AGENCY FOR INTERNATIONAL DEVELOPMENT
WASHINGTON, D. C. 20523
BIBLIOGRAPHIC INPUT SHEET

FOR AID USE ONLY
Batch 64

1. SUBJECT CLASSIFICATION
A. PRIMARY
Food production and nutrition — AS00-0000-0000
B. SECONDARY
Human nutrition

2. TITLE AND SUBTITLE
Evaluation methods for child feeding projects in developing countries,final report: Chapt. II,Perspectives for program officials

3. AUTHOR(S)
(100) Lerner,H.A.; Manly,R.P.; Eighmy,T.H.
(101) Checchi and Co.

4. DOCUMENT DATE	5. NUMBER OF PAGES	6. ARC NUMBER
1977	53p.	ARC

7. REFERENCE ORGANIZATION NAME AND ADDRESS
Checchi

8. SUPPLEMENTARY NOTES (Sponsoring Organization, Publishers, Availability)
(Chapts. I,21p.: PN-AAE-073; III,12lp.: PN-AAE-075; IV,116p.: PN-AAE-076)

9. ABSTRACT

Improved methods of evaluating supplemental feeding programs are needed. To test the efficacy of a cross-sectional survey methodology for evaluating supplemental feeding programs, a pilot study was conducted of two types of programs at five sites in each of three countries: Colombia, the Philippines, and Kenya. The study results were ambiguous. Thus the methodology tested is not recommended as an effective means of evaluating the efficiency of supplemental feeding programs. This chapter of the report on the study is written for program officials. It explains the evaluation methods the report recommends for future studies, identifies key issues involved in those methods, and describes some important technical concepts in nontechnical terms. This chapter provides a broad perspective on evaluation studies; defines the specific orientation of evaluation methods recommended; and lists the criteria considered in selecting those methods. It also discusses A.I.D.'s logical framework for program officials outside of A.I.D. and presents a logical framework designed as a master charter for the types of evaluations recommended in this report; it discusses the analysis of linkage processes in evaluating child feeding operations; examines the use of multivariate analysis and other statistical techniques; and provides a retrospective view of the chapter.

10. CONTROL NUMBER	11. PRICE OF DOCUMENT
PN-AAE-074	
12. DESCRIPTORS Children Evaluation Feeding Project planning Technical assistance	13. PROJECT NUMBER 14. CONTRACT NUMBER AID/CM/otr-C-73-199 GTS 15. TYPE OF DOCUMENT

AID 590-1 (4-74)

PN-AAE-074

FINAL REPORT

EVALUATION METHODS FOR CHILD FEEDING PROJECTS IN DEVELOPING COUNTRIES

CHAPTER II

PERSPECTIVES FOR PROGRAM OFFICIALS

A report for the
Office of Nutrition
and the
Office of Development Program Review and Evaluation
U.S. Agency for International Development

CHECCHI AND COMPANY
1730 Rhode Island Avenue, NW
Washington, DC, 20036, USA

AID Project No. 931-11-560-879
Contract No. AID/CM/otr-C-73-199 (5)

March, 1977

TABLE OF CONTENTS

Detailed Tables of Contents and Lists of Exhibits precede each chapter of this report.

PREFACE

This report was funded by the United States Agency for International Development (AID) through a contract with Checchi and Company. The pilot survey research was carried out in collaboration with Consultal, S.A., Bogota, the Institute of Philippine Culture, Manila, and the Research Bureau (East Africa), Ltd., Nairobi.

Previous working papers have specified the research design and procedures for the pilot survey, presented the survey instruments, discussed the individual sites in the three nations, and provided initial analysis of the survey data.

The principal authors of the present report are Dr. Thomas H. Eighmy, Mr. Harvey A. Lerner, and Mr. Robert P. Manly. Dr. Eighmy carried out the pilot survey data analysis. He wrote Chapter Four, presenting results of that analysis. Mr. Lerner was responsible for project management. He was principal author of Chapter Two on evaluation perspectives. Mr. Manly was co-investigator with Dr. Eighmy for the data analysis. Mr. Manly's work led to the derivation of Weight Dispersion Measures (WDM) and to the development of the Tables in Appendix A of Chapter Three. Mr. Manly also prepared the cost analysis presented in Appendix B. Chapter One and Chapter Three were collaborative efforts.

The authors wish to express their appreciation for valuable assistance and comments provided by Mrs. Patricia McPhelim, Mr. Harry Carr, and Mrs. Vicki MacDonald. The authors also wish to express their appreciation to Mr. Cae Johnston who typed the manuscript, prepared the exhibits, and was responsible for report production.

CHAPTER TWO

EVALUATING CHILD FEEDING OPERATIONS: PERSPECTIVES FOR PROGRAM OFFICIALS

CHAPTER TWO

EVALUATING CHILD FEEDING OPERATIONS: PERSPECTIVES FOR PROGRAM OFFICIALS

TABLE OF CONTENTS

Exhibit

CHAPTER TWO

EVALUATING CHILD FEEDING OPERATIONS: PERSPECTIVES FOR PROGRAM OFFICIALS

A. INTRODUCTION

This Chapter is written for program officials. It approaches evaluation from a management perspective. It explains the evaluation methods recommended in this report, identifies key issues and choices pertaining to these methods, and describes some important technical concepts in layman's terms.

A "logical framework" for evaluating child feeding operations is set forth in this Chapter, and is discussed at considerable length herein. This discussion is highly pertinent to the guidelines contained in Chapter Three. The logical framework is introduced in the present Chapter, rather than in Chapter Three, for three reasons. First, the framework should be used by administrators as a management tool in the course of chartering, overseeing, and using the results of evaluations of child feeding operations. Second, the framework itself provides a convenient and comprehensive outline for discussing the methods recommended in this report. Third, as the Music Man's competitors said, "Ya gotta know the territory." The evaluator's art should be neither arcane nor inscrutable to the program official. It should sharpen the administrator's good judgment, rather than overwhelm it. If evaluations are to serve program officials well, these officials must master the fundamentals. And that includes understanding the logical framework.

This Chapter is composed of seven sections. Section A is this introduction. Section B provides a broad perspective on evaluation studies. Section C defines the specific orientation of the evaluation methods recommended in this report and lists the criteria that were considered in selecting these methods. Section D contains some introductory observations on AID's logical framework, directed, in particular, to program officials outside of AID. Section E presents a logical framework specifically designed as a "master charter" for the types of evaluations recommended in this

report. Section F discusses the analysis of linkage processes in evaluating child feeding operations. Section G examines the use of multivariate analysis and other statistical techniques. Section H provides a retrospective view of the Chapter.

B. APPROACHES TO EVALUATION

"Evaluation" means different things to different people. The range of operational, administrative, and research questions which can be raised with respect to child feeding programs is very wide. The approaches to answering these questions may be quite varied. Thus, it is important to place the particular approach to evaluation contained in this report in a broad perspective.

There are at least four categories of approaches to evaluating child feeding programs in developing countries which contribute to understanding and/or action. These categories are not mutually exclusive. They are:

1. Basic Research is directed toward characterizing, classifying, and analyzing the phenomena associated with child feeding programs. Such research may be clinical or non-clinical in character, "longitudinal" or "cross sectional" in the survey methods it employs; and narrow or comprehensive in the range of disciplines which it utilizes. The distinguishing characteristics of a basic research approach are its use of advanced methods of analysis and highly trained researchers, on the one hand, and its objective of advancing the frontiers of knowledge, on the other. Such research may indeed yield practical insights and methodologies which can be replicated advantageously. But, such practical applications are not an explicit objective of this category and there is no firm assurance that such practical "spin-offs" are a likely outcome.

2. Applied Research is explicitly directed toward improving methods and measures used in assessment of ongoing child feeding programs. Here the attempt is to develop the practical art of evaluation. Obviously, there is a relationship between improving knowledge and improving evaluation. However, in the latter case, the emphasis is on uniformity, comprehensibility,

cost-effectiveness, and broad usefulness. The task of applied evaluation research is to identify, discover, and select evaluation tools which will be used by others to provide information and interpretation for projected direction. The task employs sophisticated knowledge to distill a mandatory degree of simplicity from a high degree of complexity.

3. Creation of Practical Evaluation Guidance includes the development and testing of forms and instructions for evaluators. This process seeks to make use of the results of applied research and to combine them with existing evaluation methods in order to provide an improved "standard approach" in the light of the state of the art, the needs of program officials, and the realities of program assessment in developing countries. This third category can be viewed as an elaboration and application of AID's Project Evaluation Guidelines.[1/] In that document, "evaluation" is defined as:

> "Measurement and comparison of actual progress vs. prior plans, oriented toward improving plans for future implementation. (Emphasis supplied.) It is part of a continuing management process consisting of planning, implementation and evaluation; ideally, each phase follows the other in a continuous cycle until successful completion of the activity." 2/

This definition of evaluation includes within its scope the tasks of designing or redesigning programs. Such design functions should, however, be distinguished from full-scale feasibility studies. Project evaluation and design can, and should, draw on the results of feasibility studies where such studies have been performed and are suitably oriented. In any event, evaluations can draw on concepts of cost/effectiveness and cost/benefit analysis on which feasibility studies are premised.

1/ See Project Evaluation Guidelines, Third Edition, Office Program Review and Evaluation, Agency for International Development, (Washington, D. C., August, 1964).

2/ Ibid., p. 36

4. Feasibility Studies are designed to provide a cost/benefit type analysis of a project before that project is installed. A feasibility study usually represents a substantial and formally structured kind of undertaking, one which employs a well-established set of techniques and conventions (e.g., shadow pricing, identification of external benefits, residual value calculations) and project feasibility criteria (internal rate of return, net present benefits, and/or cost/benefit ratios). Such techniques and criteria are believed to measure the relative economic and social merit of one project in comparison with others in the same, or different, sectors. While such studies are frequently quite thorough and elaborate, they constitute predictions of the future, oriented to "go/no go" decisions. They are seldom designed to be directly applicable to the kinds of evaluations described in Category 3: Creation of Practical Evaluation Guidelines.

C. ORIENTATION OF THIS REPORT

Each of the four categories described in the preceding paragraphs can make contributions to the art of evaluating child feeding programs in developing countries, and each is pertinent and important in its own right. However, the central focus of this report is on Category 3, Creation of Practical Evaluation Guidelines, which includes the present Chapter Two and Chapter Three--Evaluation, Guidelines, Forms, and Procedures--and on Category 2: Applied Research, which includes the analysis of Pilot Survey Data in Chapter Four.

The approach contained in this report was developed to meet the following criteria:

- It must serve the purposes of the kinds of evaluations defined in AID's Project Evaluation Guidelines.
- It must embody the thrust of the research on which it was based; namely, to test better methods of measuring the effects of child feeding operations and to use such methods for purposes of evaluation.
- It must be suitable to conditions and data existing in developing countries.

- It must be capable of being carried out by evaluators trained in survey research and quantitative analysis to the Master's Degree level, at a maximum.

- It must allow evaluation results to be presented in a form that is useful to, and readily comprehensible by, program administrators, and which are broadly comparable in similar situations.

- It should contribute to decisions of program officials which affect results of child feeding operations as part of a continuing management process consisting of planning, implementation, and evaluation.

These criteria are necessarily restrictive. They rule out, as well as rule in, a range of interesting and potentially instructive approaches to evaluating child feeding programs. It should be understood that what has been ruled out for purposes of the guidelines presented in this report, may be most useful for other purposes and in other contexts. A balanced research strategy which recognizes the importance of each of the four categories, and encourages cross-fertilization among such endeavors, will produce the best long-term results. However, it is critically important to such a process that advances in the state of the art be translated into practical applications, even as such advances open new opportunities to better understand the child feeding programs in developing countries through further research. The keynote of this report is that of practical applications.

The materials presented in this report focus on the creation of practical evaluation guidelines by combining three fundamental techniques for evaluating child feeding projects: (1) AID's "logical framework" approach, which structures a project in terms of a series of hypothesized causal linkages; (2) a systems analysis approach, which identifies relationships and describes the processes involved in these linkages; and (3) statistical analysis, which deals with important associations among specific variables in a supplementary child feeding situation.

Program officials who initiate, manage, and use the results of child feeding projects need not become experts in these techniques. But they should understand their respective rationales, advantages, and limitations. Such officials should be sufficiently knowledgeable to distinguish hard quantitative results from interpretative judgments, and interpretative judgment from speculation. The following sections of this Chapter provide a brief discussion of each of the three fundamental techniques employed in this report.

Section D provides a brief description of AID's logical framework for project evaluation. This brief description is addressed principally to program officials in developing countries and to others who are not familiar with the use of the framework. Section E presents a basic logical structure for child feeding projects. Section F describes a systems analysis approach to the various causal relationships identified in this structure. Section G describes the use of statistical analysis to examine measures believed to be pertinent to these relationships.

It should be understood that the logical structure and the accompanying discussion and guidelines represent one of several approaches to evaluation which fit within the rubric of Category 3: Creation of Practical Evaluation Guidelines.[1/]

D. THE "LOGICAL FRAMEWORK": SOME GENERAL OBSERVATIONS

This report makes a special, and somewhat restricted, application of AID's logical framework to the circumstances of child feeding projects in developing countries. The use of AID evaluation concepts does not thereby limit the recommended methods of evaluations conducted by, or under the auspices of, AID. The approach described in this report can be applied to evaluations conducted by host country governments, voluntary

---

1/ See Chapter III, Evaluation Studies, Evaluation Handbook, Second Edition, Agency for International Development (Washington, D. C.), pp. 19-32.

agencies, and other organizations with interests in child feeding programs. On the other hand, familiarity with AID's approach can be of considerable help to non-AID officials who are asked to participate in AID evaluations, or in evaluations conducted by other international development agencies employing similar evaluation methods.

Exhibit 2.1 presents a schematic representation of the structure of a technical assistance project. That structure embodies AID's logical framework.[1/] The logical framework has been formulated as a set of causal hypotheses, as follows:

-- If adequate inputs are provided in Item C-1 on Exhibit 2.1, then planned outputs (Item C-3) will be produced.

-- If these outputs (Item C-3) are produced, then purpose (Item C-5) will be achieved.

-- If purpose (Item C-5) is achieved, then a planned degree of progress toward a higher goal (Item C-7) will occur.

AID guidance on evaluation constitutes an important resource for program officials outside of AID who may be concerned with evaluating child feeding programs. These materials place considerable emphasis on collaboration with host government officials in preparing evaluation materials and in conducting evaluations. Nevertheless, the AID guidance has been prepared principally for its own staff and is oriented toward AID's own administrative procedures. The following observations are addressed to non-AID officials to whom the logical framework approach may be unfamiliar, or a bit disconcerting.

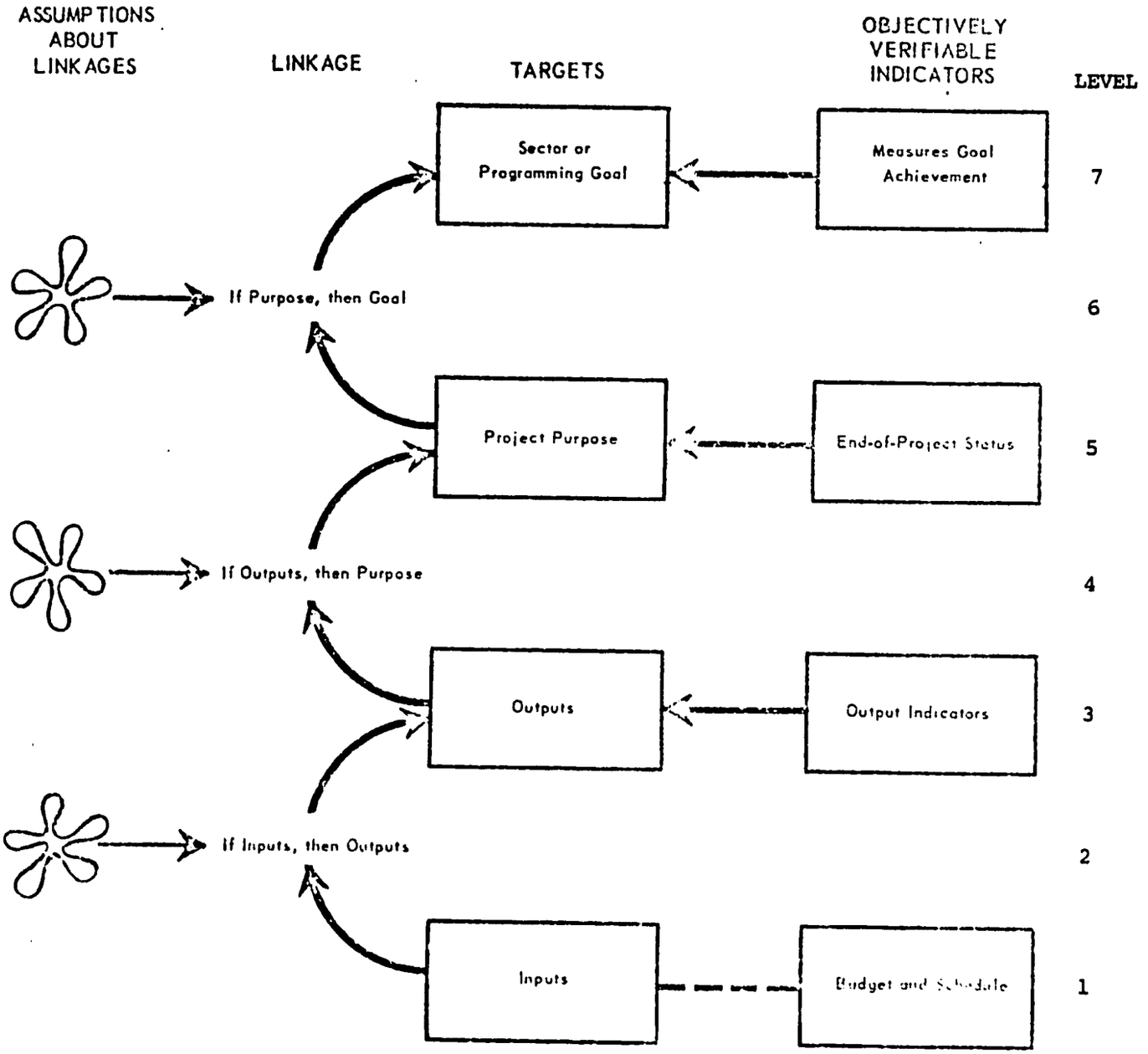
First, the framework, as presented in Exhibit 2.1, represents "a clean slate." It is a neutral instrument which can be used in a number

---

1/ Instructions for the use of the logical framework by AID officials are contained in The Evaluation Handbook, Project Evaluation Guidelines, and related publications.

Exhibit 2.1

THE LOGICAL STRUCTURE OF A TECHNICAL ASSISTANCE PROJECT



Source: Project Evaluation Guidelines, Third Edition, Office of Program Review and Evaluation, Agency for International Development, Washington, D. C., (August, 1964), p. 8.

of ways by those who apply it. As this "clean slate" is filled in, the objectives, professional experience, value judgments, and priorities of the authors (and/or the organizations they represent) are inevitably reflected in the result. When the "slate" has been filled, it is seldom neutral. It becomes an instrument for assuring that intended results are achieved and/or making changes in program design.

Second, the reality of a technical assistance project is frequently that it represents a synthesis of somewhat differing objectives explicitly (or implicitly) pursued by donors and recipients, central headquarters and field establishments, and a variety of functional units within government voluntary agencies, or other organizations concerned with a project. With the best of collaborative intentions, it is often the case that a project design ends up as the work product of a few individuals or units. In the end, a meaningful and useful logical framework must crystallize a consistent viewpoint. Such a viewpoint will subordinate some considerations to others. In the end, the creator of the framework may confront the necessity of imposing a single rationale in order to provide a clear basis for evaluation. If his design fails to resolve fundamental issues, the designer creates the equivalent of a "constitutional compromise." In effect, he "passes the buck" to others who must deal with these unresolved issues at later stages in the process. This he should not do. As a result, the "final" version of a good logical framework may capture neither the full range of the actor's motivations which generated the project in the first place, nor the synthesis of partly conflicting, partly compatible, values and viewpoints which animate its operation.

Third, just as evaluation itself should challenge all aspects of project design, so, too, should a program official challenge each element of a logical framework prepared for a project for which he is responsible. The framework should be examined for its suitability to the objectives of the particular kind of evaluation which is being conducted. The framework provides an opportunity to impose rigorous discipline, if it is tailored to specific evaluation requirements. He should feel free to suggest, or make,

changes in existing frameworks. The decision to commit resources to evaluation becomes much easier when the official responsible for that decision is comfortable with the project rationale which the evaluation seeks to apply, and is convinced it will provide him with the information and insights he needs to perform his functions in an outgoing management process.

Fourth, it should be clearly understood that most development projects are conceived, designed, installed, managed, evaluated, and brought to completion in the midst of substantial uncertainty, and often without continuity of personnel who can be held "responsible" for results. As a given framework is applied to the complex realities represented by human undertakings, the "assumptions about linkages" may become very large indeed; the "objectively verifiable indicators" may turn out to be very limited in number and reliability; and the process of evaluation may emerge as essentially an exercise in disciplined guesswork. The fact that the evaluation framework is modeled on the analogy of the classical scientific experiment does not mean that the extent of quantitative certitude sought under laboratory conditions is attainable, or expected, in evaluations. Rather, it means that evaluation should make a substantial effort to quantify whatever it is sensible to quantify, and seek to apply a reasonable degree of insight and objectivity to the rest.

Fifth, while evaluations should seek to be reasonably dispassionate and objective, they may properly be designed as purposive instruments for project control and administrative improvement. Like some inspections and audits, evaluations can cause an organization to put its house in order and to catch up on activities which have been allowed to slide. Evaluations also can clarify an organization's objectives and policies, encourage resumption of neglected planning activities, and provide an occasion for making needed changes. The extent to which a given evaluation is carried out as a disinterested search for truth, and the extent to which it is used as a purposeful instrument of administrative discipline and change, is very much a matter for management judgment.

Sixth, some elements of, and linkages in, the logical framework will be more important than others for purposes of a given evaluation. The framework is comprehensive, and the range of issues and questions which can be raised with respect to a given project will be quite varied. On the whole, however, more is to be gained from focussing evaluation efforts on a few important issues than by spreading them thinly over the entire range of possible inquiries.

In summary, there is considerable latitude for management judgment and flexibility in the creation of a logical framework and its application in project evaluation. Program officials should feel free to challenge the premises of any given framework, or a particular application of that framework in a given evaluation at project inception.

E. A "LOGICAL FRAMEWORK" FOR EVALUATING CHILD FEEDING PROJECTS

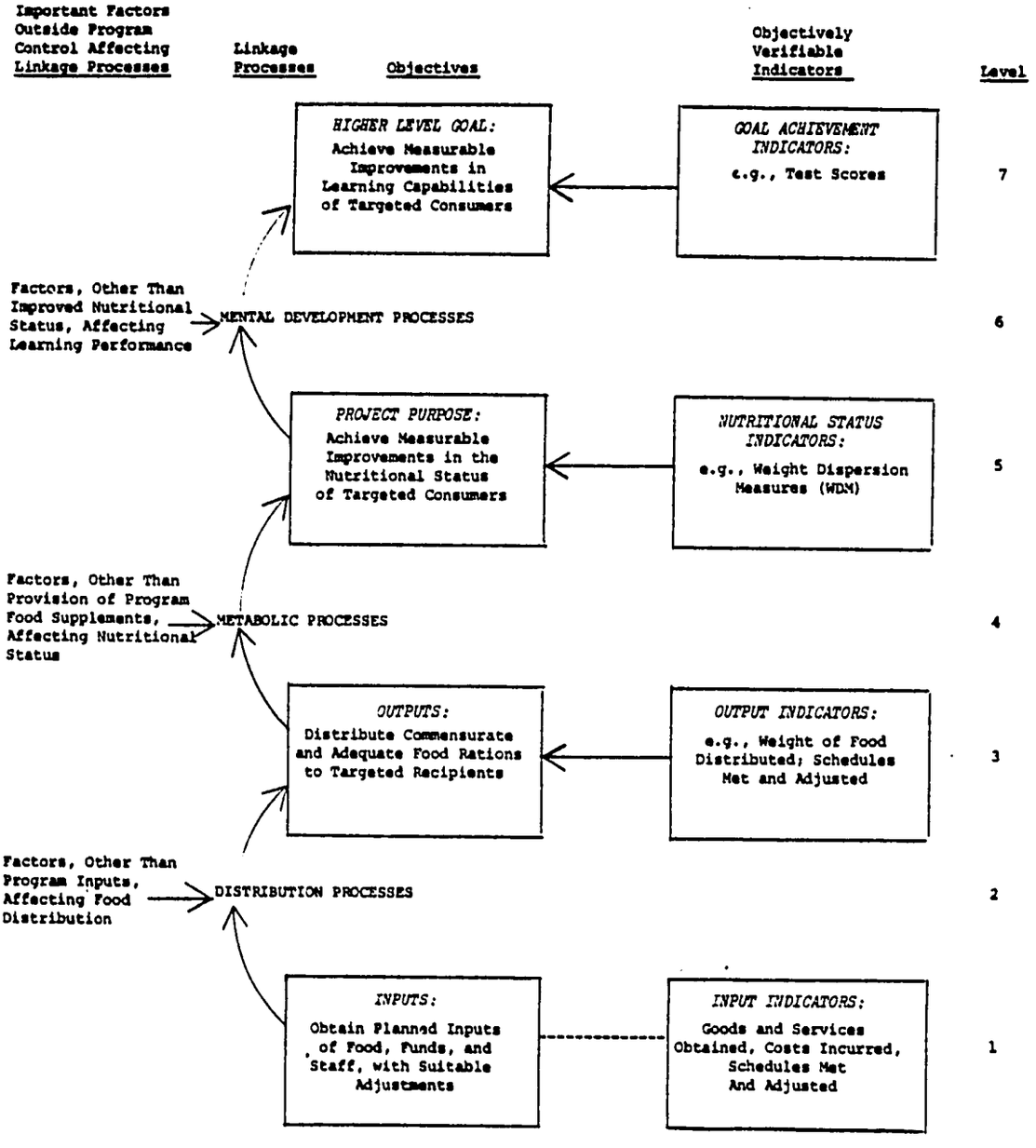
This section presents a suggested logical structure for evaluating child feeding projects. The guidelines presented in Chapter Three of this report are based in substantial measure on this suggested structure. Therefore, any official who concludes that the evaluation logic suggested in the remainder of this section should be revised in specific particulars should have his staff re-examine the guidelines presented in Chapter Three, in the light of such particulars.

We repeat here the point that was made in the previous section: AID's Evaluation Handbook, its Project Evaluation Guidelines, and related publications, are important resources to officials outside of AID who may be concerned with child feeding programs. A reading of these materials, while not mandatory to the understanding of this section, can illuminate the ensuing discussion significantly..

Exhibit 2.2 is the completed form of Exhibit 2.1. It presents a suggested framework for evaluating a child feeding project. The adjustments made in the row and column descriptions in Exhibit 2.2 do not alter the logical framework rationale. Rather, they provide a formulation of the framework which is readily compatible with the systems analysis which is

Exhibit 2.2

THE LOGICAL STRUCTURE OF
A CHILD FEEDING PROJECT:
A SUGGESTED FRAMEWORK FOR EVALUATION



Source: Checchi and Company, 1976. Adapted from Project Evaluation Guidelines, Third Edition, Office of Development Program Review and Evaluation, USAID, Washington, D. C. (1974), p. 7.

described in Section F of this Chapter, and with the statistical analysis which is described in Section G.

******

The logical progression contained in Exhibit 2.2 can be examined as follows:

-- If planned inputs of funds, food, and staff services (Item C-1 in Exhibit 2.2) are provided to a project, then intended recipients will be provided with commensurate and adequate outputs in the form of food supplements (Item C-3).

-- If recipients are provided with adequate outputs in the form of food supplements (Item C-3), the project purpose of improving the nutritional status (Item C-5) will be achieved.

-- If the project purpose of improving recipients' nutritional status (Item C-5) is achieved, then a planned degree of progress toward the higher goal (Item C-7) of improving recipients' mental performance will occur.

The logical progression presented in Exhibit 2.2 embodies a series of conceptual judgments and practical choices which were made in the course of its development. These judgments and choices have been heavily influenced by experience gained in the course of a pilot study of 60 sites in Colombia, Kenya, and the Philippines (described in Chapter Four of this report). They also reflect experience gained in conducting a variety of types of evaluations of child feeding and other development programs. They combine concern with what "should be" the various elements of an evaluation in a theoretical (or ideal) sense with what these elements "can be"--given the state of knowledge, the state of information available, the purposes for which evaluations are conducted, and the environment in which they are carried out. These choices and judgments should be made explicit. Six principal points which are of significance to program officials are treated in the following paragraphs, in order of importance.

1. By identifying improved nutritional status of children as the central purpose of child feeding operations, Exhibit 2.2 treats other effects as secondary.

The project purpose (Item C-5 in Exhibit 2.2) represents the basic objective to which the project is directed. It is to be distinguished from the "higher level" or "sector or programming goal" (Item C-7) to which the project makes only a partial contribution. From the point of view of the program official, the project purpose is generally of much greater concern and much more closely related to his key management function than is the higher level goal. Project purpose represents a "solution" fully addressed to the problem with which the project is concerned. In the case of child feeding, this problem may be stated as the low nutritional status of substantial numbers of children in developing countries. Accordingly, Exhibit 2.2 defines the purpose in terms of a solution to that problem.

Item C-5 in Exhibit 2.2 states that the purpose of a child feeding project is to "achieve measurable improvements in the nutritional status of targeted consumers." This statement has a sharp cutting edge. As noted, it emphasizes nutritional effects on targeted children as the primary purpose of the project. It relegates to a lower order of importance such "secondary," "external," "indirect," "collateral," "derivative," or "unintended" effects as the following: school attendance, test scores, community involvement in the project, nutritional impacts on non-targeted individuals, diversion of program resources to non-program purposes, socio-political impacts, effects (positive or negative) on local agricultural production, employment created by the project or by project expenditures, and income multiplier effects. This is not to say that such effects are not interesting, important, and (in some cases) even decisive. It is to say that they are not as important to an assessment of achievement of purpose as determining whether the project is actually improving the nutritional status of the targeted children, which is the central thrust of the applied evaluation procedures recommended in this report. Four practical implications of this point should be clearly understood.

First, the purpose statement in Exhibit 2.2 draws an important distinction between those who are intended to receive food ("targeted recipients") and those who are intended to consume food ("targeted consumers"). In the case of school feeding programs, the two groups are the same--children who are supposed to receive and eat program rations at school. In the case of MCH programs, one group is usually a subset of the other. MCH mothers are the intended recipients of food. MCH mothers and certain of their children (generally under five years of age) are the intended consumers. To recapitulate, at the "purpose level" (Level 5) of Exhibit 2.2, we are not merely concerned with whether the food rations are given to the persons who are supposed to receive them. Rather, we are concerned with whether the nutritional status of the persons who are supposed to eat them actually improves.

Second, the statement of purpose in Exhibit 2.2 draws a clear distinction between targeted and non-targeted persons who consume food. It does not accord the same level of significance to "nutritional benefits accruing to non-targeted people consuming a program food" that it does to nutritional benefits for the targeted population. At the purpose level, the evaluator does not distinguish between a "good" diversion of food from the targeted consumer to non-targeted consumer (as in the case of the seven-year old who pockets his nutribun and gives it to his more seriously under-nourished four-year old brother at home), and a "bad" diversion (the village chief who commandeers and eats a nutribun as his due). Rather, the evaluator asks whether the purpose of the project design is being carried out in terms of what is happening to those who are supposed to be eating the food. In attempting to determine whether project purpose is being carried out, he may weigh and measure the seven-year old to determine the child's nutritional status. The evaluator does not weigh and measure either the younger brother or the village chief. Conceivably, program officials might decide to alter the design of the project to permit transfer of food to, and consumption by, younger siblings, supervised, say, at the classroom door. That would put such siblings within target and purpose, and it would not necessarily violate the program concept.

With this change in targeting, the evaluator would seek to determine whether the health status of the fed younger siblings, in fact, improves. The case is different in the case of the village chief, taking his "tithe." Even if program officials were willing to accept this kind of behavior as inevitable, they could not "target" him as an intended beneficiary without doing violence to the basic child feeding rationale. Conceivably, the diversion could be treated as a normal distribution cost, or as an inventory loss within acceptable limits. Possibly, his behavior could be changed through punishment or education. However, it would be highly impractical to undertake an evaluation approach which would measure the increase in the chief's nutritional status, or indeed in his bank account, in an effort to capture the "totality" of program benefits or disbenefits.

Third, the definition of purpose contained in Exhibit 2.2 has special significance for MCH feeding programs. Here there is a particularly strong program rationale for providing supplemental food to the particular persons for whom it is intended: mothers and small children in very poor families. The intent of the program is to deliver such food to small children at stages in their development when nutrition is believed to be of critical importance. At the same time, there is considerable evidence that the program, in fact, operates as a "family feeding program," and that significant amounts of the food distributed is consumed by non-targeted members of the family.[1/] This is not to say that there is any practical way of assuring that small children receive all the MCH food intended for them. Nor is it to say that the way in which the poor family actually distributes the food is uninformed, irrational, or "wrong" in terms of the circumstances which it faces. It could well be that the common sharing of food is a practice essential to the preservation of the family as a social unit in its culture, or that--indeed--keeping the fathers and mothers in relatively good nutritional shape is more important to that society than avoiding the effects of malnutrition on its young

---

1/ MCH survey responses indicate that, on the average, about six people eat the ration.

children. It is possible that the food consumed by non-targeted members of the MCH family is needed by them as much as by the targeted children in the school feeding programs. Likewise, it is possible that the program is "cost effective" in terms of its targeted population if that population shows improved nutritional status despite diversions to non-targeted groups. Such diversion does not mean necessarily that food is wasted (spoiled or thrown away) if someone eats it. The diversions may be treated as "normal distribution system losses," even if they are practically impossible to track down, once the food reaches its intended recipient (the MCH mother). Even so, it would be inadvisable to revise the project design so as to legitimatize or encourage a practice contrary to the whole thrust of the program. The MCH program administrator and the evaluator may well be confronted with a dilemma stemming from the fact that the program rationale and the rationale which influences actual distribution of food within the family are different in their imperatives and values. This dilemma is essentially of the same nature as that faced by the person who undertakes to create a logical framework in the "collaborative" mode and discovers that the main actors are collaborating on a basis which enables them to maintain different orientations, priorities, and values with respect to the purpose of a project. The solution is the same in both cases. After it is determined that the viewpoints cannot be fully resolved, a fully consistent rationale is nevertheless chosen. The evaluator decides what level of effectiveness meets his criterion of project purpose under the circumstances. In the case of MCH programs, Exhibit 2.2 targets children as consumers. Officials who do not agree with this choice should make suitable adjustments in the framework and in the guidance provided in Chapter Three.

Fourth, if a program official wishes to secure information on secondary effects of child feeding operations, he has two basic choices. First, he can incorporate into the evaluation procedures a "basket category instruction" to evaluators, such as "report any data or insights you can obtain within the resources available to you on the following other

effects of the child feeding operation you are evaluating...". Or, he can mandate a rigorous approach to the data gathering analysis, and evaluation of information. The former approach is relatively inexpensive, but is likely to produce only impressions. The latter approach can be expensive, time consuming, and has little guarantee of being conclusive.

2. The purpose level of Exhibit 2.2 is particularly important because the recommended evaluation strategy makes an initial presumption that "program officials" are responsible for significant changes in the nutritional status of targeted children.

AID's Project Evaluation Guidelines describe the relative significance to management of the various linkages in the evaluation concept, as follows:

> "...the conversion of inputs to outputs--is presumed to be manageable, although the management is often very complex because of the joint provision of inputs and the subtle process of adapting imported technology. The degree of responsibility of the managers is greater for the production of outputs than for the achievement of purpose, since this achievement depends heavily on external influences beyond the control of the project personnel. (Emphasis supplied.) The responsibility of managers is even more attenuated for a goal. Evaluation is easier when managers realize that they will not be held accountable for all linkages, but will join with other interested officials to test the hypotheses that production of outputs will lead to achievement of purpose and that this achievement will contribute to the goal. The managers' responsibility about purpose is to recommend changes in outputs or purpose if the first plan is not working." 1/

This generally good advice is not fully applicable to the circumstances of child feeding operations. In the case of these operations, outputs have been defined as rations delivered to intended recipients as distinguished from rations ingested by intended beneficiaries. The evaluation approach presented in this report treats achievement of project purpose (i.e., improving the nutritional status of children) as the presumed responsibility of site officials and program personnel. Measures of nutritional status are, in fact, utilized to compensate for

1/ Project Evaluation Guidelines, op.cit., pp. 5-6

some practical problems in defining, measuring, and controlling the use of the program's outputs.

The principal evaluation strategy may be summarized as follows:

Sites are selected for repeated yearly evaluations, either on a "total universe" or sampling basis. Children are weighed and measured once each year at the sites by a measuring team. Such a team does not have to possess highly refined survey research or quantitative skills, and can perform its role rather quickly. The team uses its weight and height measurements to make on-site calculations of nutritional status indicators.[1/] The presumption is made that site operations should result in modest improvement in the nutritional status of targeted recipients over time (or, at least, maintenance of levels measured earlier). The presumption also is made that site administrators have a responsibility for the nutritional status of the children targeted to consume rations. If there are significant upward or downward departures from the expected pattern, persons responsible for child feeding operations are required to explain to the evaluators why such significant movements have occurred. Significant, and inadequately explained, <u>downward</u> movements in the indicators result in the arrival of a special assessment team with background in diagnosing problems of distribution management and administrative control. The first concern of this team is to determine whether the site has been receiving and properly distributing food rations as programmed, and whether measurements of nutritional status of targeted program beneficiaries have been properly made. The second concern is whether public health conditions, or other factors, can be identified as responsible for the ostensible decline in the nutritional status of the children measured. The remedies recommended by this special assessment team will depend on its findings. Such remedies certainly could include tighter administrative controls at the site or at

---

1/ Simple forms and clear instructions for performing necessary calculations are contained in Chapter Three.

one, or more, locations in the distribution chain through which food is delivered to the site, suggestions for improving health conditions, continued monitoring, and the like.

Significant, and inadequately explained, upward movements can also result in the arrival of a special assessment team which will include persons with such medical, sociological, statistical and administrative perspectives as may be suitable under the circumstances. In either case, the mandate of the special assessment team is to account for trends and determine whether they have been caused by normal fluctuation as a cohort passes through the supplementary feeding system, practices at the site, or by changes in the community served, which have implications for other sites. In certain circumstances, such a team might employ survey techniques or multivariate analysis as an aid to its analysis of these exceptional cases.

In the early stages of the project, it is desirable that there be an effort to establish national or regional norms and that a body of empirical evidence be accumulated. Until the establishment of such norms, the number of quality control inspections will be programmed in advance. In the first year, quality control assessments will be carried out at sites having unusually low or high absolute levels of nutritional status. In subsequent years, the concern will be changes in nutritional status.

This evaluation strategy has the following advantages:

- Because site officials know they have a presumptive responsibility for the nutritional status of the children targeted to consume the rations they distribute and because they know their practices may be subjected to scrutiny, the evaluation strategy provides a motivation for them to do the following:

  -- keep better records of food actually received and purchased, and the rations actually distributed.

  -- increase the efficiency and sanitation of site operations.

-- identify problems affecting the propensity of targeted consumers to actually eat the food provided.

-- identify community conditions causing health problems in the targeted population.

-- take other initiatives pertaining to site operations and/or community conditions which may improve the health of targeted beneficiaries.

-- provide a regular stream of quantitative and qualitative information useful to program planning, implementation and evaluation at higher administrative levels.

- It optimizes the use of evaluation resources by reserving the use of special skills to unusual cases, after sites have received initial screening.

- Ultimately, it may prove desirable to install some of the functions performed by the types of evaluation teams as regular site activities, subject to occasional inspection. The recommended strategy provides a body of experience and experienced personnel who could be utilized to install and monitor such a system.

- It results in the accumulation of objectively verified data pertaining to nutritional status over a period of years, and permits refinement of initial presumptions concerning anticipated impacts of rations provided by the site.

To summarize, this strategy relies on "project purpose" as an instrument for producing both information and change. It is not neutral or passive. It assigns presumptive responsibility to administrators who are not fully in control of the circumstances affecting the health of targeted children. And it relies, in the first instance at least, on measures of nutritional status which have been tested only within the confines of the data analyzed in this report and the limitations of the research design governing the analysis of that data.

In some respects, this is a bold strategy. It is one that could provoke resistance if it is not carefully introduced and administered with both circumspection and firmness. The program official considering the

use of this strategy should be aware that it will require continuing technical management monitoring if it is to produce the desirable results.

3. It is intended that significant reliance will be placed on the Weight Dispersion Measures (WDM) as objectively verifiable indicators of the achievement of project purpose, and in other connections as well.

Management decisions and evaluation judgments should ideally be based on measures which are reliable, easily measured, well understood, thoroughly field tested, and simply administered. They are to be distinguished from indicators which may be more pertinent from a theoretical viewpoint, but which do not meet these criteria. This report recommends three related "Weight Dispersion Measures (WDM)" for a variety of evaluation and management purposes. Given the circumstances prevailing in most developing countries, these measures meet each of the "ideal" criteria fairly well, with the exception of "thoroughly field tested."

A variety of weight dispersion measures was developed and applied in the course of the research described in Chapter Four. One stimulus for this work was the finding that measures of weight and height data were the most accurate, useful, and reliable of more than 150 variables on which data was gathered at 60 sites in three countries. Nevertheless, existing formulations of "weight-for-height" measures were found to be insufficiently sensitive for a number of research and evaluation purposes. Some innovative refinements had to be made.

The following paragraphs discuss:

-- Nature of weight dispersion measures;

-- Types of weight dispersion measures;

-- Advantages and limitations of weight dispersion measures; and

-- Applications of weight dispersion measures.

- Nature of Weight Dispersion Measures

Numerous efforts have been made to develop simple evaluation methodologies for measuring health or nutritional status on the basis of limited anthropometric techniques, which are applicable in situations where facilities for clinical and laboratory investigations are deficient or excessively costly in relation to the available resources.[1/]

A number of common indicators relative to growth and health of children are cited in the literature and range from biochemical stool analysis through as many as 18 physical measurements, including height and weight records. However, only a limited number of simple measurements of the human body appear to be suitable to practical application in field evaluations in developing countries.

Checchi investigated the technical problems of selecting an effective height-weight conversion formula for analysis of data gathered in field evaluations. Six equations for weight-for-age, height-for-age and weight-for-height, each paired by sex, were applied to the survey data and gave unsatisfactory results. It was observed in this study and others that in developing regions the age records are often incomplete and, in some communities, parents do not know the ages of their children--even to the nearest year.

The weight-for-height conversion formulas produced ratios of observed weights to reference weights for children's heights. The reference weights were based in part on those commonly referred to as Stuart-Meredith child height-weights. These benchmark data have been used in international studies and have seemed to work quite well.

1/ See, for example, A. E. Dugdale, "An Age-Independent Anthropometric Index of Nutritional Status," The American Journal of Clinical Nutrition: 24 (2/71) pp. 174-6.

The weight-for-height formulas, which are age-free, effectively set aside the immediate problem of poor age data. As a result, ratios derived from the above formula were adopted in the analytical phase of the study presented to AID at an earlier date.

The nature of the selected formulas gives a measure of overweight and underweight dispersion. However, equal weight dispersion in kilograms gives equal treatment above and below the reference norm. Consider, for example, two boys of the same height where one is a kilogram overweight and the second is a kilogram underweight. Both will have the same reference weight because they have the same height and sex. Now, if their height gives a reference weight of 25 kilograms, both will have the same four percent dispersion above or below the reference weight, 104 percent or 96 percent. The conversion formula is unskewed, or balanced.

During the present analytical work, it was observed that in the reference population data one kilogram below the reference weight of the 50th percentile spanned more individual children than one kilogram above the reference weight. In other words, the weight dispersion at the 25th percentile is less than that at the 75th percentile, so that the reference population is skewed, or unbalanced, in the direction of overweight.[1/]

Child sample populations in this three-nation study (and in other studies for India and Egpyt[2/]) show similar skewed characteristics in the direction of overweight. The conversion formula problem may be seen as a case of trying to measure a fundamental skewness using a non-skewed conversion formula that tends to mask underlying characteristics.

Thus, the second major step in the present study was to develop a suitably unbalanced, or skewed, conversion formula reflecting the skewed

---

1/ For a graphic view of this phenomenon, see Section C in Chapter Four.

2/ Height and Weight of Children in the United States, India and the United Arab Republic, U.S. National Center for Health Statistics, Series 3 Number 14, Rockville, Maryland, September, 1970, pp. 44, 46-47.

distribution of the "real world" children which could easily be applied in evaluating child feeding activities. This refinement has been achieved in the form of a ratio for which the numerator is the difference between observed and reference weights (taking a positive or negative sign) and the denominator is the difference between the reference weight and the associated 25th (underweight) and 75th (overweight) percentile weights. These factors are reduced to a simple columnar reference table, Appendix A, for field evaluation purposes. We call this measure the "Weight Dispersion Measure (WDM)." It reflects the skewness and assigns a higher dispersion measure to a child one kilogram underweight than to a child who is one kilogram overweight.

Weight dispersion data for the 25th and 75th percentiles in the population data generates elongated, bell-shaped curves as child heights increase. The bell curves are not entirely smooth because two sets of data were originally stratified by age groups and blended together for five-year olds. This gave rise to some reservations concerning the suitability of the reference population data. Fortunately, the questions thus raised are being dealt with at this time in work undertaken by the U. S. National Center for Health Statistics.[1/] This work smoothes the bell curves and confirms (on the basis of a large sample) the skewness in the direction of overweights observed in the pilot study data. The handbook working tables in our report should be revised on the basis of the new data.

Additional technical discussion of the WDM conversion formula is presented in Chapter Four. The WDM is applied in the guidelines presented in Chapter Three, for evaluation of child feeding activities in developing countries.

1/ P. V. Hamill, T. A. Drizd, C. L. Johnson, R. B. Reed, and A. F. Roche, NCHS Growth Charts, 1976, Draft, June 1976.

In summary, the weight dispersion measure (WDM) is a non-clinical anthropometric measure of child status. Such a measure is central to cost effective planning and evaluation of child feeding activities. As designed, the WDM allows a cross-sectional status evaluation of pre-puberty children both singly and collectively, according to height and weight. In addition, WDM is a convenient and low-cost method that is equally applicable in longitudinal time studies.

A group WDM directly indicates to the responsible officials and technicians the degree of potential group need for child feeding activities on a priority basis. The individual WDM, when extremely high (or low), suggests those cases in which clinical assistance may be helpful. For evaluators, a group's WDM serves as a baseline reference for quick and economical future measurements and evaluation of purpose and outputs.

A positive shift in WDM values from an initial negative position is an indicator of program benefits. The degree of such benefits relative to the WDM scale is not precise at this time and will largely depend on future testing.

- Types of Weight Dispersion Measures

Exhibit 2.3 shows three aggregate, or site level, weight dispersion measures derived from individual child measurements taken at 15 sites used as controls for school feeding operations in Colombia, the Philippines and Kenya. Since these sites represented schools where no child feeding programs were operating, the data obviously does not reflect results of program interventions. Weight dispersion data for all 60 sites surveyed is presented in Exhibit 4.19 in Chapter Four.

"WDMPROP" represents the percentage (or proportion) of the children with WDM higher than the 25th percentile of the reference population, which would show a value of 75. Thus, any number in the WDMPROP Column less than 75 indicates that a higher proportion of the site's children fell below the 25th percentile than in the reference population. In Exhibit 2.3, Colombia has two control sites which barely fit this description, Site 132 with WDMPROP of 74 and Site 152 with WDMPROP of 73. All of the Philippines control sites fall beneath the 75 level, as do four of the five control sites in Kenya.

Exhibit 2.3

WDM MEASURE
FOR FIFTEEN SCHOOL FEEDING CONTROL SITES

	Site	WDMPROP	WDMMED	WDMBAR
COLOMBIA	112	79	+007	+006.7
	122	95	+094	+086.7
	132	74	+001	-022.4
	142	82	+026	+026.1
	152	73	-041	-030.1
PHILIPPINES	212	31	-153	-123.6
	222	33	-123	-125.1
	232	62	-046	-054.8
	242	41	-136	-119.8
	252	59	-067	-080.7
KENYA	312	57	-093	-086.8
	322	46	-116	-107.5
	332	63	-051	-065.6
	342	83	-060	-076.3
	352	70	-060	-045.1

Taking weight-for-height as a measure of nutritional status, WDMPROP provides a simple, easily understood means of ranking sites within a country. In Colombia, Exhibit 2.3 shows that school control Site 122 ranks highest, with WDMPROP of 95. Site 153 is lowest with WDMPROP of 73. In the Philippines, Site 232 ranks highest, and Site 222 lowest. In Kenya, Site 342 is highest and 322 lowest. WDMPROP is an overall measure of the nutritional status of the school.

"WDMMED" represents the value of the median child measured at the site, as compared with the median of the reference population, where it would take a value of '0'. WDM is designed such that a child falling on the 25th percentile value of the reference population takes a value of -1.0 and a WDM measure of a child at the 75th percentile is designated at +1.0.

"WDMBAR" represents the Site Average (not an individual, such as WDMMED). However, the reference point in the standard population from which it is measured is the median of the standard reference population.

WDMPROP, WDMMED, and WDMBAR correlate highly with each other and with the traditional "percent of standard weight-for-height," as would be expected. WDMPROP and WDMBAR are perhaps the most convenient for program officials to use, since they are based on percentages and averages representing the school as a unit. WDMMED is used principally for statistical calculations requiring the use of a median, and because it represents an individual typical child. Of the three measures, WDMPROP is the one most likely to be understood and preferred by a wide audience, and is the most useful single measure of nutritional status at a site.

Readers interested in a full discussion of these measures should refer to Chapter Four. That chapter presents both mathematical and graphic descriptions of the measures.

- Advantages and Limitations of Weight Dispersion Measures

The WDM measures, as initially derived for this study, were tied to the Stuart-Meredith Scale, which (in turn) is based on reference populations sampled in the United States. WDM measures could be keyed to

other reference population data: data currently being prepared by the National Center for Health Statistics for reference populations gathered in particular developing countries; or even data for reference populations gathered for particular ethnic groups within (or among) countries, if such should become available.

It should be understood that the 25th percentile used in WDM represents a convenient statistical reference point, which permits quantitative description of children above, and below, "a line." That line is not necessarily clinically significant. The nutritional significance of the line, as it relates to a reference population in the United States, may well vary from country to country, or from population group to population group, depending on genetic, climatic, and other circumstances. Thus, it is possible that a clinical evaluation would show that nutritional status of the control sites in the Philippines, shown in Exhibit 2.3, is uniformly superior to that shown for control sites in Kenya. It is also possible that if WDMPROP were keyed to good reference population data in each of these countries, the 25th percentile in each country would permit a higher degree of comparability for purposes of cross-national comparisons.

Most evaluations, however, do not require cross-national comparisons among sites. WDMPROP and other WDM measures permit the establishment of a baseline and examination of variations from a baseline at a given site. They also permit thoughtful comparisons among sites within a country. Whatever the state of other information concerning those sites may be, if substantial variations in genetic endowments, or other pertinent circumstances are thought to exist, such factors can be noted and taken into consideration. If such other information is not available, the WDM measures constitute a consistent quantitative input to management judgments which must be made in any event.

As WDM is applied in a country over a period of time, evidence will accumulate on the clinical significance of the statistical demarcation made at the 25th percentile and on genetic and other factors which may cause

variations among sites. Such evidence should lead to a growing understanding of the context in which the measure operates and possibly to some adjustment in the percentile used as a point of reference.[1/]

Implicit in the formulation of WDM is the presumption that a principal concern of child feeding operations is with the children who are most nutritionally deprived in a given feeding environment, who have been identified for statistical convenience as those in the fourth--or lowest--category below the 25th percentile of the reference population, having WDM values less than -1.00. This presumption may not be well founded in particular cases, or even as a general proposition. It could turn out, for example, that children in the third, or next-to-lowest category (between the 25th and 50th percentile with values between -1.00 and 0) are those who benefit most from nutritional supplements because families in the fourth category make significant downward adjustments in their children's home rations in response to the provision of nutritional supplements, and third category parents do not. As knowledge accumulates, it thus may be desirable to develop other formulations of weight dispersion measures. For the present, the indicators presented in this report are deemed sufficient.

- Applications of Weight Dispersion Measures

We envision the following principal applications of weight dispersion measures:

-- Selecting sites for new school or MCH feeding operations;

-- Establishing baseline conditions for new school or MCH feeding operations;

-- Determining nutritional status for purposes of ongoing evaluations;

-- Determining end-of-project status; and

-- Terminating operations at a site, when faced with scarcity of sources or lack of continuing demonstrable need.

WDM indicators provide a basis for making decisions for locating new child feeding operations where selections are required among competing candidates. The ways in which these indicators can be used for siting

---

1/ These same comments apply to standard weight-for-height measures as well as WDM.

naturally depend on the decision-making context. Suppose the decision-maker's problem was to select three sites from the five shown for the Philippines. If he had no information other than that shown on Exhibit 2.3, he might decide to target the three with the lowest WDM indicators: Site 212 (WDMPROP: 31), Site 222 (WDMPROP: 33), and Site 242 (WDMPROP: 41). However, he might also choose to "hedge his causal assumptions" by substituting a site with higher nutritional status for one of the three low status sites. This might well be advised if the sites are to be subjected to repeated intensive evaluation for purposes of exploring causal relationships. He might then substitute 232 (WDMPROP: 62) for Site 242 (WDMPROP: 41), providing a site selection with considerable ostensible nutritional variability.

When new child feeding operations are introduced, the WDM indicators become the baseline which represents starting nutritional status at the site. A projected improvement in nutritional status of the site over time becomes "end-of-project status," and periodic applications of WDM indicators measure progress toward this goal. Terminations of site feeding may come about because of reduced program resources, determinations that it is desirable to concentrate available resources in fewer locations, absence of continuing need at a given site, and for other reasons. Using WDM indicators for designating end-of-project status may serve as a signal either for the end of all feeding or a particular kind of intervention.

Some program officials may object vigorously to the concept of "end-of-project status." School-feeding operations (as distinguished from foreign AID contributions to school feeding) may not be designed to end, as a matter of policy. In any case, supplementary feeding projects seldom directly ameliorate conditions in the society or the community which create the need for food supplements. In some cases, they, in fact, may engender dependence upon such supplements.

Another way of assessing beginning and ending conditions is to examine the status of a cohort of children as they enter, participate in, and leave the project. If such a cohort is matched, in addition, against a

cohort of "control" children, a careful analysis holds a promise of sorting out feeding impact with somewhat greater certainty than a yearly comparison of the status of the cohort at the site, taken by itself or compared with other program sites. In a good many circumstances, it may be more appropriate to answer the question "What nutritional improvement do we seek to accomplish in cohorts of children during the course of their exposure to feeding at this site?" than to ask, "How do we know when it is time to end supplemental feeding operations at this site?". In any event, the Item C in Exhibit 2.1, "End-of-Project Status" has been changed to the "Nutritional Status" in Exhibit 2.2, permitting the use of either (or both) of these approaches. The program officials charged with conducting evaluations can make their own judgments as to how they wish to treat this item.

Use of WDM indicators will require refinement over time. In the early stages of application, estimates of long-term outcomes will necessarily represent educated guesses. A body of information and experience will emerge gradually. In order to gain maximum understanding of what the evaluators are learning about program sites and cohorts, it is desirable that WDM measures should be gathered at non-program sites. Program and control data can be matched on the basis of WDM status rather than on the basis of socioeconomic status indicators, as is normal practice in research design. As better data accumulates on rations actually distributed, food actually eaten, number and regularity of meals actually served, and length of exposure to the program, a clearer picture of nutritional impact should emerge.

Program officials should be particularly alert to opportunities to provide meaningful ways to analyze the effects of feeding. While it generally may be desirable to use random methods to select non-program control sites, there are three clear exceptions. First, it would be desirable to assure that school sites are chosen which are close to MCH sites, so that graduates of the MCH program could be followed through the system to discern possible long-term impacts of their earlier supplemental feeding. Second, where a non-program school is under consideration for installation of a new feeding operation, evaluators have an interest in gathering WDM indicators

for a number of years in advance. Similarly, where a school is being considered for termination, it would be desirable to gather WDM indicators for a number of years in advance and to continue measurements after the program is terminated. The advantages (from the point of view of accumulating information which is useful in evaluations) must, of course, be balanced by program officials against the need for action.

4. "Outputs" are treated as rations delivered to targeted recipients rather than as food ingested by targeted consumers.

Program outputs are treated as the products of a distribution/food conversion system which add utility to physical inputs by means of transportation, storage, breakbulk operations, and food preparation. That concept treats an output as roughly equivalent to what a customer receives at the table or the take-out counter of a chain restaurant in the United States.

But, the analogy is not exact. From a theoretical evaluation viewpoint, it would be preferable to measure, through time, the amount of program food which is actually eaten by each intended beneficiary of a child feeding operation. If such a measure were available, it would serve as a valuable indicator which could be related to improvements in nutritional status, as measured by WDM indicators. Such ideal ingestion measures are seldom available outside of a clinical environment. In point of fact, it is very difficult to obtain fully adequate and reliable information on the volumes and time profiles of rations delivered to intended recipients, let alone the volumes and time profiles of rations actually eaten by intended beneficiaries. The lack of suitable indicators dictates an evaluation strategy designed to bridge an analytical gap.

At the output level (Level 3 in Exhibit 2.2), the evaluation approach in this report focusses on whether program outputs are commensurate with program inputs, and whether they are adequate to produce intended nutritional results. The term "commensurate and adequate outputs" has thus been substituted for the term "planned outputs" in the standard AID formulation. The practical implication of this change is to take the site official off the

hook for meeting ration distribution targets established for a site in advance by other levels of management, and to put him on the hook for distributing rations which are commensurate with what he actually receives from higher levels of the distribution system, and with additional resources available to him at the site. It puts all program officials, and the evaluation system they utilize, on the hook to determine whether rations distributed to intended recipients are adequate to produce change in the nutritional status[1/] they are seeking and to explain what is happening in the analytical gap. In terms of evaluation concepts, this treatment shifts the determination of "adequacy" upward from Level 1 in Exhibit 2.1 to Level 3 in Exhibit 2.2. The question of whether planned levels of distribution flow are being achieved is shifted downward from Level 3 in Exhibit 2.2 to Level 1 in Exhibit 2.1.

One practical upshot of this treatment is to encourage site level officials throughout the system to report what is actually happening, as distinguished from what is supposed to happen at the site level. The information gathered during the pilot study indicated that ration sizes in a given program (MCH or school feeding) in a given country could vary among five sites by a factor of as much as 7 to 1. Presumably, a good deal of such variation is real. However, some of it clearly reflects estimating methodology and site reporting based on plans as distinguished from performance. The better the data which is evoked on actual ration distribution, the better will be comparisons with measures of nutritional status and provision of inputs. Since such comparisons provide the keys to improved management of the system, program officials should lay considerable emphasis on getting at the reality of output delivery during evaluations.

1/ Included in the ranks of program officials are many people who possess a deep personal dedication to the cause of achieving tangible improvements in the health of children. Such people are already "on the hook" in terms of their personal values. The evaluation approach presented here should provide administrative reinforcement for those critically important human motivations.

5. Provision of inputs is treated as a function of an ongoing distribution management system.

The definition of input objectives shown in Item C-1 of Exhibit 2.2 emphasizes two types of management performance: (1) meeting planned targets in providing food, funds, and staff to the distribution system, and (2) making suitable adjustments in these planned inputs in the event that changes are required by circumstances inside, or outside, the system.

Inputs are thus regarded as resources requiring management judgment at the input level, not merely as postulated givens which may turn out to be "adequate" or "inadequate." It is pertinent to inquire as to whether the system adjusts its intake to constraints and opportunities, as well as whether it is performing at planned levels. It should be borne in mind that the "outputs" of the system have been defined as "rations to be delivered to intended recipients." It is at this level (Level 3 in Exhibit 2.2) that the root question of adequacy of resources is addressed. Responses to that question, as well as changes in physical conditions within the distribution system, in prices of program inputs, and in program funding availabilities, can require adjustments in magnitudes and schedules of planned inputs. The skill with which shifts are made are a proper subject of inquiry in an evaluation.

Input management clearly has a cost side as well as a physical, "goods and services" side. Indicators should be selected which allow comparisons between these two ways of measuring resource flows for purposes of analyzing the efficiency of the input function in distribution management.

The title of Item D-1 has been changed from "Budget and Schedule" in Exhibit 2.1 to "Input Indicators" in Exhibit 2.2 in order to accommodate the treatment of inputs as a two-sided phenomenon. Nevertheless, for a given evaluation, program officials may be interested only in making summary comparisons of input costs with outputs. In such a case, the evaluator's task is to ascertain that the input costs utilized are complete and conceptually consistent. Examination of physical flows provides a cross-check and a means of estimating the costs for any missing cost categories.

However, such comparisons are not mandatory if the cost data appear to be complete and there is no desire to evaluate the efficiency of input management.

Again, program officials have a choice. They can look at the input function as a dynamic of the system or they can treat inputs as a flat summary cost item. This choice should be explicitly treated in the evaluation design.

6. The "higher level goal" of child feeding programs has been defined as improving the mental capabilities of intended consumers of food supplements within the ambit of the school system.

Item C-7 in Exhibit 2.2 has been labelled "Higher Level Goal," as distinguished from "Sector, or Programming Goal" used in Exhibit 2.1. It has been noted previously that the higher level objective is of less concern to a program official than are project purpose and outputs. The administrator's responsibility for goal achievement is "attenuated," because he has less direct responsibility for its achievement.[1/]

Four options were considered in connection with the definition of a goal for child feeding operations. First, it is possible to treat the site as the unit at which project purpose is pursued and to define a nationwide programming goal as the sum of the purposes of site operations. Thus the national programming goal is to improve the nutritional status of targeted children throughout the country, while the site goal is to improve nutritional status at the site. Conceptually, this would make "purpose" and "goal" identical, except in their respective geographical scopes. The measure of both purpose and goal would be changes in the nutritional status of children, with aggregate statistics being provided on the national level.

A second alternative is to "truncate" the matrix: to treat purpose and goal as identical, and leave it at that. The argument for this

---

1/ Project Evaluation Guidelines, op. cit., p. 6.

treatment is that improving the nutritional status of children is an end in, and of, itself, and that the first alternative is simply a facile way of avoiding this underlying reality.

A third alternative is to define "goal" in very broad terms, such as "improving human performance." It is certainly true that programs to improve the nutritional status of children are undertaken because it is believed that nutritional status has an impact on children's potential for leading useful lives and contributing to society. However, such a definition is rather wide in its potential sweep. It founders when a search is made for "objectively verifiable indicators" which could be applied in the real world of evaluation.

A fourth alternative is to define a more limited higher level goal which is not identical to project purpose but is closely related to the environment in which program officials operate. That is the approach taken in Exhibit 2.2. In Item C-7, the "Higher Level Goal" is defined to be that of achieving measurable achievement in the mental capabilities of targeted consumers.

This definition places "goal" squarely within the functions of the school system. If we think of the principal of a program school as the site level official responsible for school feeding, it is a goal which is readily understandable, and relevant to the site official's regular duties. It is also a goal subject to objective measurement by means of tests, and (at least theoretically) grades. In the case of MCH feeding, tests of mental acuity are difficult to administer to small children during their exposure to the program. However, such children might be tested when they enter school and changes could be measured thereafter. It is for this reason that we suggested, earlier in this chapter, that some control schools might be selected on the basis of their proximity to MCH centers.

To repeat, the advantage of defining the higher level goal as achieving measurable improvements in mental capabilities is that this goal well

fits the role of the school system. However, three final points should be made in connection with this treatment. First, this approach elevates to the status of "Higher Level Goal" a program impact (improved mental capabilities) which might also be considered a "project purpose," or a "secondary benefit" by some evaluators. Second, the procedures recommended in this report do not attach high priority to the measurement and analysis of indicators of mental capabilities. Such measures can be expensive to administer and may not be conclusive. Finally, mental capabilities are not the only effects of improved nutritional status which could be measured in the schools. For example, tests of physical prowess might also reveal a "higher level" effect of supplemental feeding.

Once again, the program official has a choice. In practical terms, his decision will most likely be whether he wishes to incorporate measurements of goal in a given evaluation design. If his decision is negative, his selection of a "higher level goal" is of limited immediate importance.

F. EVALUATING LINKAGE PROCESSES

This section discusses the linkages between inputs, outputs, purpose, and goals in terms of "systems analysis." The emphasis is not on the technical details of systems analysis, but rather on concepts and practical implications for evaluation. Exhibit 2.2 makes the following conversions in the characterization of linkages shown in Exhibit 2.1:

- Level 2, if inputs, then outputs in Exhibit 2.1, becomes distribution processes in Exhibit 2.2;
- Level 4, if outputs, then purpose in Exhibit 2.1, becomes metabolic processes in Exhibit 2.2;
- Level 6, if purpose then goal in Exhibit 2.1, becomes mental development processes in Exhibit 2.2.

These conversions are certainly not mandatory: in a theoretical sense they are rather oversimplified. Nevertheless, they serve to point out important consideration for program officials. One of them is that child feeding programs link together "subsystems" or "linkage processes." It is worthwhile to examine these processes one at a time.

In Point 4 of the previous section, we observed that an ideal measure of output would be an indicator accurately reflecting the amount of program food ingested by intended beneficiaries. At that point, the bemused program official may well have exclaimed, "That's not an output, it's an input!" In fact, each of the four main elements of the logical framework (inputs, outputs, purpose, and goal) can be regarded as a synapse, or joining point, at which the "output" of one system, or linkage process, becomes the "input" to the next. The logical framework, then, becomes one section of a long chain. Behind "inputs" lies an agricultural system. Ahead of "goal" lies a socio-economic system in which mental capabilities are put to use. This way of looking at the framework emphasizes the importance of understanding and evaluating what is happening in the linkages. In terms of Exhibit 2.2, it treats child feeding as a program in which food distribution processes (Level 2 in Exhibit 2.2) are "connected to" the metabolic processes of children (Level 4), and the metabolic processes of children (Level 4) are "connected to" their mental development processes (Level 6). Thinking of these processes as the links in the chain permits the program official to size up evaluation functions in fairly concrete terms. It enables him to deal, in familiar terms, with such questions as, "What kinds of functions are we really evaluating here?" and "What kinds of people are qualified to evaluate those functions?" It is in such concrete terms that each of the three processes shown in Exhibit 2.2 are described below.

1. The Distribution Process

What are called "distribution processes" in Exhibit 2.2 are in fact, distribution and food preparation processes. The types of people who know about distribution systems are not necessarily the same types of people who know about food preparation. It is believed that in most evaluations it will be the distribution system problems that occupy the center of the stage. Solutions to distribution problems are likely to afford the greatest opportunity for making substantial improvements in operations, reducing costs, and obtaining vital information concerning what rations are really being distributed to recipients at the program sites. On both the

cost and physical distribution side, the whole system can be viewed as a system of pipes and reservoirs, subject to a certain amount of leakage. The job of system management is to manipulate "input" and "output" valves (procurement and food distribution at the site), watch the reservoirs (stores and inventories), plug the leaks (distribution losses), and monitor the whole system to make it work (distribution system management). When the system is "loaded" there is trouble if an input valve is opened, and an output valve is not. When the system is empty and someone opens an output valve, nothing will come out until someone opens an input valve. Communication and coordination are important, whatever the system's status.

In the case of food distribution, there is no single skill, such as plumbing or hydraulic engineering, which spans the whole system. The functions involved are transportation, logistics, storage, record keeping, communications, and the management of flows of physical goods through the system in ways that minimize costs as they meet planned, or adjusted, schedules. It is in <u>flows</u> through the system that waste, losses and inefficiencies can occur. It is in managing and establishing controls on these flows that substantial improvements in data and performance can be made. There will be some factors outside the control of program officials (natural disasters, strikes, absence of adequate communications, division of responsibilities among organizational units, and the like); but the relationships within the system are well understood and fundamentally manageable. In evaluating a distribution system, a tough-minded practitioner, with experience in logistics and in analysis of accounting records, can be worth a hundred good nutritionists. However, if the problem being evaluated is one of menu or food preparation, one good practical nutritionist can be worth a hundred logistical types.

2. <u>Metabolic Processes</u>

Metabolic processes build and destroy protoplasm in a living child. They produce the chemical changes in living cells by which energy is provided for the vital processes and activities, and new material is assimilated for body growth. Improved nutritional status--the purpose of child feeding programs--is an outcome of metabolic processes. These processes influence, and

are affected by, many phenomena other than food supplements alone. These other phenomena include non-program food, water, climate, human activity levels, and the onset of disease. Food goes in. Energy and waste go out. The metabolic processes determine whether the body builds its inventory of protoplasm, or depletes it. Metabolic processes thus constitute a kind of biochemical management system, the history of whose operations accounts for the body's stock of protoplasm at any given time. The essence of the problem is to determine the impact of specific inflows (supplemental feeding) on an inventory level (protoplasm).

In a clinical setting, the analysis of impacts of feeding on metabolic processes lies in an interdisciplinary borderland, occupied jointly by researchers in the fields of nutrition, biochemistry, and medicine. For the purposes of evaluating child feeding operations in a non-clinical setting, experts in survey research, statistics, and quantitative methods may justly lay claim to that territory. The factors influencing the process examined are many in number. Unlike food distribution system processes, the relationships involved as food is converted to protoplasm and becomes part of the body's inventory, are imperfectly understood. Reliable indicators are few in number and the program official is not in control of "the system." The problem is to detect subtle associations among phenomena which have been measured with very crude instruments and, hopefully, to establish the existence of quantifiable causal relationships. This is the province of quantitative analysis. If he can successfully establish and quantify relationships between food supplements and nutritional status for purposes of field evaluations, one statistically-oriented research analyst can be worth a hundred practitioners.

3. Mental Development Processes

The number of factors presumed to affect mental development is rather large. Such factors include genetic inheritance, family values, motivational and sociological factors, and education and training. It is at least theoretically possible that supplemental feeding could affect mental development through "routes" other than (or in addition to)

improved nutritional status. Quite apart from their nutritional impacts, supplemental rations may act as a "placebo" which motivates school attendance and therefore possibly accelerates mental development through additional classroom exposure. It is also possible that feeding has short-term "energy expenditure" effects on learning. Greater attentiveness could be such an effect. For purposes of basic research, these presumed impacts fall within the province of many disciplines: neuro-physiology, psychology, sociology, and the like. For purposes of field evaluations of child feeding projects, statistical analysis may prove to be the only practial means of identifying important relationships.

******

In summary, the program official should understand the nature of the systems being evaluated. He should select evaluators with skills suitable to such systems and problems typical of the processes being analyzed.

G. STATISTICAL ANALYSIS

This section deals with the use of statistical analysis in evaluations. However, no attempt is made to give the program official "a short course in statistics." It should be well understood that mastering the types of statistical techniques required for analyzing subtle and complex relationships in child feeding operations requires "the long course." Accordingly, this section deals briefly with three areas in which program officials may be required to exercise judgment: making decisions on the scope of data to be gathered for multivariate analysis, interpreting assessments by quantitative analysts, and implementing a long-term strategy for gathering and analyzing data.

1. Scope of Data Gathered for Multivariate Analysis

Multivariate analysis provides a means of examining whether association exists among factors which may affect, or be affected by,

child feeding operations: the size of the child's ration, the length of his exposure to the program, the socioeconomic status of the child's family, school attendance, the alertness of the child and (most importantly), the nutritional status of the child. It is important to bear in mind that the more indicators and factors are analyzed, the greater the potential for explaining variability in the associations that may exist among them. This consideration argues for casting a wide net for data, for gathering information on as many variables as possible. In addition, some persons may urge the "wide net" approach because of their special interests in particular kinds of data or out of intellectual curiosity.

But there are strong countervailing considerations. First, in most developing countries, the number of variables on which reliable data is available, or can be measured inexpensively, is quite limited. Second, once the data are gathered, the costs of exercising a high degree of quality control can be formidable. The consequences of <u>not</u> exercising a high degree of quality control can be quite serious from the viewpoint of the quality of results. Third, it may be possible to gather information on many interesting variables only through interviews, observations and follow-up visits. The longer the questionnaire, the lower the accuracy, the greater the cost, or the smaller the sample.

The results of the pilot survey discussed in Chapter Four indicate the prudence of concentrating on a few key variables. It is suggested that evaluation designs intended to make use of multivariate analysis take account of this hard-won wisdom.

2. <u>Interpreting Assessments Made by Quantitative Analysis</u>

In Section F, it was stated that evaluating the effects of child feeding on the linkages represented by Level 4 of Exhibit 2.2 (Metabolic Processes) and by Level 6 (Mental Development Processes) may fall within the province of quantitative analysis. It was suggested that one quantatively-oriented research analyst could "be worth a hundred practitioners" if he succeeds in capturing the subtle quantitative effects of child feeding operations within the web of his statistical techniques. <u>If</u> he succeeds.

In fact, the quantitative analyst may be dealing with data that is so imprecise or unreliable that his methods reveal little concerning subtle effect--or he may generate impressive results which are, in fact, spurious artifacts of bad data. Under such circumstances, the experienced eye and intuitive judgment of the practitioner (the program official, nutritionist, or clinician who has had exposure to field conditions over the years) may reveal more, and constitute a better basis for, judgment than does extensive processing of poor data.

But the talents of a quantitative analyst may not be limited to the application of the formal disciplines of his trade. He may have a practiced eye and a fine intuition concerning the potential implications of numbers. His experience and intuition enable him to formulate hypotheses concerning the presence of subtle effects which go beyond associations which he can establish through the rigorous application of his professional disciplines to the data. Indeed, he may have <u>two</u> practiced eyes--one of them trained on the numbers, the other recording his impressions of what is happening in the real world from which the numbers may have been imperfectly drawn.

However important such insights may be, the program official should clearly understand the difference between a situation in which a quantitative analyst presents findings deemed to be meaningful in terms of the rigorous application of <u>statistical disciplines</u> and one in which he reports judgments based on his perceptual skills, his intuition, and his experience. The former is statistical evaluation; the latter is "eyeball" (or judgmental) evaluation, however it may be clothed in the language of the profession.

There is no reason why a quantitative analyst should not report his impressions, intuitions and judgments as well as his professional conclusions concerning the meaning of the data. He may, after all, be the best "eyeballer" available. But the quantitative analyst should distinguish clearly between the results of statistical rigor and the results of "eyeballing." And so should the program official who utilizes the

analyst's work products. In the one case, the official is dealing with "objectively verifiable indicators;" in the other he is confronted with speculations concerning "assumptions about linkages." In the one case he is dealing with a methodology which is replicable and results which can be confirmed by anyone with suitable professional training. In the other he is confronting interpretations by a particular individual.

The program official, and indeed the quantitative analyst, may not always be able to make a clear distinction between these two types of evaluation work products. Under such circumstances, a bit of cross examination may be in order.

There are two kinds of questions which the program official can put to the quantitative analyst under such circumstances.

-- "On the basis of what hypotheses did you analyze the data? Did you apply a null hypothesis[1/] in your analysis? If not, why not? If so, with what results?"

-- "Please identify for me which of your findings rest on interpretative personal judgments. Please describe to me, in layman's terms, what the nature of those judgments was."

The answers to these questions will lead naturally to others, and the program official should be able to emerge from the discussion with a clear impression of the basis on which the findings have been developed.

3. An Analytical Strategy

Chapter Three of this report suggests a temporal approach to sorting out the effects of child feeding programs, which emphasizes the gathering of data on a few key variables over time. Section D of the present Chapter describes an evaluation strategy designed, among other things, to accumulate data on children's nutritional status, to obtain more accurate information on the rations actually distributed at child

---

1/ For a brief discussion of the null hypothesis, in this respect, see Chapter Four, Section B.

feeding sites and to provide regular feedback from site officials concerning factors affecting the nutritional status of intended beneficiaries.[1/] Such a strategy can be supplemented with the use of multivariate analysis to examine a few environmental variables and diagnose unusual problems.

The approach recommended in this report is designed to produce both short-term payoffs and long-term results. The short-term payoffs include helping officials to make other decisions on establishing and terminating site operations, improving the efficiency of distribution systems, and providing quantitative and qualitative information on site conditions. But, the analytical strategy of determining the effects of feeding on the nutritional status of children is essentially long-term. It consists of building up good information on nutritional status at program and control sites, and in making substantial improvements in information concerning what rations are actually delivered to intended recipients, and when these rations are delivered to them. As such, a data base is built up over time. It can be used with increasing effectiveness to address issues of causation.

H. IN RETROSPECT

At the beginning of this Chapter we quoted the salesman's injunction, "Ya gotta know the territory." The approach to evaluation recommended in this report takes both the feeding program and its evaluation very seriously. Child feeding operations are important undertakings with important purposes. The procedures recommended in this report are designed to enhance the management of these programs, to improve understanding of what they are in fact accomplishing, and to help assure that the goods are in fact delivered.

The implementation of the recommendations in this report requires commitments: commitments in the form of resources devoted to management controls; gathering of information, and analyzing that information; and commitments in the form of continuing support from program officials, faced with competing demands for their attention. It is believed that such commitments are merited and will contribute significantly to the achievement of the purpose of the programs.

1/ See Point 5 in Section D of this Chapter.