

PW-AAU-229

936-4062

---

# Methods for Evaluating the Nutritional Impact of Food Aid Projects: Lessons from Past Experience

David Sahn

---

Reprinted from  
Food and Nutrition Bulletin  
Vol. 6, No. 3, September 1984

000668

000780

**INTERNATIONAL  
FOOD  
POLICY  
RESEARCH  
INSTITUTE**

1776 Massachusetts Avenue, N.W.  
Washington, D.C. 20036

# METHODS FOR EVALUATING THE NUTRITIONAL IMPACT OF FOOD AID PROJECTS: LESSONS FROM PAST EXPERIENCE

**David Sahn**

Visiting Research Fellow, International Food Policy Research Institute, Washington, D.C.,

## PREFACE

The ideas and views expressed in this paper are based on discussions with the staffs of the World Food Programme and the Food-for-Peace Program and on a review of current documents (from the past three years) on efforts to evaluate the nutritional impact of food aid projects (see the list of references). As most of the documents examined were seriously flawed, and as only two weeks were available for their review, my personal experiences as an evaluator were of great importance in reaching conclusions. Therefore, I acknowledge that the nature of the document is partially subjective. It is designed to emphasize general issues that need to be addressed in order to move forward on improving efforts to assess the nutritional impact of food aid projects.

## INTRODUCTION

Food aid has undergone a transition from being surplus commodities in search of a purpose to being a valuable development resource with an opportunity cost approaching that of financial aid. This has resulted in a growing interest in evaluating the impact of specific projects administered by the World Food Programme (WFP) and the Food-for-Peace Program (US Public Law 480, Title II). Commensurately, the past few years have witnessed an evolution in the role of evaluation. The teams of experts sent out to the field to perform process-oriented, qualitative assessments have given way to a search for systematic and comparable evaluation methods that provide quantitative information on impact.

This paper will discuss the obstacles and misconceptions that have plagued the art and science of evaluation in

the recent attempts to quantify the nutritional impact of food aid projects. Gaining insight from previous experience is intended to provide guidance to those who embark on future evaluations, with all the obstacles and pitfalls that task portenos. This document will also address widely held preconceptions and misconceptions concerning evaluation. In doing so it will respond to the arguments of some that evaluation is purposeless and a waste of resources and of others who have become obsessed with searching for project impact.

This paper is limited to addressing experiences in evaluating nutritional impact of *project aid*, excluding emergency relief. Project aid includes food aid provided on a grant basis. It is used to support specific and defined activities to promote economic and social development, usually in the form of maternal and child health, food-for-work, and school feeding projects. It is distributed through differentiated market channels. Project aid is juxtaposed with *programme aid*, which is generally sold on a highly concessional basis and is not tied to a specific set of activities. Rather, the value of the transfer in the form of programme aid accrues to a government treasury that has a high degree of freedom in how it is spent. Attention is primarily focused on vulnerable group or supplementary feeding projects. Nutritional improvement is a nearly universal and prominent objective of these interventions.

Other types of food aid projects, however, also fall within the purview of this paper. These include the use of food as wages (i.e., in-kind payment) to develop infrastructure and assets, as an incentive to resettle new lands and adopt new agricultural practices, and as inputs to spur agricultural development at a low cost to the government in terms of foreign exchange and budgetary support (for example, dairy or oilseed development projects). These measures all have *indirect* nutritional consequences, which may be positive or negative. They are mediated by the general process of economic development. They can and should be measured primarily and observed in the long term, after the active distribution of food in the community has ceased.

---

This paper was prepared for a symposium on the Nutritional Impact of Food Aid, tenth session of the UN ACC Sub-committee on Nutrition, Rome, 5-9 March 1984.

## FRAMING THE PROBLEM

The resource of food is inextricably linked to nutrition. The result is that there is characteristically a search for nutritional impact of food aid projects, even in those that do not explicitly state nutritional improvement as one of the primary objectives. A review of food aid evaluations, however, tells a rather dismal story. Nutritional impact has rarely been substantiated. In those few instances where evaluations determine that nutritional improvement was observed, a critical review of the methodologies employed inevitably leads one to suspect the conclusions.

The obvious question is: Why have the evaluation experiences yielded unsatisfactory results, and what can be done to improve upon past performance? A review of previous and current evaluation studies of the WFP and PL 480, Title II, provides the data to answer this question. In combination they elucidate the reasons why there has been so much difficulty and controversy in evaluating nutritional impact of projects. These common problems, which are observed in a number of settings, will be expanded upon below. I must stress that not all of the issues addressed are applicable to *all* projects or evaluations reviewed. The purpose is to recount briefly the recurrent difficulties that plague the evaluation of nutritional impact, thereby laying a foundation for improvements in the future.

### Poorly Conceived and Designed Projects

The historical role of food aid in meeting multiple objectives, coupled with the fact that most formulas for the use of food aid were developed to expedite rapid disposal with minimal financial and political costs, has conditioned the current operation of projects. The result is that there are serious deficiencies in the design and theoretical foundation of food aid projects. These explain to a great extent the reason impact evaluations have failed to successfully show marked nutritional benefits. One could make a strong argument that the scarcity of efforts to evaluate impact is appropriate, given the inadequate conceptualization of food aid projects. Just as a moral conviction prompts the impulsive reaction that handing out donated commodities must be beneficial, especially when distributed directly through differentiated market channels, a review of the design of most projects makes it equally easy to adopt the sceptics' viewpoint that nutritional impact is an unrealizable goal for *most* supplementary feeding projects and may at best be a faint long-term hope in other projects designed to promote economic growth and development.

The problem is most clearly manifested in the conceptual

weakness of many projects that fall in the domain of targeted feeding interventions. A number of underlying assumptions remain largely unsubstantiated. They must be accepted before one can reasonably argue that supplementary feeding projects can be expected to reduce malnutrition. These include that (a) it is feasible to identify those at greatest nutritional risk in the community and subsequently encourage their regular participation; (b) through the provision of food commodities, along with efforts involving moral suasion, household nutrient availability can be increased significantly; and (c) the perception of needs within the household can be altered to encourage a new allocation of food commodities among family members.

To illustrate, the value of the transfer from most supplementary feeding projects is usually less than 5 per cent of annual total household expenditures. Empirical evidence also indicates that a 1 per cent increase in income will lead to an increase in calorie consumption of up to 0.7 per cent for the poorest of the poor. This figure tends toward 0.3 to 0.5 per cent for the more typical indigent household. Recent research has shown that income elasticities of demand for calories among the poorest of the poor households is 0.3 in Brazil (1), 0.61 in India, 0.67 in Bangladesh (2), 0.74 in Indonesia (3), and 0.71 in Sri Lanka (4). Similar figures for households near the poverty line (although not the poorest) show markedly lower elasticities of 0.35 in Bangladesh, 0.37 in Indonesia, 0.44 in India, and 0.43 in Sri Lanka. There is little doubt that the participants in food aid projects are poor and at risk. However, participation is less likely by the poorest, who are most inaccessible and hardest to reach because of where they live, time constraints, lack of education, and so forth. Thus, it is assumed that the realistic calorie elasticity figure is between 0.3 and 0.5.

Some have argued that, since the income transfer is in the form of food and typically accrues to the women heads of household, consumer behaviour is affected, thereby raising the marginal propensity to consume (MPC) calories. Arguments working in the other direction are that, to the extent that in-kind income is viewed as transitory or the value of the donated commodity is discounted because it is used for resale or bartering of goods, such transfer will actually lead to a lower MPC, and that nutrition education at the feeding centre effectively alters consumer tastes and preferences or intra-household food allocation. Despite the early research that indicated that sources of income do affect consumption behaviour (5), more recent research (6) indicates that the MPC for the food ration is not different from cash income in Panama. However, both studies acknowledge the need for more rigorous research in this area. Therefore, given the typical quantity and composition of the food ration, one must question the expectation that

most food transfers will result directly in a measurable impact on net household food consumption.

It is even more uncertain that the intakes of targeted individuals (i.e., mothers and children) will be raised significantly despite increased household calorie consumption. This is especially the case for take-home feeding projects, where both sharing and substitution of the ration are likely. In addition, even if a feeding project increases household and individual caloric consumption, the link between that and measurable changes in anthropometric measurements remains tenuous (7). This is because of the complex aetiology of malnutrition. For example, utilization of the incremental nutrient intake may be low because of infection, or higher calorie intake may result in increased activity or metabolic adjustments, thereby increasing the requirement for energy. Thus, it should come as no surprise that few evaluations have proved nutritional impact.

Indeed, qualifying circumstances may raise appreciably the likelihood of observing improvements in nutritional status. First is the case when ancillary services, such as health education, encourage better sanitary practices, use of oral rehydration solution in cases of diarrhoea, and prenatal care. Past studies have indicated that such efforts can bring about impact mediated through factors such as breaking the malnutrition-infection synergism. Nevertheless, experiences to date reveal that in most circumstances adequate attention is *not* given to these non-food aspects of programming. Furthermore, in those cases where food acts mainly as an incentive, a whole series of other questions are raised that have yet to be asked, let alone studied. One must consider the alternatives to food aid that can animate the community and encourage participation in self-help schemes, education programmes, and other services that improve access to health care and promote better hygiene and child-feeding practices.

The second qualification is the renewed interest in targeting the household, not the individual, in feeding projects. Some recent projects have been designed to provide a transfer sufficiently large to have a meaningful impact on household expenditure levels. Evaluations have yet to examine this strategy. Nevertheless, its adoption has clear implications, not only for factors such as selecting the most cost-effective commodities, but for broader policy issues. These include (a) the wisdom of various alternative delivery systems for the transfer (e.g., ration shops, subsidized marketing costs using normal distribution channels); (b) alternative forms of assistance (e.g., the shipment of commodities in bulk, rather than the inordinately expensive packaging techniques

used for Title II and WFP commodities); and (c) the use of cash rather than a food transfer that may be discounted or have negative effects on local agricultural development. If the goal is to improve nutritional status by increasing incomes, fundamental reconsideration of how to do so is required. This should go beyond the selection of commodities and consider the distribution and administrative systems as well.

Targeted supplementary feeding projects are not alone in suffering from fundamental weaknesses in their design. Inherent conflicts in food-for-work projects, such as selecting participants to maximize utility of the food transfer and selecting participants to maximize the marginal productivity of the project are often not reconciled in the project plan. To illustrate, projects installed in regions with greater natural resources (e.g., fertile soil) or that have achieved a critical level of social and economic infrastructure (e.g., market access) will often have greater returns for investments. This is because of the synergism of project inputs with existing endowments. However, the poorest communities are in drought-prone and more isolated areas, where investments may be less economically sound. The result is that both welfare and development objectives are compromised. This is manifested in problems such as true objectives differing from those stated publicly or projects not providing the necessary complementary resources (i.e., tools, equipment, managerial support).

Planners and politicians must address explicitly these contradictions. If this is not done, confusion will inevitably hinder project implementers and functionaries. This will reduce the potential for any measurable achievement. Evaluators may also finally come to assessing project performance based on short-term welfare objectives rather than long-term economic growth.

This is not the place to resolve the issues discussed above. Rather, they are raised for three reasons. *First*, many evaluations of impact have not proved fruitful because of poor planning and the inadequate attention to underlying assumptions that form the foundation for project design. Technical limitations of evaluation methods are often not the primary reason for the failure to capture nutritional effects.

*Second*, untested assumptions which form the basic logic of a project's design should be addressed through carefully planned research studies, as juxtaposed to evaluation of operational projects. *Third*, and most disconcerting, is that many of the design and conceptual problems of food aid projects have come to our attention as a result of previous process and impact evaluations; however, most remain to be acted upon.

Much study has focused on these ingredients for success: the imperative of integrating nutrition education, improving targeting and outreach, promoting diarrhoeal disease control as a component of supplementary feeding projects, and the necessity to increase availability of complementary inputs and managerial support for food-for-work projects.

The fact is, however, that acting upon much of the knowledge gained to date requires further commitments of resources to bolster the quality of projects. Many donors and recipients alike have been reluctant to assume the costs involved. These are mostly financial to alleviate manpower constraints and provide non-food resources. There are, however, political costs that revolve around closer co-ordination between donors and host governments. Therefore, the first item on the agenda is to make use of the knowledge that already exists concerning constraints to achieving sound project design and performance. Undoubtedly the cost of doing so will extend far beyond those incurred at present. On the other hand, if the evaluations are to be taken seriously, there is an imperative for acting vigorously on their conclusions. We must not allow the call for better and more evaluation in the future to obscure the need for action in the present.

#### Implementation Constraints

Food aid projects are often not implemented as planned. This problem was noted in a number of countries where evaluations were performed. It is difficult to gauge the actual dimensions of this problem. Experience suggests it is large. But regardless, it is amply evident that one cannot find nutritional impact in environments where projects are not operating effectively.

This issue of poorly functioning projects provides a number of lessons. It is a reminder that process evaluation must precede or be carried out simultaneously with the search for impact. Rather than abandon evaluation efforts altogether, it will be more fruitful to link impact studies more closely with efforts designed to improve project performance.

For example, severe logistical problems were identified in Malawi, a site chosen by the WFP for an in depth evaluation (8). The response was first to resolve logistical problems, thereafter, procedures to determine impact would be developed and instituted. The evaluation team separated the development of a process-oriented management information system to improve project performance from impact evaluation activities. Similar recommendations have been made to WFP by others who argue that "ongoing management and operational

evaluations" should no longer fall within the domain of the Evaluation Service, who should focus on "providing information important to policy formation and project design" (9). This is the wrong approach. Instead, impact indicators should be incorporated into routine data collection procedures that are part of the management process from the outset. This will have numerous advantages. It will create an opportunity to collect baseline information that will prove vital to further evaluation efforts. Similarly, a well-developed management information system will provide secondary data that can subsequently be used to evaluate projects (10-12). And most important, the development of a goal-oriented management system, in which impact data are key elements, is the best source of positive feedback on project performance, and serves as a powerful motivating force for the field staff.

#### Posing the Appropriate Question

An impact evaluation is only as good as the questions asked and hypotheses posed. Often objectives are ambiguous or do not follow logically from the project inputs and outputs. The result is that evaluators set out to assess *the wrong type of impact*. This problem is manifested by the use of inappropriate indicators for measurement and comparison. As an illustration, note the case of school feeding projects in general and the recently completed evaluation of the Jordan project in particular (13). The quasi-experimental evaluation study was well designed. Findings indicated that the nutritional status of children in both treatment and non-supplemented villages in Jordan deteriorated, only more so in the treatment villages.

In drawing inference from the lack of reported impact, a few important points must be considered by the reader. First is whether school feeding programmes should or can be justified on nutritional grounds and, thus, whether the evaluation was measuring goal achievement in terms of an appropriate objective. If nutritional improvement in the community is the overriding goal, inevitably resources are better allocated to women and preschool age children, who are the most vulnerable groups. Even the author of the Jordan study acknowledges that "the school children of this study were not at an age where growth can be significantly affected by a food supplement." I would further argue that not only is it of questionable validity to define school feeding objectives (and thus assess impact) in terms of nutritional objectives but, if such an approach is used, indicators other than anthropometry are required.

Justifying school feeding projects on nutritional grounds is precarious. Previous experience indicates that searching

for nutritional improvement will yield little or no encouraging results, regardless of the quality of the study design. In the long run, the inadequate proof of goal achievement will reduce or eliminate support for such projects. Instead, educational objectives, such as improved attendance, enrolment, and academic performance, should be espoused for school feeding projects. Impact studies must focus on these questions. The higher goal (although not to be measured) is promoting economic development through investments in human capital.

The problem of asking the right question, which leads to evaluating impact based on the proper criteria, is perhaps nowhere more apparent than with food-for-work projects. There are two difficulties: the first is the need to distinguish between short term and long-term impacts; the second is the inherent conflict between welfare and growth-oriented objectives of food-for-work projects.

Concerning the former issue, with the exception of the study under way currently in Bangladesh, the few attempts to examine the nutritional implications have focused on short-term effects. The appropriateness of this approach is conditioned by the second issue: the extent to which the project is oriented towards relief versus construction of infrastructure and assets that generate a stream of economic benefits. For many projects that *indirectly* affect nutritional status through the process of economic development, searching for nutritional impacts, especially in the short term, is the wrong exercise.

This is well illustrated in the context of four prominent types of food-for-work (FFW) projects: (a) agricultural adjustment projects in which farmers who innovate are provided food to reduce risk and uncertainty; (b) resettlement schemes in which food, along with a number of other incentives, encourages households to migrate to new areas where land is more plentiful and/or opportunities greater; (c) projects where food serves as an input to encourage local production of milk or oilseeds; and (d) price stabilization schemes. There is little reason to expect these projects to improve nutritional status markedly during the short-term when the food assistance is actually being distributed. A combination of factors reduce the likelihood of improvement. These include the following circumstances: (a) the expected stream of returns on investment does not begin to flow immediately; (b) the poorest of the poor are not likely to be recipients (especially in projects designed to construct viable assets and infrastructures); (c) participation is often irregular and intermittent; (d) the income is considered transitory, thereby causing a low marginal propensity to consume; and (e) recipients are often drawn from other low-productivity areas of employment and are rarely completely idle.

Furthermore, it is likely that in the absence of the project, other adjustments will take place to maintain food consumption levels. They include (a) the sale of assets (e.g., livestock), (b) consumption of inferior commodities (e.g., an increase in consumption of famine foods such as wild tubers), (c) consumption of seeds earmarked initially for the next planting, and (d) lower levels of physical activity (especially of the household head) that reduce household demand for energy (as well as possibly alter intra-household allocation in favour of children or women).

In general, then, the practice of assessing the nutritional impact of food-for-work projects that are designed to promote economic or agricultural development should not focus on the short-term. This has been proved by previous experiences. As remarked in a recent review, none of the studies examined that were undertaken during the construction (food distribution) phase "have shown that a FFW project resulted in nutritional improvement in the participants or their families" (14). While this partially reflects the gross deficiencies in the study designs employed, even more compelling is the argument that the goal of such development projects is longer-term in nature, and thus evaluations should reflect that perspective.

Before examining the issue of nutritional impacts over the longer term when the return from investment is expected, one must first acknowledge that FFW is often initiated as a relief mechanism to cope with events such as seasonal hunger or critical shortfalls due to crop failures and man-made disasters. Even for these projects, however, it is of questionable value to judge project achievement (i.e., welfare) on the basis of anthropometric or other more complex nutritional indicators (immunocompetence, biochemical tests).

Rather, in order to assess the short-term welfare impact of relief-oriented food-for-work projects, more direct indicators, such as prices of staple foods and their rate of disappearance from the market, or consumption data from a sample that examines the levels of intake of traditional famine foods or the mix of starchy staples in at-risk households, may prove more appropriate. They will be more responsive to the issue at hand – whether the project is reaching intended relief objectives – than the collection of nutritional status data that reflect a phenomenon of complex and multiple aetiology.

In examining the long-term (during the operational phase) growth and distribution effects of FFW, a number of questions arise as to the appropriateness of measuring nutritional impact. Other than a well-designed study under way currently in Bangladesh (15), there are virtually no good experiences to draw upon where

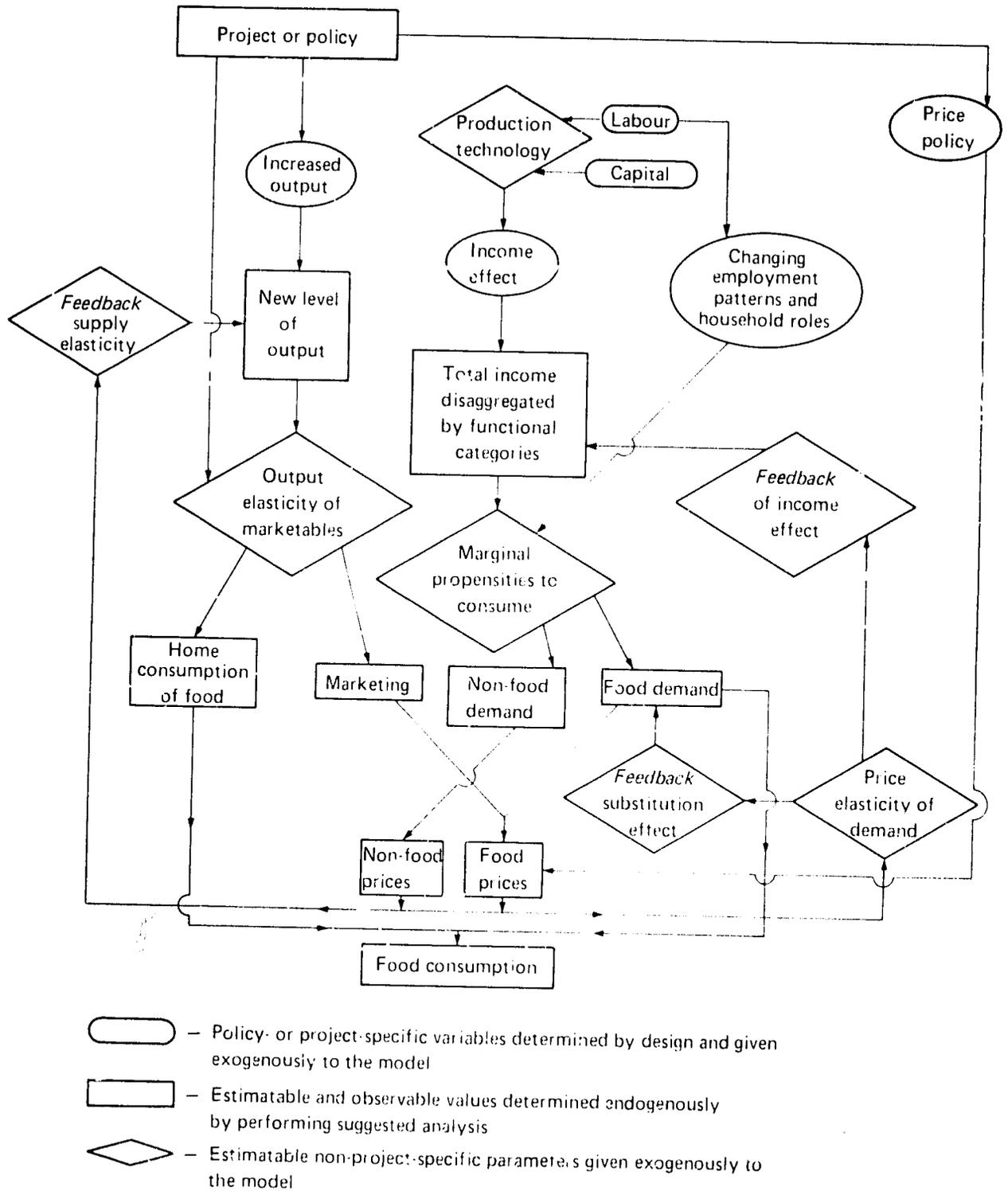


FIG. 1. Conceptual Diagram for Modelling Nutritional Consequences of Policies or Projects (Source: D. E. Sahn and C. P. Timmer 1984, based on work by Anderson 1981)

nutritional changes are measured. This is no surprise. The data requirements and rigours of such a study preclude widespread application of the methodology. This suggests that attempts should be made *ex ante* to assess the nutritional consequences of FFW, as in all development projects. Estimating the effects on key variables such as commodity prices, incomes, and the ratio of home consumption to marketing, disaggregated by a functional classification that distinguishes the nutritionally vulnerable households, should be a normal part of the project appraisal process. Figure 1 presents a scheme for doing so. Integrating nutritional concerns into FFW projects need not become a burdensome and highly quantitative exercise. Rather, it will generally suffice to simply understand the direction of anticipated changes in prices and incomes. Thereafter, qualitative judgements can be made as to whether a project will improve or represent a hazard to the nutritional status of at-risk groups.

Effective food-for-work projects will indirectly raise nutritional levels of the poor. The hallmarks of such projects are increased food availability that moderates food prices and higher income among the poor. In a few cases these objectives may be achieved within a single agricultural cycle. In other cases the time horizon will be longer. In either case, I take exception to the recommendation that nutritional status is an appropriate indicator for standard of living because its measurement "is relatively straightforward and unambiguous" (14). The truth is that there are too many factors that reduce the likelihood that a food-for-work project will translate into traditional indicators of nutritional improvement. Although the occasional research project may determine whether a project results in expected nutritional outcomes, emphasis should be placed on formulating a project design that seems likely to increase net food consumption of vulnerable groups. Thereafter policy makers should concentrate on the myriad of operational problems that have been characteristic of food-for-work schemes.

A review of food aid evaluations provides several instances, other than the school feeding and FFW examples cited above, where questions concerning impact are formulated incorrectly -- for example, the use of food as a rehabilitative tool or to prevent malnutrition at the first sign of growth faltering. The food is only distributed to children falling beneath a cut-off point of a reference growth curve or to children who fail to gain weight for two or three consecutive months. Other services are also provided. These include intensive educational efforts (e.g., to improve feeding practices and hygiene; to encourage the use of oral rehydration therapy) and such activities as providing resources for home gardens. The intent is not just to rehabilitate the malnourished child, but to

prevent reoccurrences. However, the evaluations reviewed did not examine whether fundamental changes in knowledge or behaviour had any impact on preventing children from becoming severely malnourished so that rehabilitation feeding in the form of food aid becomes obsolete.

These problems reinforce the need to define objectives clearly, based on a logical project design. If this is done, evaluations will undoubtedly respond accordingly.

### Selecting the Correct Indicators and Performing Accurate Measurement

Measuring nutritional status is a complicated undertaking. No single measure is completely sensitive and specific to the distinction between well-nourished and malnourished. Selecting the correct indicators and collecting data accurately add considerable complexity to evaluating nutritional impacts. The fact is that the relationship among indicators, and between indicators and changes in nutritional status, was not addressed by many evaluators. The implication is that, even if improvements do take place in nutritional status, they may not be detected.

To illustrate, consider the following examples of the limitations of indicators often employed (16):

- Arm circumference does not respond in the short or medium term to changes in nutritional status. This measurement will therefore probably not reflect improvements due to an intervention.
- Weight-for-age, which is a composite of stunting and wasting, may be low due to deficits incurred years previously and not to present status. Children may be misclassified as malnourished even if their status has improved, since evidence exists that chronic malnutrition during certain susceptible time periods may result in stunting without subsequent catch-up growth (17).
- Weight-for-height measurements are not sensitive to improvements in mildly or moderately malnourished children. A normal ratio of these measurements will be maintained by reduced activity and metabolic adjustments until the child is severely deficient in intake (18).
- Little is known about the dose response of increased caloric intake and how this will be manifested in terms of improvement in growth indicators. Lack of growth response despite increased energy intake, may be explained by increases in the level of physical activity and metabolic rate (19).

Without going into further detail, the above illustrates that the wrong indicator will result in failure to capture the benefit of a programme. In the future, more careful consideration must be given to identifying the nature

and extent of the nutrition problem, the expected scope for improvement, and which indicators would best capture this improvement.

Similarly, there is evidence that *random* errors in field measurements may have serious implications for the findings of an evaluation. If children in both control and treatment groups are randomly misclassified due to inaccurate field measurement, in most cases this will result in a low estimate of the difference in rates of malnutrition between the two groups (20). This further reinforces the importance of using good measuring equipment and training procedures.

A final methodological issue that has limited the usefulness of evaluations reviewed is that analytic procedures such as choice of standards (anthropometric or dietary), cut-off points (below which one is considered underfed or malnourished), and statistical techniques (e.g., the use of percentiles of median values versus Z-score cut-off points) are different from one study to the next and often not based on sound rationale. The implications of these choices for the prevalence of malnutrition are enormous. For example, Drake et al. (21) showed that almost twice the number of children were classified as second- and third-degree malnourished when the Harvard, rather than the Gomez, growth standards were applied to a population in Brazil. Similarly, Sahn (22) found that, even when employing the same growth standard (23), the percentage of children classified as wasted in Sri Lanka was almost twice as high when a -2 Z-score cutoff was used, rather than the traditional 80 per cent of the median weight-for-height cut-off point. As a consequence of these differences in standards and cut-off points, comparisons between impact studies are spurious. Therefore, it is urged that agencies move towards standardizing analytical procedures.

### Research Design and the Difficulty of Attribution

One further obstacle is often observed in many evaluations. This concerns documenting change and comparing that change to what is expected *in the absence of the project*. There are a number of factors that characteristically impede the attribution of improvements in nutritional status to project activities. Such confounding factors can be negative or positive. If negative, successful interventions seem ineffectual; if positive, unsuccessful interventions appear to have improved nutritional status. Thus, the wide range of competing explanations (including the project) for observed changes in nutrition often result in indeterminate evaluation results.

The most predictable competitor with a project as an explanation for nutritional changes, which was overlooked in many studies reviewed, is the phenomenon of population

aging. Rates of malnutrition, according to anthropometric indicators, vary markedly by age. Failure to take this into account (i.e., controlling for the age distribution of the population and limiting comparisons to six-month groupings) will lead to spurious results.

A second phenomenon observed in malnourished populations concerns the contention that those with the worst nutritional status benefit most from a programme. This statement was nearly universal in the evaluations reviewed. However, naturally occurring changes in populations, including regression toward the mean, can once again explain such findings. That is, there is a spontaneous movement of individuals who fall at extremes of a measurement scale (e.g., grade III malnourished) to regress towards the population mean. This is logical. If you have grade III malnourished children, they either stay the same or improve (or alternatively die, which in most cases means that they are no longer part of the sample). On the other hand, you would expect a certain number of grade II malnourished children to improve and the remainder to worsen, becoming grade III.

A variety of other confounding influences were often not accounted for in the evaluations reviewed. For example, there was a failure to consider the addition or attrition (i.e., in- or out-migration) of participants that may have altered the composition of a treatment or matched control group. Furthermore, most evaluations reviewed did not consider the social and economic factors or physical phenomena that are competing explanations for nutritional changes. The complexity and volatility of environments in which programmes are operating require that consideration be given to these historical and environmental factors. Unfortunately, when outside evaluators enter unfamiliar environments and have little knowledge of what secular events transpired during the course of the intervention, it is difficult to account for such confounding factors.

There are numerous other competing explanations and threats to internal and external validity not addressed above. The literature discusses these issues and cautions the reader to be sceptical of evaluation findings that stem from poorly conceived methodologies (24). The question is what to do, given these problems.

In an ideal world, future impact evaluations would employ classical experimental design characterized by randomization in selecting treatment and control groups. Doing so would leave little doubt that observed differences were attributable to the project. Unfortunately, an experimental research protocol in the context of operational projects is all but impossible. Reasons include: (a) programme planners and staff may resist randomization as a means of allocating treatments, arguing for some other

criterion, e.g., need or merit; (b) the randomization process is difficult to carry out correctly in highly dynamic environments, resulting in non-equivalent (test and control) groups; (c) there is a high likelihood of spill-over effects from the treatment to control population; and, most persuasive, (d) the expense of running good experiments precludes their use on a wide scale.

Quasi-experimental or non-experimental evaluation designs are therefore relied upon. Most popular among these are:

- *the one group pre-test/post-test design*, in which initial measurement is performed on the population, followed by the delivery of services and post-programme measurement to see what changes occurred;
- *the static group comparison design*, in which a project is initiated and, at some time after the programme has been operational, measurements are taken on the treatment group and on other similar populations not receiving services and are then compared;
- *the non-equivalent control groups design*, in which matched treatment and control groups are both selected and measured prior to the beginning of the project, and then have their nutritional status re-assessed at some time after the project has been operational;
- *the recurrent institutional cycle design*, in which the nutritional status of children who have been participating in the project for a meaningful length of time is compared with that of individuals who have recently enrolled.

The evaluator should be familiar with the strengths and weaknesses of these and all other possible options. With all these techniques it is also plausible to use multivariate statistical techniques (e.g., regression analysis) to control for non-project variables that may be important determinants of nutritional status.

Quasi-experimental designs, however, have met with only limited success in controlling for negative and positive confounding variables. This reflects partially the inherent limitations of the methods. However, there is considerable room for improvement in the application of these techniques. It is imperative that future evaluation studies (a) should place greater emphasis on reducing the extent to which there are competing explanations for relating the delivery of services to nutritional outcomes (i.e., on minimizing the threats to validity), and (b) should acknowledge and discuss candidly the many biases that creep into analytic procedures. Thereafter biases can be dealt with in a qualitative fashion by involving those individuals most familiar with and knowledgeable about the project and the local environment in the interpretation of the data. An example of adhering to these principles is found in the evaluation of the weaning foods project in Sri Lanka (10). Following the data analysis, the outside evaluators engaged local workers in a dialogue. The original

interpretations of the data were altered significantly in response to discussion with those on site.

The implications of these principles do not militate against attempting to measure the nutritional impact. They do, however, warn against hastily conceived methodologies used in the past that may grossly over- or under-estimate changes in nutritional status.

## RECOMMENDATIONS FOR FUTURE EVALUATION ACTIVITIES

The difficulties encountered in substantiating impact are indeed discouraging. Regardless of whether they are caused by ineffectual programming, unsound evaluation methods, or simply the problems inherent in performing social science experiments in volatile and complex village settings, the cost of proving the nutritional impact of food aid projects on anything approaching a global basis will be high.

This is amply illustrated by the resources required just to carry out those few sound impact study designs, such as the proposed WFP evaluation in Honduras (8) and the food-for-work study under way in Bangladesh (15). (These costs say nothing of the need to infuse resources into the improvement of the projects themselves.) Expectations from these major research endeavours should be limited, however, given that previous research efforts on food aid projects have yielded equivocal findings. In any event, the cost of applying such methods on a wide scale is prohibitive. One must also question the value of such studies in terms of what is considered the primary pitfall of food aid projects – poor design and breakdown in implementation. Therefore, the question is: Where do we go from here? It is suggested that a two-tier strategy should be adopted. First is to develop built-in evaluation procedures to assist in designing and implementing projects. This will increase the probability of achieving objectives; simultaneously, it will facilitate assessing nutritional impact. Second is to design and undertake a series of small-scale operational research studies to test vital hypotheses rather than try to determine impact *per se*.

### Building Evaluation into Standard Operating Procedures

The development of *built-in* monitoring and evaluation systems should be accorded the highest priority for all food aid projects. The building-in of evaluation procedures begins before operations commence, i.e., when the project is being conceived. This will encourage integrating the collection and interpretation of process and impact data into standard operating procedures.

The attributes of an evaluation system built into the project are described briefly below. The adoption of such a system for all projects is urged.

### *Building Evaluations In at the Planning Stage*

A strategy for impact and process evaluation should be developed concurrently with designing the project. This involves defining the basic components of the monitoring and evaluation systems. According to Miller and Sahn (16) these are:

- *a data system* that describes the actual variables to be collected, how often, and on what population; these objectively verifiable indicators will include key process and impact elements;
- *an analytic methodology* that delineates the computational algorithms used to describe and interpret data; this involves defining explicitly the statistics employed to summarize a data set, the nature and methods for aggregating data, and the types of procedures used to interpret the statistics;
- *information flow* that details the forms to be used and flow of information through the system, including what is passed up to higher levels of management and what information is fed back to functionaries in the field;
- *a management structure* that outlines explicitly how the information is to be used at each level of management hierarchy, i.e., the range of decisions and types of actions to be taken.

Defining the monitoring and evaluation system as part of the initial project design will yield a variety of dividends. This procedure will encourage improvements in the diagnosis of the nature and extent of the nutrition problems. Encouraging the definition of objectively measurable units (i.e., data elements) for process and impact evaluation will immediately focus attention on objectives and the logical progression of events that will lead to their realization. It will also enable the collection of benchmark (baseline) data on strategic impact indicators. Furthermore, delineating acceptable levels of achievement or changes in these indicators before starting a project will provide quantitative targets to serve as a source of comparison in the future. In this regard, the use of tools such as the logical framework is recommended. However, such tools are often used in a perfunctory manner. They run the risk of becoming an excuse for not thinking through the hard questions. Caution should be taken to avoid "filling in the boxes" at the expense of performing the requisite analysis.

### *Improving Logistics and Cost-Accounting*

It is not unusual for logistical breakdowns to impair food supply lines. Similarly, complementary inputs such as tools

(e.g., shovels), equipment (e.g., scales for weighing), and raw materials (cement for road construction, oral rehydration salts for diarrhoeal disease control) are often not delivered to the project site on time, or at all. An information system that carefully monitors the stocks and flows of project inputs serves as an essential ingredient in understanding and explaining impact indicators. For example, consider two villages. In the first, the monitoring of inputs shows logistical breakdowns and impact indicators fail to substantiate nutritional changes in the community; in the second, the information on stocks and flows records few breakdowns and impact indicators show marked improvements. In yet another village, with even greater nutritional improvements, the monitoring system may indicate either that a different set of inputs was provided or that they were applied with different intensity, for example different ration sizes. The availability of these data on inputs represents vital information for constructing a story as to *why* and *how* the project succeeded in reaching its goals. The accumulation of knowledge over the extended life of a project will provide a formidable data base not only for assessing impact but for explaining it as well.

Monitoring project inputs will also encourage the involvement of local functionaries in rationalizing their flow. This will help guard against characteristic stories such as spoilage of food commodities or pharmaceuticals, workers sitting idle awaiting the delivery of inputs, or the arrival of the monsoon causing ditches to fill up with silt because the delivery of irrigation pipes was late.

Perhaps more important in the context of impact evaluation, *a well-developed monitoring system forms the basis for cost-accounting*. Experiences in evaluating nutritional impact display a neglect of the relationship between costs and effects. This is especially the case for supplementary feeding programmes. It is stressed that the village-level project managers are not necessarily responsible for the careful quantification of the value of inputs. This task will be beyond their ability and time availability. Inputs should be valued at their opportunity cost. Considered judgements will be required to determine shadow prices and to distinguish which portion of shared costs should be assigned to the project. Nonetheless, simple and well-maintained monthly inventory reports will form the basis for technical experts to assign cost figures to project inputs when doing their analysis.

### *Achievement of Output Objectives*

Accounting for project inputs leads to the next aspect of a built-in monitoring/evaluation system – documenting that the resources were used as planned. Monitoring that inputs were transformed into outputs focuses on whether

and what services were actually provided and to whom they were delivered. The type of information included may be, for example, the size and number and composition of food rations delivered; the economic or social characteristics of the individual participants (to assess effectiveness of targeting strategies); the number of miles of roads built; the attendance at the weekly nutrition and health education classes; or the number of new enrollees and drop-outs in the project. In addition, it will often be appropriate to collect key non-project-specific data. These would cover variables exogenous to the project (e.g., staple food prices) that are considered vital to addressing confounding factors for nutritional changes.

Given the historical problems in the operational aspects of food aid projects, not only are data on achievement of outputs necessary to interpret and understand impact indicators, but their collection is commended by the need for project management to have information that can be used to enhance the quality and extent of supervision. This suggests that the selection of output indicators should be largely a function of identifying those variables that are most amenable to change or control by field staff. In addition, the information system, like the project itself, should be viewed as dynamic. *It should be revised in accordance with the needs and experiences of field workers and policy-makers.*

#### *Determining Project Impact*

The existence of a monitoring system that substantiates the causal hierarchy of project events (i.e., that inputs were provided, then transformed into outputs) lays the foundation for the evaluation of impact. The basic premise is that a self-evaluation system will be superior to expensive quasi-experimental research. This is supported by a number of compelling arguments.

*First* is that the steps discussed previously will enhance the probability that projects are well designed and performing efficiently. Those doing data gathering and recording for a project will understand its purpose and be encouraged to interpret and act upon available information.

One of the most difficult aspects of designing development projects is the dynamic nature of the environment in which they are implemented. Characteristically, a standard formula, or at least an inflexible set of instructions, dictates the nature of the services provided. The latitude given to those on site is minimal. Given that the duration of most food aid projects is measured in years, a key ingredient to project success is the ability to make changes in the nature of services provided and level of intensity in application. This involves the ability to respond to factors such as seasonal fluctuations in the food supply or prices, or

the evolving context of the project site over time. These types of decisions must be made at the local level rather than in Rome or Washington. The discretion of the local staff based on knowledge gathered through a well-functioning self-evaluation system is therefore indispensable. Just as a field worker must be able to rationalize the stock and flow of project inputs to avoid shortfalls in supplies or wastage, it is important that the monitoring and evaluation system trigger programmatic changes, as appropriate, based on sound and timely data (21).

The data included in a built-in evaluation system may be of a wide variety, ranging from anthropometry to food prices. As an illustration, note that there may be a need to adjust the size of the food ration if no nutritional impact is being observed. The time may also come, according to data collected on the use of oral rehydration solution, to shift emphasis from the efforts to control diarrhoea to promotion of improved hygiene practices to prevent infection. Similarly, the number of work days provided by a FFW project may have to be adjusted based on seasonal fluctuations in labour demand or food availability. Although this may seem difficult, its feasibility has been demonstrated in the context of the Early Warning Information and Intervention System in Indonesia (25). A variety of data collected regularly indicate or predict changing patterns in food availability and consumption that trigger a range of interventions, depending on the severity of the signals from the information system.

*Second*, a monitoring and evaluation system will afford greater opportunity for outside experts, in conjunction with local field staff, to draw meaningful inferences concerning impact from available data. (As noted previously, the best evaluations performed were those that tapped into an existing information system.) This is attributable to various causes. It is likely that the quality of data will be improved because field workers use them in the context of project management. Longitudinal sequences on routinely gathered data elements will represent an enormously rich data set. Indeed, the number and complexity of impact indicators will have to be limited in order to maintain the manageability of the system. For example, in a supplementary feeding programme, it would indeed be preferable to collect a battery of measurements (e.g., height-for-age and weight-for-height data). Doing so may prove impracticable in a self-evaluation system. It may only be feasible to use a composite indicator such as weight-for-age. The necessary compromise, however, will be more than compensated for by the completeness of coverage and the availability of key process indicators.

A *third* advantage of a built-in evaluation system relates to the difficulty of controlling for confounding variables (i.e., threats to validity). Involving local personnel in the

interpretation of data will usually resolve what appear to be surprising or contradictory outcomes of data analysis. External evaluators may be involved in a determination of the changes that have occurred. They may even limit the possible explanations for such changes. In the final analysis, however, on-site personnel must distinguish among the competing explanations for the findings. Only their intimate knowledge of the project will be appropriate for such a task.

It is appropriate to acknowledge the arguments against promoting an internal information system. First is the cost. Undoubtedly a well-functioning information system adds significantly to the cost of the project. Most of the additional costs will be in designing the data systems and training local personnel in their use. In fact, five to ten per cent of project costs may be absorbed by a monitoring and self-evaluation system. However, such systems are considered vital not only to the assessment of project performance but to the achievement of impacts as well. Monitoring and evaluation systems form an integral component in any attempt to improve operational performance. The cost of developing such systems is not only justified but essential.

A second argument against built-in evaluation systems is that project personnel are incapable of collecting and recording the data. The information from the system would then be unreliable. My response is that, if a field worker is unable to measure and record accurately data elements vital for assessing project operations and achievement of objectives, one must question the appropriateness of the entire intervention scheme. Previous experience suggests that it is precisely those operationally superior projects that have well-functioning information systems. Thus, breakdowns in the information system serve to warn country officers or headquarters of potential implementation problems.

The other possible reason for poor data quality is that there will be a great deal of subjectivity and bias in a self-evaluation system. This assumption is not valid, especially when built-in monitoring and evaluation systems are viewed as constructive enterprises by the staff in their operational and management roles.

With a well-functioning monitoring and evaluation system, it will be feasible to adopt techniques such as management by exception. Project sites with abnormally good or bad performance will distinguish themselves. The former can be examined in detail to learn characteristics of successful projects; the latter can become the focus of special correction measures. In the final analysis, the type of built-in information system discussed above will promote adherence to principles such as goal-oriented management and evaluation.

## Operational Research

Complementary to promoting built-in monitoring and self-evaluation systems, there is a need to identify and test a variety of underlying assumptions that form the foundation of food aid projects. Despite the enormous opportunity cost of project food aid, in both donor and recipient countries, many of the basic premises upon which projects are based have not been examined. This problem stems from the historical, although changing, perception of food as a surplus commodity; it was simply not appropriate to spend significant human and financial resources to guide its disposal. Now that food aid is recognized as a valuable resource, the time has arrived to do the necessary, although belated, research.

The purpose of operational research, like most research, is not defined in terms of improving the welfare of the study community; it is not even intended primarily to determine whether or not a given project achieved its objectives. Instead, the task is to garner generalizable information that expands the boundaries of knowledge concerning the potential uses of food aid. Interest should be focused on hypothesis-testing in an operational environment. The viability and replicability of a given intervention strategy should be assessed. The insight gained is of primary use for policy formulation at the highest levels of programme responsibility.

Operational research should be directed towards identifying and testing a variety of hypotheses in a select few project sites. The precise questions to be asked should be formulated by policy-makers within WFP and Food-for-Peace, with consultation from appropriate experts. *The research should examine the assumptions* that link each level of the logical progression of project activities and events that are thought to bring about expected impact.

To amplify, there are many assumptions linking inputs to the achievements of outputs for typical feeding projects. These are relatively straightforward. The verification of these assumptions (as well as that inputs were provided and outputs achieved) is the domain of built-in evaluation systems discussed above. Data collection on these processes should be undertaken at all project sites. In contrast, the achievement of project purposes and goals is based on linking less apparent assumptions. For example, an assumption linking outputs to purposes for a supplementary feeding project is that the ration will not be shared with other siblings or substituted for commodities already provided. For a food-for-work project, an assumption linking outputs to purposes is that increased mobility of labour markets and greater integration of commodity markets that result from infrastructure development will increase employment opportunities for

the poor and improve household food security. Verifying these assumptions, which form the theoretical basis for projects, is beyond the scope of a built-in monitoring and evaluation system. Rather, they must be the subject of empirical research performed in an operational setting.

A few important qualifications are required. First is that *not all* assumptions at the purpose or goal level must be tested. Only strategic questions that remain a source of concern or controversy should be addressed. Second is that, in order to test hypotheses, it will often be necessary to rely on data concerning project impact — i.e., objectively verifiable indicators at the purpose or goal level. Ideally, if a project is functioning well, much of the data should be available as part of a built-in information system. However, special surveys or data collection on impact may be required. It is emphasized that the data requirements for operational research should be clearly defined.

To date, a variety of activities undertaken by Food-for-Peace and WFP, as well as other agencies, fall within or near the boundaries of hypothesis-testing through operational research. There are some excellent examples found in studies now under way. These include determining whether the marginal propensity to consume food out of in-kind income differs from that with wage income (6, 7), the development of a simulation model to test the cost-effectiveness of various targeting strategies (26), examining whether and to what extent random measurement error results in under-reporting of project impact (20), and the proposed testing of the theory underlying the recommendation that the choice of food aid commodities be selected on the basis of maximizing the value of the income transferred to the recipient (27). Similarly, past studies such as Project Poshak, Ivarangval, and CARE Phase II and III studies were attempts to examine some fundamental assumptions that form the conceptual foundations for the nutritional impact of food aid projects. In all these cases, decision-makers formulated clear compelling questions. Thereafter, they called on researchers to find answers in an operational environment.

The proposed WFP "in-depth" evaluation studies only partially fall within the domain of the type of operations research discussed above. For example, the well-designed Honduras study (8) is framed around determining the impact of a specific project. Undoubtedly, some generalizable knowledge will be gained. But the motivating question is whether the project improves nutritional status. Although important for justifying budgets, a well-functioning built-in evaluation system would do equally well in answering this question. This is especially so given the differing opinions concerning how well the

project itself is operating and the high cost and difficulties of performing social experimentation in this field.

In-depth research should instead focus on examining the linking hypotheses. If these can be substantiated, it will enable one to assume, with confidence, that if the services are delivered as planned, and the achievement of outputs substantiated, there is little question that the project will reach nutritional objectives. The time is right to draw up an operational research agenda, mutually agreed upon by donors and recipients alike. This should give co-ordinated action. Doing so will provide the knowledge to use food aid intelligently. The scope and potential for achieving nutritional impact will become more clear. This, coupled with the development of built-in information systems, should form the organizing theme for moving forward on the issue of evaluating the nutritional impact of food aid projects.

## REFERENCES

1. C. Williamson-Gray, "Food Consumption Parameters for Brazil and Their Application to Food Policy," Research Report No. 32 (International Food Policy Research Institute, Washington, D.C., 1982).
2. O. K. Knudsen and P. L. Scandizzo, "The Demand for Calories in Developing Countries," *Amer. J. Agric. Econ.*, Feb. 1982.
3. C. P. Timmer and H. Alderman, "Estimating Consumption Parameters for Food Policy Analysis," *Amer. J. Agric. Econ.*, 61: 5 (1979).
4. D. Saha, "An Analysis of Food Consumption and Expenditure Patterns in Sri Lanka" (International Food Policy Research Institute, Washington, D.C., and Community Systems Foundation, Ann Arbor, Mich., USA, 1984).
5. S. Kumar, "Impact of Subsidized Rice in Food Consumption and Nutrition in Kerala," Research Report No. 5 (International Food Policy Research Institute, Washington, D.C., 1979).
6. M. P. Demousis, M. W. Harrell, and D. C. Franklin, *Income Effect of Title II Commodities in Rural Panama* (Sigma One Corporation, Raleigh, N.C., USA, 1982).
7. E. Kennedy, "Determinants of Family and Preschooler Food Consumption," *Food and Nutr. Bull.*, 5 (4): 22 (1983).
8. World Food Programme, *Research Proposal: Process and Impact Evaluation of World Food Programme Assisted Project Honduras 2523* (World Food Programme, Rome, 1983).
9. E. B. Hogan, "Scope of Work for a Study of Potential Improvements in the WFP Programme System" (World Food Programme, Rome, 1983).
10. W. D. Drake, J. Gunning, A. Horwitz, R. M. Iler, H. Rice, and G. Thenabadi, "Nutrition Programs in Sri Lanka Using U.S. Food Aid (An Evaluation of P.L. 480 Title II Programs)" (Community Systems Foundation, Ann Arbor, Mich., USA, 1982).
11. J. W. Gilmore, C. Adeiman, A. Meyer, and M. Thorne, "Morocco: Food Aid and Nutrition Education," AID Project Impact Evaluation Report No. 9 (US Agency for International Development, Washington, D.C., 1980).
12. S. F. J. J. E. Flores, M. B. Flores, M. Mack, M. Norton, and R. Pooley, "The Impact of P.L. 480 Title II in the Philippines, 1970-1980" (US Agency for International Development, Washington, D.C., 1981).
13. R. C. Weisell, "Study on the Nutritional Status Changes between 1979 and 1981 among 1500 School Children in Mafraq and Madaba (Jordan) and the Possible Impact of WFP School Lunch Food Aid in Mafraq" (Food and Agriculture Organization, Rome, 1983).

14. A. Wijga, "The Nutritional Impact of Food for Work Programmes: Report of a State-of-the-Art Study" (Wageningen, Netherlands, 1983).
15. International Food Policy Research Institute (IFPRI) "Proposal for an In-Depth Evaluation of WFP-Aided Food for Work in Bangladesh Including Impact of CIDA and World Bank Financed Appurtenant Structure" (Bangladesh Institute of Development Studies, Dhaka, Bangladesh, and IFPRI, Washington, D.C., 1980).
16. R. Miller and D. Sahn, "Built-in Evaluation System for Supplementary Feeding Programmes - Why and How," in D. E. Sahn, R. Lockwood, and N. S. Scrimshaw, eds., *Methods for the Evaluation of the Impact of Food and Nutrition Programmes*, Food and Nutrition Bulletin Supplement 8 (United Nations University, Tokyo, Japan, 1984), p. 265.
17. R. Martorell, "Improved Nutrition and Its Effects on Anthropometric Indicators of Nutritional Status," *J Nutr. Rep. Internat.*, 21: 2 (1980).
18. J. P. Habicht, "Some Characteristics of Indicators of Nutritional Status for Use in Screening and Surveillance," *Amer. J. Clin. Nutr.*, 33: 3 (1980).
19. G. Beaton, "Energy in Human Nutrition," *Nutr. Rev.*, 41: 11 (1983).
20. W. D. Drake, R. Miller, and R. Timmons, "Effect of Classification Errors on Evaluation Outcomes: Nutrition Programs Revisited" (Community Systems Foundation, Ann Arbor, Mich. USA, 1984).
21. W. D. Drake, R. Miller, and M. Humphrey, *Analysis of Community Level Nutrition Programs*, vol. 1 (Community Systems Foundation, Ann Arbor, Mich., USA, 1980).
22. D. E. Sahn, "An Analysis of the Nutritional Status of Preschool-Age Children in Sri Lanka, 1980-81" (Community Systems Foundation, Ann Arbor, Mich., USA, 1983).
23. *NCHS Growth Charts* (US Public Health Service, Health Resources Administration, Rockville, Md., USA, 1976).
24. D. E. Sahn and P. Pestonk, "A Review of Issues in Nutrition Program Evaluation," AID Evaluation Discussion Paper No. 10 (US Agency for International Development, Washington, D.C., 1981).
25. J. B. Habicht and D. Dapice, "An Indonesian National Nutrition Surveillance System" (mimeo., Cornell University, Ithaca, N.Y., USA, 1984).
26. R. J. Timmons, "Targeting: A Means to Better Intervention" (Community Systems Foundation, Ann Arbor, Mich., USA, 1983).
27. S. Reutlinger, "More Effective Aid to the World's Poor and Hungry: A Fresh Look at U.S. Public Law 480, Title II Food Aid," World Bank Discussion Paper No. ARUB (World Bank, Washington, D.C., 1983).

## NOTE FOR THE RECORD: Comments on the Paper "Methods for Evaluating the Nutritional Impact of Food Aid Projects: Lessons from Past Experience" by David Sahn

Hossein Ghassemi

Senior Nutrition Adviser, UNICEF, New York

Dr. Sahn's paper pinpoints many problem areas in programming and administration of food aid. It states that past definition of objectives has not always been clear. There have been problems in programme design. Sometimes programmes are not based on a solid hypothesis. There have been poor implementation and various methodological problems in evaluation. The paper also concludes that evaluations have often been qualitative, incomplete, and inadequate.

The paper recommends two important lines of action: The first is to design and conduct a number of in-depth studies in order to gain a better insight into programme costs, effects, and proper indicators for evaluation. It also proposes building into ongoing projects a solid, carefully designed element of monitoring and evaluation and suggests that the additional costs of such action are entirely justified.

In reviewing the lessons in evaluation, there are several issues to be raised:

First, it is important to distinguish between good projects that were poorly evaluated and poor projects that had a favourable evaluation.

Second, we need to ask, Evaluation for what? Evaluation has not been approached broadly enough. In general, there is a problem of utilizing evaluation for various purposes. Evaluation can and should be conducted for three different reasons:

- to improve decisions on choice of policies and programmes,
- to improve efficiency and effectiveness,
- to bring advocacy for generating enough food aid for the needs of developing countries.

Evaluation activities in the past have concentrated more on the questions of efficiency and effectiveness and very rarely have been designed in relation to programme/policy choices, and they have certainly not been effectively utilized for advocacy purposes. As a matter of fact, the field of food aid in the international development community has evolved a rather bad image as a resource that has been poorly used, and some negative effects, such as disincentive for food production, have been of great concern. It is not often understood that food aid is a resource; it is the way it is used that makes the difference.

Therefore, the suggestion is that the field of evaluation should be clearly studied and directed towards meeting all three of these important needs. For major agencies in charge of programming and administration of food aid, there is an urgent need for better knowledge in programme policy choice and also for the whole area of advocacy.

Third is the question, Evaluation by whom? Most evaluation studies have been conducted by the donors or as a part of special studies. It is important that programme management strongly encourage the authorities in the countries to take monitoring/evaluation seriously into account in proper design and implementation.

Fourth, considering all the constraints and uncertainties raised by Dr. Sahn, what can be done now? I am raising this question with some specific suggestions that I believe are feasible and practical. What can we do now, based on the available knowledge and skills? The following immediately come to mind:

1. There is a need for better *integration of food aid projects into overall community development activities*. It is common practice now for food aid projects to take a rather isolated, individual approach and as an operational channel within the system. This implies that food aid is not being effectively considered as another resource available in the community that can be pooled and effectively used within the overall frame of priorities and programming in the community, especially in the areas of nutrition and health needs of vulnerable groups. Problems of objectives, implementation, participation, etc. all stem from lack of integration.

2. A great deal of improvement can be achieved in developing better, clearer, and *more concrete definitions of objectives* for the use of food aid.

3. There also needs to be *improvement in implementation*. Three options exist:

- a. Improve management capacity, logistics, targeting, supervision, etc.
- b. Shift responsibility for implementation to the community.

- c. Modify the programming approach. This option is obviously of a more long-term nature.

If one has to make a choice between strengthening programme management — i.e., trying to do a better job through more efficiency on the part of management — and shifting the responsibility to the recipients, it is quite likely that the second option promises much more hope for a better outcome and would be less costly. However, this is a hard decision to make.

With regard to modification of the approach to programming:

First, for quite some time suggestions have been repeatedly made to change the concept from food distribution to food as medicine, specifically targeting on advanced cases of malnutrition. In that concept, the health system has the responsibility of prescribing food but not necessarily distributing and administering it.

Second, food aid can be used through women's programmes and income-earning projects, especially those targeted on increased local food production. This is an approach that would turn the whole operation around and make it possible to use food aid for increasing food production rather than competing with it. This can also be effectively used for helping women to earn more income, and for reasons that I will explain later, this can have a very strong incentive effect on better design and implementation of primary health care/nutrition programmes.

Third, food aid can be used in a monetized form for supporting either the initial or the recurrent costs of community development programmes.

It is very clear that concentration on these three possibilities, among many, would change the design and operational aspects of food aid to a great extent and would give it a much more effective role as a major resource element in the total development efforts supported by the international community.

It is important to recognize that costs and benefits of a programme will be strongly influenced by the socio-economic circumstances relative to advancement in access, infrastructure, income, etc. Such circumstances in turn affect relative size of deprivation and therefore needs and the cost of meeting those needs, and also the magnitude and durability of benefits achieved. It happens that the most disadvantaged are the top priority for humanitarian reasons, but it would often cost more to meet their needs. Hence, it would mean the same benefits for higher costs. Also be aware that small input in relation to large need (e.g., food gap or calorie gap) is very unlikely to produce a measurable impact. This is where the resources are

spread too thin. Such cases should be seen as resource allocation, project design, and implementation problems and not as evaluation problems, unfortunately, they are not rare.

A very important point is that the annual cost of eradication or control of hunger in the world has often been cited as five billion dollars and, as the discussion in the Sahn paper indicates, the value of food aid currently available for the specific needs of vulnerable groups such as mothers and children and for nutritional purposes easily adds up to about one billion dollars per year. This is a substantial resource, and proper use of it could have a major impact if we are willing to take a serious look at the possibilities of better use of it and to be prepared to make hard decisions.

Among the very low-income households, there is substantial evidence that programmes that tend to effectively reduce poverty at the household level — such as income-generating activities — would be extremely effective in helping the poor to elevate themselves above the poverty line. Such improvements would then make them much more responsive and active participants in organization and operation of nutrition/primary health care activities in the community.

Therefore, use of food aid in cash (through monetization) for income generation, together with local food production with particular focus on women, offers excellent potential for an integrated, development-oriented approach to nutrition improvement.

Hard decisions must be made in the following areas:

1. Strong efforts are needed to change the programming approach.
2. More money should be spent on evaluation and promoting evaluation in participating countries. There have been many arguments that one should be very careful about the amount expended for evaluation. Sometimes the administrators are hesitant to consider using major portions of resources for evaluation, arguing that this is actually the

food or resource being taken from the needy. This is a false argument because there is enough evidence that spending an additional 1 or 2 per cent of the total resources on evaluation would contribute to improvements in the whole understanding of programme policy choices, in efficiency, and especially in advocacy, that should help improve the image of food aid and the potential for more flexible aid. In both the short and the long run the return would be much greater than the cost.

In this context, it is also important that major attention be given to generating enough political will among the governments to prepare the overall climate for such major changes.

3. Decisive initiatives should be taken to sharpen the focus on the purposes of evaluation, improvement of project design, and evaluation for advocacy.
4. There is need for support of in-depth studies. It should be made very clear that in-depth study is not specifically designed for better evaluation of projects. It is a complex exercise designed to develop better, simpler, and more practical methodology at the field level, and this is an expert attempt to improve the methodological aspects of evaluation, which is quite different from assuming that in-depth evaluation at much higher costs is needed to evaluate every project.
5. Finally, it is important to differentiate between what is possible now and what needs to wait for the future. For example, in the area of evaluation at the project level, it is much more practical at this point to improve and emphasize the process evaluation aspects of the project and give more time to more definitive and effective impact evaluation. This is primarily because of some methodological problems, especially in the area of indicators, that need some time for clarification of a number of controversial issues. At this time, impact evaluation should focus on the best use of reliable methodology and wait for necessary improvements in order to add to the precision of impact measurement.