the wife, rather than the husband, was able to earn outside the household. With respect to food expenditure surveys, it was acknowledged that they can be very useful but costly. At the same time, national consumer expenditure surveys disaggregated by region are frequently available and often contain food expenditure data. These surveys may identify kinds and frequencies of food purchases. Costs can be significantly reduced by using existing databases to provide baseline information or cross-sectional comparison data.

Much of the discussion that followed addressed the relative advantages and disadvantages of consumption data compared with data on nutritional status. Nutritional data were seen as a proxy indicator for measuring impact on food consumption, in part because such data are easily collected and aggregated. However, as noted by Frankenburger, the collection of consumption-oriented data through a low-cost, add-on component to food production data collection efforts already underway is more likely to be acceptable to those conducting the survey than is the collection of data on nutritional status. People oriented toward agricultural production are unlikely to collect data on nutritional status, but will sometimes recognize the value of collecting food consumption data.

Many participants, however, thought that anthropometric measurements as an indicator of nutritional status may be even more useful than consumption data in assessing impact on individual well-being. However, such measurements must be used with some care. For example, a change in the nutritional status of young children may provide unreliable evidence of an agricultural project's impact, because it can be influenced by factors external to the project, such as sanitation conditions, activity levels, and education. Still, for any given population, anthropometric indicators (specifically, age-height and height-weight ratios) will provide considerable information about short-term changes for a very modest investment. One participant noted that in the case of relocation projects impact is successfully measured almost exclusively using anthropometric indicators.



Most participants agreed that no single approach focusing exclusively on food consumption or nutritional status will provide satisfactory results. Measures that attempt to capture changes in food consumption and nutritional status should be combined.

## 6. ASSESSING\_IMPACT\_ON\_NATURAL\_RESOURCES\_AND\_THE\_ENVIRONMENT

### 6.1 Introduction

Tom Gilbert of the Graduate Program of Ecology of the University of Tennessee initiated the discussion on assessing environmental impacts by calling attention to a portion of the focus statement that refers to the need to maintain and enhance the natural resource base. Some of the problems associated with assessing the effects of development projects on natural resources are different from those associated with measuring impacts on income and food availability and consumption. In particular, natural resource impact assessments require a highly integrated systems approach that incorporates numerous elements; comprehensive baseline data, and longer time horizons to capture change. For example, unsound agricultural practices hundreds of miles inland can cause significant ecological damage to coastal swamp areas and marine fisheries. Here, the interplay of a broad range of elements over a large geographical area produces specific outcomes. Adequate monitoring and evaluation of such impacts requires substantial baseline information and a reasonable understanding of the natural processes involved and their linkages.

Gilbert suggested that the concept of natural resource systems can provide a basis for developing models to assess site-specific and regional impacts. Such models can be developed based on the types of ecosystems that are affected and the key indicators of ecosystem conditions that can be identified. The characteristics and scale of a project provide an additional list of variables that can be incorporated into a model. Gilbert provided an example of a project and the types of information that would be needed to assess site-specific impacts. Suppose a project is located in a tropical forest and that local inhabitants gather wood for their fuel and housing needs and forage for various plants and animals to supplement their diets. To assess change in the natural habitat, general data on physical, biological, and social aspects of the setting must be collected at appropriate intervals. Site-specific information would also be needed on natural resource users and how their actions affect wood supply, wild plant and animal stocks, and water supply.

Knowledge of traditional resource management schemes can provide extremely useful insights into the dynamic relationship between human communities and the environment.

Gilbert stressed the need for increased international cooperation between donor agencies, host governments, and international organizations, for both management and information-sharing purposes. He described three programs that could be of benefit to A.I.D. First, Gilbert mentioned the United Nations Education, Social, and Cultural Organization's (UNESCO) Man and Biosphere (MAB) program, which could help A.I.D. develop methodologies and train staff to conduct inventories, to construct baseline studies, and to derive models useful for assessing impacts on natural resources. Current MAB program objectives include (1) conducting studies of ecosystems that are experiencing varying types and intensities of human impact and (2) analyzing different investment choices and development schemes as they relate to natural resources. The program is now conducting 266 biosphere assessments in 70 countries that will provide useful baseline data on a variety of ecosystems. Second, Gilbert cited the Conservation Data Center of the International Union for the Conservation of Nature and Natural Resources as another useful program for development agencies. Finally, Gilbert mentioned the Expanded Information Base project: Integrating Natural Resource Management and Economic Development, a cooperative A.I.D./National Park Service effort as a useful source of information for assessments of changes in the natural resource base.

## 6.2 Discussion

Participants agreed that evaluating natural resource aspects of projects presents several challenges. First, the task is difficult because the unit of account at the project level is not comparable to the unit of account in the ecosystem. For example, project activities in the upper watershed create impacts downstream, as has occurred in the sedimentation of reservoirs in Morocco. Second, linkages between project interventions and environmental impacts are difficult to anticipate. Thus, the need to look beyond a project's stated objectives is especially critical. By understanding these linkages, project managers and evaluators can better anticipate otherwise unexpected outcomes. The example was mentioned of leaking canal water in the Indus watershed, which led not only to increased salinization (an adverse impact), but also to aquifer recharge, making possible increased use of groundwater through tubewells (a positive, but unexpected, outcome).



Some participants gave examples of successful models that were developed to monitor many aspects of the natural resource base, including erosion rates, water quality, impact on fisheries, and maximum sustained yield. Others noted that some models are increasingly capable of using remotely sensed data. Participants described an erosion-productivity calculating model developed from a plant-process model and relevant components from other models. It was stressed that plant-process models used in the United States can be adapted to developing countries fairly easily. Such models are driven by basic information on soil type, weather, and crops and residue associated with different cultivation practices. Finally, the comment was made that models should often incorporate social and institutional factors into the analysis. All too often, the behavior of local populations is left out of the planning process, especially their behavior with regard to the environment and natural resource use.

Participants stressed that cooperation between host countries and donor agencies is necessary, especially for projects that are likely to have broader impacts. Most current efforts to assess environmental impact are fragmented and suffer from a lack of long-term commitment. Participants mentioned that examples of well-planned cooperative efforts do exist. Several noted the Endangered Resources for Development project in Kenya, where A.I.D. joined with UNESCO, the United Nations Environmental Programs (UNEP), and the International Board of Plant Genetic Resources to assist the Kenyan Government in gathering data on threatened natural resources and developing policies and strategies to conserve important ecosystems. These efforts resulted in the Kenyan Government's improved understanding of alternative policy options, as well as improved baseline information on a comprehensive set of natural resources.

Some participants strongly agreed that the lack of baseline data was one of the most critical problems facing evaluators attempting to examine impacts. It was noted that quite sophisticated methods have been developed to generate data on environmental impact, but most efforts are directed toward engineering questions. Adequate guidance relevant to problems of natural resource management (e.g., as it relates to impacts on agricultural production) is usually not available. Many large gaps remain in our knowledge of what plant and animal species exist (especially in tropical environments); how they interact with each other, their physical environment, and human communities; and what measures can be followed to conserve and sustain various natural systems.

In conjunction with the concerns expressed about the lack of data, the alternative view was expressed that much information does exist but that it is not well utilized. Two reasons

for this situation were discussed. One is that information is presented in highly technical terms, rather than translated into a more usable form for policymakers and project managers. Much of the data used to evaluate natural resources impact is of limited interest to policymakers, who must assess impacts on agricultural production and the costs and benefits of different policy options.

A second reason cited is the lack of agreement on assessment standards, both within the scientific community and between donor and host country agencies, for defining tolerable levels of impact for various kinds and degrees of intervention. A lack of consensus also exists in defining effective strategies, and many models remain controversial or inappropriate for tropical and semitropical environments. There is little value in measuring change if one is unable to tie it convincingly to an overall strategy. It is difficult to discuss impact assessment when effective strategies and models for natural resource conservation and enhancement remain unclear.

While recognizing that controversy does exist with respect to strategies, model building, and assessment standards, participants acknowledged that indicators can be identified that are useful for assessing specific impacts on various elements of the natural resource base. Regional impacts can be measured using models that describe conditions of particular ecosystems (i.e., rangelands, wetlands, and upland forests) by analyzing certain variables like vegetative cover, animal populations, and habitat deterioration. Site-specific impacts can be measured using models that analyze the status of soil and water resources. Indicators are available to measure impact on soil erosion, nutrient deficiencies, and mulch management, as well as ground and surface water supply and quality.

Many participants expressed their deep concerns about A.I.D.'s involvement in environmental issues, because the standards and concerns of donor agencies in general are not necessarily in accord with host government development priorities. Host government agencies often perceive a trade-off between increased production and income on the one hand and the conservation of natural resources on the other. Many nations facing large debts do not have the luxury of worrying about longer term natural resource management issues. For example, data requirements for monitoring increasing desertification were developed for the Sudan and led to several good recommendations for action, but none of these was ever carried out by the Sudanese Government.

## 7. ORGANIZATIONAL\_ISSUES

### 7.1 Introduction

Ram Malhotra, Director of the Monitoring and Evaluation Division of the International Fund for Agricultural Development (IFAD), introduced this final discussion topic by describing IFAD's experience with monitoring and evaluation systems. He suggested two key areas of concern, one structural and the other procedural. Structural issues involve concerns about whose needs should receive priority when the system is designed and which party should be responsible for monitoring and evaluation activities. Procedural issues relate to concerns about the composition of the evaluation team and the methods of investigation to be used to keep data collection costs low and still provide the needed information. Malhotra further observed that lessons learned from evaluations are often not disseminated widely enough to influence the design of future projects and programs.

Malhotra stressed that impact findings are of interest to several groups, including project managers, host governments, A.I.D. staff located in Missions and in Washington, and the U.S. Congress. However, the information needs of different groups vary, a situation Malhotra described by using A.I.D.'s Logical Framework, which contains several levels of information, as an example. The project manager needs information on the achievement of targets, which are the "outputs" found in the Logical Framework. The host country implementation agency needs information on agricultural production or numbers of farmers visited by extension workers to measure achievement of "purpose" in the Logical Framework. Agency planners need information on who benefited, which is the "goal" found in the Logical Framework. Malhotra said that this hierarchy of information needs should be addressed in impact evaluations.

Malhotra next focused on how to determine where to locate monitoring and evaluation units. He described a two-tiered system that separates monitoring and evaluation functions. Monitoring should occur at the project level, whereas responsibility for evaluation should be located higher up within the line agency.

Another issue Malhotra discussed was the importance of institution building, and the need to develop a project impact assessment capacity within host government agencies. Host country involvement in evaluation is essential to ensuring the wider utilization of evaluation findings and recommendations.

IFAD has learned through experience of the need for good collab-

oration. When helping set up systems for monitoring and evaluation, IFAD focuses on the host country's needs and not those of the donors. Establishment of monitoring and evaluation units at the project and ministerial levels is an effective means for building host country capacity for improved project management.

Finally, Malhotra suggested that evaluation activities are best carried out by those most closely involved. And the most important methodological issue may be how to bring beneficiaries into the evaluation process.

## 7.2 Discussion

Several participants disagreed with some of Malhotra's comments. They pointed out that the hierarchy of levels implicit in A.I.D.'s Logical Framework simply does not exist in practice. The categories mentioned in the Logical Framework are analytical and not empirical. For instance, it is inaccurate to say that while A.I.D./Washington is primarily interested in goal-level issues, the main preoccupation of project managers is with problems at the input and output level.

Some participants also noted that it may be misleading to keep structural issues separate from procedural ones. Procedural issues, like those related to data collection efforts, are closely tied to structural concerns. For example, using imaginative ways to collect data will often require developing new organizational relationships. Various possibilities were noted, including encouraging schools to collect anthropometric data by providing them with scales, urging farmers to keep their own accounts, being aware of seasonal needs for data, and attaching additional data collection modules to existing efforts.

One structural issue, that of whom impact evaluations should be designed for, was raised by participants on several occasions. Several stakeholders were identified who have, or should have, an interest in the impact of A.I.D.'s projects. These include project beneficiaries, project managers, host country and donor agency officials, national planners, and program funding agencies like the U.S. Congress. Participants generally agreed that each set of stakeholders needs different information about project impacts. Thus, for example, information that is useful to the project manager may not be useful to a member of Congress.

Participants mentioned key questions that should be considered when developing an information system for assessing project

impacts. Such questions include: Who will use the information

gathered? How will they use it? How can we reconcile the various information needs of different stakeholders for any given project? What type of data will be required? How will data be gathered and analyzed? There was a general feeling that information systems are often established without careful thought or preparation.

In this connection, participants stressed that impact evaluation and monitoring functions should be separated. This is not to argue against self-evaluation, but to distinguish it clearly from impact evaluation, which must be done by an objective outside body. Monitoring involves the use of data and information that are routinely collected during project implementation for purposes of management. Although impact assessments may well be able to use much of that same information, a separate entity, not linked to the project management, must be responsible for evaluation efforts.

Two opposing views were expressed concerning the role of project managers in impact evaluation. Some participants thought that there was a great need for project managers to examine the impacts of project activities. Managers who are simply concerned with the immediate outputs of a project have too narrow a view of their duties and responsibilities. Good managers should constantly seek to understand whether their activities are contributing to the long-term goals of development assistance efforts (i.e., improvement in the overall quality of life of target groups, or the establishment of adequate natural resource conservation measures).

While some participants agreed that project managers should be and are interested in the impacts of their activities, it was suggested that they should not be concerned with systematic data collection activities that support efforts to assess impacts. Project managers have many responsibilities, and all of their energy and outstanding skills in personal relations are needed to ensure that the project runs smoothly. Especially during the project start-up phase, with many new people to account for and new relationships forming, it is difficult for the project manager to spend the time needed for data collection and analysis efforts.

Several other reasons were mentioned by participants in support of the argument that impact assessment responsibilities cannot be entrusted to project managers. First, managers do not usually possess the necessary expertise in data collection and analysis and are therefore unable to make the right decisions. Second, they have a personal stake in the project, and it is in their interest to present a positive picture of project impacts. Their ability to remain objective is always in doubt. Third,

many intervention impacts may not become visible during a

project's lifespan. Therefore, data collection efforts in many cases must continue for some time after donor assistance to a project ends.

Participants also discussed a variety of problems encountered in conducting impact evaluations. Despite statements to the contrary, it was noted that there is a lack of political support for evaluation activities. In many instances, USAID Mission staff and host country officials are not enthusiastic about becoming involved in such efforts. Moreover, time-series data are usually not available, and when they are, their quality is questionable. Impact evaluations require considerable investment of resources, which are often not forthcoming. A.I.D. staff also face many logistical problems in organizing and fielding evaluation teams overseas and providing them with necessary data collection and analysis support. There was a consensus that despite these problems, impact evaluations conducted by donor agencies have richly contributed to our understanding of the overall effects of agricultural and rural development assistance. Impact evaluation findings are reaching a wide audience of national-level policymakers and are often taken into consideration by those involved in the design of new projects and programs.

Finally, participants discussed the importance of developing host country capabilities to manage impact assessment activities. Several participants thought that donors had done rather poorly at developing capacity at the national level to support the impact evaluation and other information gathering and analysis needs of projects. In some cases, the host government unit responsible for statistical efforts has no credibility within its own agency. The result is that each donor operates independently to fund its own impact studies, which leads to wasted resources and duplication of effort. In other cases, donors have supported the establishment of national-level information systems, but just as these systems are becoming operational, donors pull out prematurely because funding for the project ends. For example, in the case of Kenya's rural roads impact monitoring system, the host country agency's efforts began to wane as donor support lessened, with the result that data collection efforts continued, but no attempts were made to analyze the data.

Participants pointed out that donor agencies have also failed to assign the task of impact assessment to local agencies that have demonstrated the capacity to undertake such responsibilities. In some cases, donor agencies have selected institutions that have not performed well in other cases, donor agencies have overlooked some very capable local governmental and nongovernmental institutions.



## APPENDIX B

### LIST\_OF\_WORKSHOP\_PARTICIPANTS

Dr. Duane Acker Agency for International Development Bureau for Science and Technology, Directorate for Food and Agriculture Washington, D.C.	Professor Vernon (Tom) Gilbert Adjunct Professor Graduate Program of Ecology University of Tennessee Gatlinburg, Tennessee  Ms. Paula Goddard Agency for International Development Bureau for Program and Policy Coordination Center for Development Information and Evaluation Washington, D.C.
Dr. Annette Binnendijk Agency for International Development Bureau for Program and Policy Coordination Center for Development Information and Evaluation Washington, D.C.	Dr. Douglas Horton Head, Social Science Division International Potato Institute C.I.P. Lima, Peru
Dr. Dennis Casley Chief, Monitoring and Evaluation Division Central Operations Department The World Bank Washington, D.C.	Dr. Stanley Johnson Director Center for Agricultural and Rural Development Iowa State University Ames, Iowa
Mr. Paul Duane Senior Evaluation Officer Operations Evaluation Department The World Bank Washington, D.C.	Dr. Krishna Kumar Agency for International Development Bureau for Program and Policy Coordination Center for Development Information and Evaluation Washington, D.C.
Dr. Timothy R. Frankenburger Office of Arid Land Studies University of Arizona Tucson, Arizona	
Mr. John Gaudet Agency for International Development Bureau for Africa Office of Technical Resources Agricultural and Rural Development Division Washington, D.C.	Mr. Ram Malhotra Director, Monitoring and Evaluation Division International Fund for Agricultural Development (IFAD) Rome, Italy



Mr. W. Haven North  
Agency for International  
Development  
Bureau for Program and  
Policy Coordination  
Center for Development  
Information and Evaluation  
Washington, D.C.

## RAPPORTEURS

Mr. Allen G. Turner  
Ms. Patricia Fisher-Harris

Dr. Beatrice Rogers  
School of Nutrition  
Tufts University  
Medford, Massachusetts

Dr. Thayer Scudder  
California Institute of  
Technology  
Pasadena, California

Dr. Alfredo Sfeir-Younis  
Senior Evaluation Officer  
Operations Evaluation  
Department  
The World Bank  
Washington, D.C.

Ms. Emmy Simmons  
Agency for International  
Development  
Bureau for Africa  
Office of Development Planning  
Policy Planning and Evaluation  
Division  
Washington, D.C.

Mr. Parmeet Singh  
Program Coordinator  
National Household  
Statistical Capability  
Project  
Statistical Office  
United Nations  
New York, New York

Ms. Nena Vreeland  
Agency for International  
Development  
Bureau for Program and  
Policy Coordination  
Center for Development

Information and Evaluation  
Washington, D.C.

## BIBLIOGRAPHY

- Casley, Dennis, and Krishna Kumar. 1987. *Project Monitoring and Evaluation in Agriculture*. Baltimore: Johns Hopkins University Press.
- Johnson, Pamela, Ahmed El Dahry, Richard Dekmejian, Eugene McJunkin, and Robert Morrow. 1983. *Egypt: The Egyptian-American Rural Improvement Service: A Point Four Project, 1952-63*. A.I.D. Project Impact Evaluation Report No. 43. Washington, D.C.: Agency for International Development, April.
- Kumar, Krishna. 1987a. *Conducting Group Interviews in Developing Countries*. A.I.D. Program Design and Evaluation Methodology Report No. 8. Washington, D.C.: Agency for International Development, April.
- Kumar, Krishna. 1987b. *Rapid, Low-Cost Data Collection Methods for A.I.D.* A.I.D. Program Design and Evaluation Methodology Report No. 10. Washington, D.C.: Agency for International Development, December.
- Norton, Maureen, and Sharon Pines Benoliel. 1987. *Guidelines for Data Collection, Monitoring, and Evaluation Plans for A.I.D.-Assisted Projects*. A.I.D. Program Design and Evaluation Methodology Report No. 9. Washington, D.C.: Agency for International Development, April.
- Organization for Economic Cooperation and Development (OECD). 1986. "Methods and Procedures in A.I.D. Evaluation." Paris: OECD.
- Verma, Vijay, Tim Merchant, and Chris Scott. 1988. "Evaluation of Crop-Cut Methods and Farmer Reports for Estimating Crop Production." London: Longacre Agricultural Development Center.