# Monitoring and Evaluation

## A Guidebook for Nutrition Project Managers in Developing Countries

F. James Levinson
Beatrice Lorge Rogers
Kristin M. Hicks
Thomas Schaetzel
Lisa Troy
Collette Young

September 1999

Human Development Network
The World Bank

Prepared for The World Bank, Human Development Network by
International Food and Nutrition Center
Tufts University School of Nutrition Science and Policy

# Contents



**Acronyms** . . . . . **vii**

**Glossary** . . . . . **ix**

**Introduction: How to Use this Guidebook** . . . . . **xiii**

**1** **An Overview of Monitoring and Evaluation** . . . . . **1**

**2** **Clarifying Project Goals, Objectives and Information Needs** . . **23**

Pre-Design Tasks . . . . . 23

Clarifying Project Goals and Objectives . . . . . 23

Mapping the Project . . . . . 26

Determining Information Needs . . . . . 31

**3** **Designing a Monitoring System** . . . . . **35**

Monitoring for Project Improvement with a Management Information System (MIS) . . . . . 36

Responding to Monitoring Results . . . . . 47

Monitoring the Budgetary Health of the Project . . . . . 51

Planning the Flow of Information . . . . . 53

**4** **Selecting an Evaluation Design** . . . . . **59**

Designs for Evaluation . . . . . 60

**5** **Preparing for Evaluations** . . . . . **73**

Planning an Evaluation . . . . . 73

Determine the Sample Size . . . . . 74

Identify a Control Group . . . . . 75

Choose a Sampling Frame . . . . . 77

Collecting Pre-Project Information through a Baseline Survey . . 79

**6 Selecting Indicators . . . . . 81**
What Is an Indicator? . . . . . 81
Characteristics of a Good Indicator . . . . . 81
Types of Indicators for Measuring Program Effectiveness . . . . . 86
How to Select Indicators . . . . . 88

**7 Deciding Data Collection Strategies . . . . . 91**
Qualitative vs. Quantitative Data Collection Methods . . . . . 91
Maximizing the Efficiency of Data Collection . . . . . 99
Controlling for Bias through Data Collection Methods . . . . . 100

**8 Developing Data Collection Instruments . . . . . 103**
Steps in Designing M&E Instruments . . . . . 103

**9 Analyzing the Data . . . . . 115**
Analyzing Quantitative Information . . . . . 115
Analyzing Qualitative Information . . . . . 129
Returning to the Conceptual Framework . . . . . 131

**10 Maximizing the Usefulness of Results . . . . . 133**
Ownership of the M&E Process . . . . . 133
Effective Presentation of Findings and Recommendations . . . . 136
The “Usability” of Results . . . . . 137

**Annexes**

1: Program Constraints Assessment . . . . . 143
2: Conceptual Framework Examples from Three Nutrition Projects . . . . . 145
3: General M&E Checklist . . . . . 149

4: Illustrative Data Collection Forms Which Facilitate "Management by Exception" . . . . . 151
5: Sampling Tips . . . . . 175
6: Instrumentation Checklist . . . . . 179
7: Illustrative Use of the Conceptual Framework Model for Backwards Mapping . . . . . 181

**References . . . . . 185**

# Acronyms



| | |
|---|---|
| BMI | Body mass index |
| CNP | Community nutrition promoter |
| CNO | Community nutrition organizer |
| CNW | Community nutrition worker |
| CS | Cluster sampling |
| FFW | Food-for-work |
| HAZ | Height-for-age Z-score |
| ICDS | Integrated child development services |
| IDD | Iodine deficiency disorder |
| LBW | Low birth weight (< 2,500 grams) |
| MIS | Management information system |
| M&E | Monitoring and evaluation |
| NGO | Non-governmental organization |
| PRA | Participatory rural appraisal |
| RRA | Rapid rural appraisal |
| SD | Standard deviation(s) |
| SRS | Simple random sampling |
| SS | Systematic sampling |
| StS | Stratified sampling |
| TINP | Tamil Nadu Integrated Nutrition Project |
| UPGK | Family Nutrition Improvement Program in Indonesia |
| VAD | Vitamin A deficiency |
| WAZ | Weight-for-age Z-score |
| WHZ | Weight-for-height Z-score |

# Glossary



**Benefits**—the broader, sustainable changes in public health or economic status that a program seeks to achieve but which are inevitably influenced by a wide range of other factors.

**Constraints Assessment—**the systematic identification of constraints inhibiting project effectiveness. The constraints are then organized to permit the identification of technical, policy, research, and training means of addressing them.

**Control Group**—a group of individuals in an evaluative study who share the same characteristics as a participant group.

**Cost-Delivery Analysis**—study of the cost incurred to deliver a specified set and quantity of goods and services (outputs) to a targeted population.

**Cost-Effectiveness Analysis**—study of the cost incurred to achieve a specific change in nutritional status (impact) in a targeted population.

**Focus Groups**—small group discussions led by a trained moderator who introduces a topic and facilitates participation by all group members.

**Goals**—the broad aims of the project, the significant, longer-term changes that planners expect to occur as a result.

**Indicator**—an objectively verifiable measurement which reflects the activity, assumption, or effect being measured.

**Key Informant Interviews**—a face-to-face meeting between a trained interviewer and a person selected to represent a certain group whose knowledge, attitudes or practices are being monitored or evaluated, or a person likely to offer informed views.

**Proxy Indicator**—a measurement used as a substitute when true indicators are too difficult to measure directly.

**Impacts**—changes in the condition of the target population which generally reflect the primary objectives of the project.

**Inputs**—the materials, goods and actions necessary to carry out the primary project activities.

**Input assumptions**—the expectations regarding the effectiveness and quality of the project inputs

**Evaluation**—a process of data collection designed to assess the effectiveness of the project in attaining its originally stated objectives, and the extent to which observed changes are attributable to the project.

**Experimental Design**—a rigorous evaluation design which includes a control group, randomization, and pre-post project data.

**Management Information System**—a tool, often computerized, which is used to compile and analyze monitoring data.

**Monitoring**—the ongoing collection and review of information on project implementation, coverage and utilization of inputs.

**Objectives**—operationalized goals which specify the results and the level of change expected.

**Outcomes**—the intermediate effects, often behavioral, resulting directly from project outputs that may be necessary to achieve a desired impact.

**Outputs**—the provision of project goods and services to the target population. The primary project activities.

**Output Assumptions—**expectations regarding the ways goods and services (outputs) will be used by the target population.

**Sample**—subset of a population which is used to represent the entire group.

**Sensitivity Analysis**—a means of exploring how plausible changes in assumptions about uncertain variables affect conclusions.

**Special Studies—**studies to investigate issues raised before or during project implementation which can not be addressed through ongoing project monitoring.

**Stakeholders**—individuals or organizations associated with or affected by a project.

**Quasi-experimental Design—**evaluation designs that eliminate competing explanations of project effects without the benefit of a true control group.

# Introduction: How to Use this Guidebook



In international nutrition, as in other development fields, there has been a growing recognition of the need for monitoring and evaluation (M&E) systems. Monitoring is closely linked to project management and designed to assess and improve project performance. Evaluation, additionally, permits decision-makers to assess whether project objectives are being met. The absence of M&E in large numbers of nutrition projects, despite continued evidence of their value, suggests that, beyond resource constraints, some project staff may not yet have the necessary skills or confidence to develop and operate such systems.

This guidebook is designed specifically to assist World Bank and other nutrition project task managers with responsibilities they are likely to have for project monitoring and evaluation. Even if these responsibilities do not include the actual development of M&E systems, effective task management is likely to require an understanding of how such systems work, the key factors that render some M&E systems better than others, and the critical questions to ask of M&E operations.

While much has been written on the monitoring and evaluation of social services, these materials are often general in nature or geared to services operating in industrialized countries, and not always useful for specific developing country applications. We will present recognized M&E processes in ways that they can be easily understood and put into practice. In addition, we use examples to demonstrate how these processes have been used in programs addressing malnutrition in Asia, Africa and Latin America.

We begin with an overview that defines monitoring and evaluation and presents the basic framework that will be used throughout the book. The core of the guidebook then leads the reader through the steps involved in developing and implementing monitoring and evaluation systems. These sections are supplemented by a set of annexes which provide application of M&E tools to specific examples and additional information to be used for M&E.

The guidebook has been prepared for use either as a self-teaching manual or as the basis for group training programs. Where they are needed, references to other information sources are included. In the case of formal evaluations, particular steps are likely to require the assistance of a professional statistician. Nonetheless, Sections 4 and 5 present, in summarized fashion, the necessary steps involved and the common pitfalls encountered in setting up an evaluation—this based on the belief that project planners and managers need to be more involved in these issues than has normally been the case.

In addition, though we provide a broad array of nutrition project examples, we do not provide detailed M&E procedures for each category of nutrition intervention, particularly those addressing individual micronutrient deficiencies. Fortunately a growing number of excellent manuals on the monitoring of particular micronutrient interventions are being developed by UNICEF, WHO, the Micronutrient Initiative, OMNI, PAMM, and the micronutrient consultative groups* to help address this need. By contrast, there are numerous examples from the first Tamil Nadu Integrated Nutrition Project (TINP I) which, to date, has provided the most thorough, and probably the most effective example of a monitoring and evaluation system associated with a large scale nutrition project.

We have tried to keep the guidebook brief without sacrificing clarity and examples, and to assume an audience of professional managers with some background both in nutrition projects and in project planning and design. Readers familiar with the fundamentals of monitoring and evaluation may find the guidebook useful primarily as a check list for nutrition project M&E systems. Others who are less familiar with such processes may find it more useful as a step-by-step guide.

*The micronutrient consultative groups are the International Vitamin A Consultative Group (IVACG), the International Counsel for the Control of Iodine Deficiency Disorders (ICCIDD) and the International Nutritional Anemia Consultative Group (INACG).

Finally, we seek to counter the idea that M&E need to be a harsh, user-unfriendly regimen imposed by outsiders on overworked project implementers. In fact, M&E can and should be an adaptable and participatory process. When monitoring and evaluation are used properly, programmatic efforts to improve nutritional well-being or, more broadly, to improve the human condition, can be strengthened—an aim central to most development practitioners.

# 1 An Overview of Monitoring and Evaluation



Many professionals who work in social service programs cringe when they hear the words “monitoring” and “evaluation”. When asked to write their impressions of these terms, participants in a recent M&E workshop responded with descriptions such as “cumbersome”, “a waste of valuable time”, and “something imposed from above”. After learning more about the benefits and techniques of M&E, however, these practitioners were converted into enthusiastic supporters. Why the change of heart?

The changes in attitudes came from:

- an understanding that the primary purpose of monitoring and evaluation is *project improvement*;
- a recognition that monitoring and evaluation can be tailored to fit the specific needs of a project and usually its budget;
- a better understanding and mastery of actual M&E processes.

But the prior skepticism of these professionals was not unwarranted. Traditionally, M&E have been tacked on to a project, and in a manner often quite removed from planning and implementation processes. Monitoring systems, where they have been in place, often have been used almost solely for the production of national level reports. Evaluations frequently have been limited to external teams or individuals who arrive at the completion of a project to look at existing data, speak with individuals involved, collect impressions and write a report—all often within a few weeks. The motivation for even these limited efforts usually has been the requirements of a government or donor.

This manual reflects a very different orientation.

**What monitoring and evaluation can offer**

Monitoring and evaluation can . . .

- assess the quantity, quality and timeliness of project inputs **(M)**
- identify operational constraints to project effectiveness thus helping planners and managers improve implementation **(M)**
- determine if a process or service, e.g., food fortification, is meeting national or some other accepted / set standard **(M)**
- determine whether a project is serving intended beneficiaries **(M)**
- provide information to improve targeting **(M)**
- help to identify effects that are attributable to a project **(E)**
- provide information which will permit cost-effectiveness comparisons with other projects seeking to accomplish the same objectives **(E)**
- meet donor accountability requirements **(M, E)**
- serve as a vehicle to increase community participation **(M, E)**
- inform decision-making on the future of a project **(M, E)**

### *What Are Monitoring and Evaluation?*

The body of social science known as ***evaluative research*** is the systematic collection of information on the design, implementation and effect of projects on targeted populations (Rossi and Freeman, 1993). Ideally, the process is divided into an ongoing ***monitoring system*** and ***periodic evaluations*** with some ***special studies*** designed to answer specific

questions about the project. Monitoring and evaluation are distinct, though related efforts, with different overall objectives, and, therefore, require differently designed systems.

***Monitoring***[1] is concerned primarily with the ongoing collection and review of information on project *implementation,* coverage and use. By collecting information on a regular basis throughout the life of a project a monitoring system can be used to assess the quality of project inputs and services, the timeliness of service delivery, the degree to which the targeted individuals and communities are reached, the acceptability and actual use of services, the costs involved in implementing the program, and the extent to which actual implementation coincides with the project's implementation plan. An effective monitoring system also provides an important input for project staff supervision.

Monitoring data are often entered into a management information system (MIS) which, in turn, provides information in an easy-to-use format to keep track of project activities, budgets, and personnel. Information generated from a monitoring system provides valuable clues as to *where* problems are occurring, *why* operations are succeeding or faltering, and *which specific aspects* of a project need to be adjusted to improve targeting, coverage and implementation. Moreover, since monitoring information is collected and reviewed at regular intervals, areas of concern can be addressed as they arise and corrective measures can be instituted, thus improving the chances for project success.

One important characteristic of nutrition project monitoring is that , in a well-functioning project most data needed for an MIS is already being collected for programmatic purposes so that establishing a monitoring system should not impose an additional burden. In an integrated commu-

1. Monitoring is sometimes referred to as process evaluation or implementation evaluation.

nity-based project, for example, this would include ongoing growth monitoring data, as well as information on activity attendance, supplement distribution, and clinic referrals.

While monitoring information is used primarily for management decisions, it is also important for providing contextual information for evaluations.

***Evaluation*** seeks to measure project effects, i.e., whether and to what extent the project's inputs and services are improving the quality of people's lives. Evaluations provide information on the changes in the behavior and conditions of targeted communities and individuals (Rossi and Freeman) by assessing the effectiveness of the project in attaining its originally stated intermediate and overall objectives. As with project monitoring, however, evaluations may also reveal unexpected findings, both positive and negative, which can be used to alter and improve project design and implementation.

Generally, following an initial baseline survey, one or two midterm evaluations take place in the mid to late stages of a project and an endpoint[2] evaluation is conducted upon project completion or at the end of a funding cycle. Ideally, there should be an efficient ongoing monitoring system in place from the start of the project. If such monitoring indicates that implementation is proceeding reasonably well, the formal evaluation can be limited to (a) the verification of the monitoring system and (b) the provision of information on outcomes and impacts.[3] In the absence of a

2. The term "endpoint evaluation" will be used throughout this guidebook to denote studies which are conducted upon the completion of a project to measure outcomes and impacts. Other terms commonly used are summative or impact evaluation.

3. In cases where monitoring indicates that implementation is seriously substandard, evaluations of project effect make little sense. In one recent situation in southern Africa, formal evaluations were canceled despite the existence of baseline data and control groups because delivery of project inputs was so flawed that no impact could be expected.

## Field Insight: Neglecting the Reasons for Change

In Vietnam, community health and nutrition workers observed that, despite comparable socioeconomic status, some children were growing adequately while most suffered from varying degrees of under-nutrition. Using the positive deviance* approach, the behavioral characteristics of "successful" households were identified. These behaviors included the collecting of shrimp and crabs while working in the fields and supplementing a child's diet with them. An educational campaign promoting such positive deviant behaviors was implemented. After several months, an evaluation was conducted to investigate the effect of the project on nutritional status. Evaluators were pleased to discover that nutritional status had substantially improved since the baseline. However, the evaluation collected only anthropometric data and neglected to determine whether the adoption of positive deviant behaviors had, in fact, taken place. While the overall improved growth of children in the project area is cause for celebration, the conclusion that this resulted from the educational campaign may have been incorrect. In fact, deworming medication had been introduced into this area during the same period and may have contributed substantially to the decreased rates of under-nutrition. Simply examining impact relegates the reasons for change to a "black box". Unfortunately, this approach to evaluation is all too common: improvement in impact indicators, where it is observed, is assumed to be attributable to the project without examining process (what we later will define as output and outcome) indicators.

*Positive deviance refers to situations in which individuals or households are doing better than would be expected given their social and/or economic circumstances. Their time and resource allocation strategies may be worth disseminating more broadly.

reliable and comprehensive monitoring system, however, an endpoint evaluation would have to include an explicit assessment of the implementation process to determine the extent to which the target population actually was reached and services delivered. Without this information, any absence of positive impact will leave unanswered the question of whether the problem was a structural defect or faulty implementation. Similarly, even positive results cannot be attributed to the program interventions when information about the process is absent.

While project monitoring clearly serves the interests of program funders, program managers and staff, and beneficiaries, all of whom benefit from a process that improves project operations, each of these groups might resist an impact evaluation out of concern that the result might be negative. Funders might have to admit to mistakes in judgment; program managers and staff might consider their jobs threatened; beneficiaries, if they are receiving food or other goods or services may fear their loss. Overcoming such resistance to evaluation is not always possible, but experience suggests that resistance can be reduced if the stakeholders of a project are involved in planning for the evaluation and reviewing evaluation data as they are being compiled.

Since monitoring data are essential to effective management, all projects should be monitored. Most projects have some form of monitoring system in place for precisely this reason. Far fewer projects conduct regular evaluations. Of 97 feeding programs in Latin America analyzed by Musgrove (1991), only 10 included an evaluation, and, of these, only three used generally accepted evaluation procedures. Humanitarian workers' attitudes that every available penny must go to feeding hungry children, and program managers' concerns with time and financial constraints, were explanations most often given why so few evaluations are conducted (Musgrove, 1991). In other cases, project funders and implementers believe that the project is so obviously beneficial, and the potential for negative effects so small, that evaluation is simply a waste of scarce resources. They may argue that evaluation resources—not just

money, but staff time and disruption—would be better spent to expand the project.

This reasoning may be dangerous. There are many examples of projects that have proven to be ineffective, or, in some cases, have even had negative effects, despite high expectations for their success. Resources spent to evaluate a project may result in far more effective use of the remaining resources available to the project. For example, one ineffective component of a generally effective project may be scaled back, saving resources that can then be used for the more effective components. Or an evaluation may find a project is effective in addressing the needs of one target group but not another, so that some resources can be redirected in more effective ways. Or negative side effects of a generally effective program can be reduced or eliminated by suitable program modification.

Committed project staff and management often believe strongly in the value of the project they operate; they will focus on the successes and perhaps neglect to see the less successful aspects of the project. Only systematic evaluation can truly verify or modify these positive impressions. For example, a comprehensive review of supplementary feeding programs, conducted in the 1980s, showed that targeted supplementary feeding of malnourished infants and preschoolers *in the absence of complementary health services* showed no effect on the nutritional status of this target group (Anderson, 1977; Beaton, 1982). The evaluation was considered virtual heresy at the time, but it eventually led to careful exploration of the reasons for this result. As a consequence, supplementary feeding programs are now far more likely to be implemented in the context of comprehensive, integrated health and nutrition services. But at the time this evaluative review was carried out, many of us were so set in our conviction about the automatic translation of food supplements to improved nutritional status that we would have argued (and some did) that evaluating such programs was a waste of resources, virtually taking food from the mouths of hungry children.

Evaluations need not always be elaborate, lengthy, or costly. If monitoring data strongly suggest the existence of positive effects, an evaluation may simply verify that these effects are attributable to the project, by comparing current with baseline data, and by comparing beneficiaries in the target area with comparable individuals or households in areas without the project. The complexity of an evaluation depends in part on resource availability and in part on the complexity of the project itself, but often a focus on a few measurable impact indicators is sufficient, if ongoing monitoring data are sound. This underscores the need for effective and comprehensive monitoring, which serves both management needs and the needs of evaluation at key points in the life of a project.

***How M&E Fit into a Project***

While monitoring and evaluation are complementary, they are two distinct processes. Monitoring follows a management model with a focus on improving day to day operations. Evaluation uses a research model to assess the extent to which project objectives have been met or surpassed. However, monitoring and evaluation are most effective as interwoven activities. Together they can provide information that will help decision-makers choose an appropriate course of action for the future of the project or on the direction of future projects. Depending on the M&E findings, decision makers may decide to:

- *continue the project*, either as it is currently implemented or with revisions;
- *expand the project* by increasing the target population;
- *replicate the project* in a new setting; or
- *curtail the project* and reallocate the resources elsewhere.

Often, as indicated at the outset, evaluations, or more accurately, assessments, are exercises tacked on to the end of projects to examine project implementation and impact. As will be made clear in the sections which

## Field Insight: The Benefits of Constructing a Comprehensive M&E System

By designing a comprehensive M&E system during the planning phase, staff from the first Tamil Nadu Integrated Nutrition Project (TINP) were able to use M&E data in a number of useful and innovative ways. The system consisted of the following six components.

1) Ongoing monitoring of the quality, delivery, coverage, acceptance and utilization of the services provided
2) Ongoing monitoring of project costs
3) Ongoing monthly impact "snapshots" using the growth monitoring data
4) Longitudinal data collection of a sub-sample of households or individuals to track the participation and benefits accruing to potentially under-served groups
5) Formal evaluations of 1% of the targeted population (consisting of a baseline, two midterm evaluations, and a final evaluation)
6) Other special studies

This M&E system gave the project a comprehensive feedback system which allowed staff to (a) make timely management decisions rather than having to wait for evaluation results, (b) monitor on an ongoing basis changes in the nutritional health of the population, (c) calculate costs for services delivered or impacts achieved which could then be compared to other programs with similar inputs and objectives, (d) gain valuable insights on characteristics of drop-outs and non-participants, (e) draw conclusions about the effectiveness of the project in producing short-term outcomes and long-term impacts, and (f) gain deeper insights into the internal dynamics of the project.

*Source:* Adapted from World Bank. 1980. *Tamil Nadu Nutrition Project Implementation Volume.* Washington, DC: The World Bank, Population, Health and Nutrition Department.

follow, this process is rarely capable of evaluating a project's impact, and is, in fact, the type of exercise likely to create resentment among stakeholders. By contrast, monitoring and evaluation should be built into a project during the design stage. *Project planning should always include the development of parallel M&E systems.* By incorporating M&E from the beginning, project staff will be providing themselves with a thorough and ongoing feedback system that will allow them to make timely management decisions without having to wait for the results of an evaluation. At the same time, early planning means that a valid baseline survey can be conducted and control groups established, significantly increasing the likelihood that the findings of the endpoint evaluation will be credible. Initiating an evaluation after the project is under way makes it more difficult to attribute changes in behavior or condition to the project or quantify the magnitude of the change. Both monitoring systems and evaluations are most useful if they are incorporated into a project from its inception, but both are valuable even if introduced later.

Some ways in which M&E can be used throughout the life of a project are summarized in Table 1.1.

### *Who Should Be Responsible for Monitoring and Evaluation*

There are three basic options for structuring M&E responsibilities:

- contracting external[4] monitoring and evaluation personnel
- having a mix of external and internal (project) personnel
- relying on project personnel alone

4. The term "external" here means external to the project.

## *Table 1.1* The Role of M&E throughout the Life of a Project

| | Planning or Redesign Phase (Monitoring and Evaluation) | Implementation Phase (Monitoring) | Late Implementation or Post-Project Phase (Evaluation) |
|---|---|---|---|
| Focus is on: | The design of the project and how it will improve the lives of a particular population group. | Project coverage, delivery, costs, intermediate outcomes, and other management concerns. | Determining the intermediate outcomes and more substantial impacts of the project on people's lives. |
| Types of Questions to be Answered by Monitoring and Evaluation | Are the goals, objectives, and activities appropriate in light of the project's context?<br><br>Are the project inputs and activities (including training and materials) likely to achieve these objectives?<br><br>Will the project's monitoring and evaluation system produce the information needed for critical decision-making?<br><br>Are the criteria used for targeting appropriate? | Are the specified inputs and services reaching the targeted populations, and on time?<br><br>Are inputs of the desired quality?<br><br>Are inputs being well used by the population?<br><br>Do actual project activities correspond with those spelled out in the project design or implementation plan?<br><br>What are the project costs and do they correspond to the budget plan? If not, what components of the project are over and under budget?<br><br>Is there evidence of short-term, intermediate outcomes that will produce long-term impacts? | What, if any, are the outcomes and/or impacts of the project on the targeted populations?<br><br>Have the originally stated objectives and goals been met by the project?<br><br>What other effects, intended or unintended, did the project have on local communities, project staff, or government policies? |

In general, the more “external” the process, the more objective it is likely to be. At the same time, wholly external evaluations often are out of touch with project realities and with ongoing monitoring processes. Wholly internal processes, conversely, will assure full familiarity with the context and its nuances, but are often considered inadequately objective by decision makers and other observers. Additionally, wholly “in-house” evaluations may not have all of the expertise necessary for such a process.

Decisions about the internal/external balance in an evaluation will necessarily vary from project to project. In seeking the ideal balance for a given project, the following three scenarios may be helpful:

***Scenario I: An ongoing external evaluation presence***

In large, expensive nutrition projects, it may well be worth contracting with an external institution which would be actively involved in evaluation-related activities throughout the life of the project. Such an entity, working closely and interactively with internal M&E staff, could be responsible for the following:

- selection of control groups
- collection of baseline data on a representative sample from the project and control populations
- subsequent collection of midterm and end-point data (primarily on outcomes and impacts) on participants and controls
- regular quality checks on the monitoring data being collected internally by the project
- periodic disaggregation of the monitoring data to assure that particular groups (e.g., religious, caste, food-insecure, females, those residing on the outskirts of villages) are not excluded and are sharing in project benefits

- special studies identified at the design stage or during project implementation

- periodic assessments of the perceptions of service providers and beneficiaries on project effectiveness, constraints and means of addressing them (using the program constraints assessment methodology described in Annex 1); and periodic assessment of field worker job satisfaction

- analysis of evaluation data together with internal staff

Scenario I has the advantage of reducing the workload of internal staff which can then devote their energies to project implementation and monitoring. It also assures an integration of evaluation activities with ongoing monitoring. (Several recent state level external evaluations of the Indian ICDS program did not interface with ongoing monitoring efforts, thereby not only depriving themselves of crucial "process" information, but also resulting in confusing conclusions.) Assuming a fully competent and responsible external institution, the quality of the entire process is likely to be high, but the evaluation costs will also be high. (As a rule, monitoring and evaluation costs should total 3–5% of total project costs. If they are much higher, less expensive options should be considered. If they are much lower, the M&E process is probably being short-changed.)

***Scenario II : Periodic external presence***

In medium-size projects, it may not be necessary to have an external institution involved in all of the above. Instead the external entity, working at all times with internal staff, could take responsibility for designing the evaluation, assisting in the identification of control groups, participating in the baseline data collection and the midterm and end-point evaluations, and participating in the analysis. In this scenario, quality checks on monitoring data would be carried out, at least occasionally, by internal M&E staff, and special studies would be contracted out or conducted internally.

***Scenario III: External presence at beginning and end only***

In smaller projects, an external evaluator, often a single individual with M&E expertise, would be present at the beginning of the project to advise on the M&E system as a whole and specifically issues of control group (or a reasonable substitute), sample size and critical indicators, and, in some cases, provide necessary orientation for staff who will be responsible for data collection and analysis. The external evaluator would then rejoin internal M&E personnel at the conclusion of the project to review the monitoring and evaluation data collected and the analysis carried out, and would meet with project managers, service providers and groups of beneficiaries to discuss the process and the conclusions.

Whenever an external entity or individual is utilized, care should be taken to provide clear terms of reference and necessary project documentation.

Regardless of which M&E staff structuring is employed, relevant training of those responsible for M&E operations is essential to ensure quality data collection, analysis and interpretation and effective action. In nutrition projects, this is likely to include skills in nutritional assessment plus more generic ones associated with interviewing, focus group facilitation, and data processing.

***What to Monitor and Evaluate***

This guidebook breaks down monitoring and evaluation systems into four principal components[5] that appear particularly well suited for utilization in

5. Though in this guidebook the words *inputs, outputs, outcomes* and *impacts* are used to describe the principal project components to be monitored and evaluated, there is a wide range of terms currently being used in project M&E. Other frameworks have used terms such as performance/processes/impacts; provision/utilization/coverage/impact; and inputs/processes/outcomes/impacts. Regardless of which terms are used, it is necessary that definitions be clear and that those involved in the M&E of a particular project (e.g. project staff, external evaluators, and donors) understand which vocabulary is being used.

## A Framework for Monitoring and Evaluation

Inputs ⟶ Outputs ⟶ Outcomes ⟶ Impacts ⟶ Benefits

| | Assumptions | | Assumptions | | | |
|---|---|---|---|---|---|---|
| Resources used to support the primary activities of the project. | Expectations regarding the effectiveness and quality of the project inputs. | The delivery of goods and services. | Expectations regarding the ways these goods and services will be used by the target population. | Changes in behaviors/ practices.<br><br>Links provision of goods and services to impact. | Nutritional status measures.<br><br>Effects resulting **directly** from project outputs or **indirectly** through outcomes. | Broader effects.<br><br>Effects resulting from the achievement of impacts, usually in combination with other factors. |

nutrition projects. Beyond its specific M&E utilization, such a framework is a useful tool for project design because it provides a means for planners and other staff to (a) articulate how they anticipate project inputs and activities will achieve the desired effects, (b) reach consensus on the details of the project, and (c) clarify the terminology that will be used for their particular M&E system. Dividing a project into various components also makes it easier later to identify the specific constraints to project effectiveness.

Monitoring focuses on the appropriate and timely provision and use of project resources focusing primarily on **inputs** and **outputs**; evaluation focuses on whether the expected **impacts** were achieved. Both monitoring and evaluation systems assume that before the project was implemented, the designers conducted a problem or situation analysis, determined the proximate and underlying causes of the problem they wanted to address, and developed the project to deal with these specific causes, linking project ***inputs*** with desired ***outcomes*** and ***impacts***. In the context of this analysis of the linkages from inputs to outputs to outcomes to impacts, financing is considered a given; the project's inputs are the resources and services purchased with project funds.

***Inputs*** are the materials, goods and actions necessary to carry out the primary project activities. These include items to be delivered to the target population (e.g. food, micronutrient capsules or injections, education materials), training of project personnel, and preparation of project sites or equipping factories. Timely availability of adequate equipment and supplies needed for project implementation—such as trucks and gas for the delivery of food supplements or warehouses for food storage—are also inputs into the implementation of the project. Monitoring should yield information that can answer questions regarding the procurement, production, delivery (to project staff) and costs of these resources. For example, are vitamin A capsules being delivered to project staff on time and in the proper quantities? What percent of targeted community nutrition workers (CNWs) have received training? To what extent do the project's actual input costs coincide with the budget plan?

There often will be more than one task associated with a particular input. The delivery of training to community nutrition workers, for example, requires prior recruitment, development and production of materials, and perhaps the training of trainers. Each of these sub-components can be monitored under the heading of "inputs". (As discussed in Section 2, more complex and sequential input systems may require their own flowchart or input tracking system such as that illustrated in Table 2.1.)

***Input assumptions*** are the expectations regarding the effectiveness and quality of the project inputs (e.g., vitamin capsules have not lost potency; fortification equipment is installed properly) *and* the expectations regarding the process of getting these inputs to the output or delivery stage (e.g., adequate numbers of weighing scales and growth charts have been delivered to project sites; the target population has been properly identified).

With respect to effectiveness and quality assumptions, the following example may be useful. In a project that includes training of village health workers there is an assumption in the overall design of the project that

the workers will understand the concepts and techniques taught and will be able to convey this information in a useful way. If that assumption is false, the likelihood of achieving effective outputs will, in turn, be adversely affected. Accordingly, it might be useful to monitor this assumption by collecting information periodically on the effectiveness of training activities. Similarly, if an intervention uses food supplements, it is assumed that the rations produced are of adequate quality and caloric density. In the case of nutrition projects that focus on behavioral change through nutrition communication, it is crucial that the educational methods be appropriate for the target population, and that they be targeted to the behaviors that need to be changed. As with inputs, input assumptions can be systematically monitored.

With respect to process assumptions, monitoring may include a mechanism to track the placement of necessary staff and the delivery necessary inputs at each service delivery point. Monitoring might also include checks on target populations selected. This may be less important where target groups are more easily identified, e.g. pregnant women, children under two years of age, but more important where the target is food insecure households, requiring a transparent process with reliable and well understood indicators.

***Outputs*** refer to the provision of project goods and services to the target population; these constitute the primary project activities. The types of questions that can be answered with information on outputs include, How many of the children eligible for the project were given vitamin A capsules last month? What percent of pregnant mothers in the project area received iron folate supplements? How many of the targeted school age children received deworming medication in the last six months?

***Output assumptions*** are those made about the target population and their utilization of the goods and services received. Though the delivery of inputs and services may run smoothly, positive outcomes will only result if certain assumptions about the target population are met. Do they

understand the messages? Do they have the resources to put them into practice? Can the environment support the intervention? For example, do project participants consume enough fat to permit the efficient conversion of beta-carotene into retinol? Does the provision of food supplements for women and children increase their total daily caloric intake? (Even efficient delivery of a supplement—the **output** will not lead to improved growth—the **impact**—if the supplement substitutes for food normally consumed at home rather than confirming the **output assumption**, namely that the supplement will be additive and increase total daily intake.)

Although many output assumptions will have been addressed in the design stage of a well prepared project (e.g., pre-testing of nutrition education messages will have addressed resources, limitations and literacy concerns), their critical importance and the possibility of changed conditions may suggest the value of periodic monitoring.

Information on both inputs and outputs should be collected routinely, and information on input and output assumptions at least periodically, to monitor a project's operations and thus inform management decisions. In cases where monitoring data yield shortcomings, these indicators can help pinpoint design and implementation weaknesses of the project. Monitoring inputs and outputs is also necessary to provide context for midterm and endpoint evaluations.

The terms "outcome" and "impact" are commonly used by development practitioners in an interchangeable fashion. It may be useful, however, to distinguish between intermediate outcomes and more substantial impacts. ***Outcomes***, as defined here, are the intermediate effects resulting directly from project outputs that may be necessary to achieve a desired impact. In many nutrition projects, outcomes take the form of behavioral changes in the target population, such as improved child feeding practices or more equitable intrahousehold food distribution resulting from Nutrition Communication efforts. These behavioral change outcomes may

then translate into improved nutritional status, which would be considered the impact.

Outcomes can also include intermediate changes in the conditions of the target population. If, for example, a nutrition project designed to improve the growth of children supplies deworming medication in addition to other inputs, an intermediate "outcome" would be a decrease in parasitic load. The **outcome** in a take home food supplementation program for children would be the actual consumption of that food by the child.

In general, even nutrition projects that have been conscientious about M&E, often fail to collect information on intermediate outcomes. It should be noted that for some nutrition interventions there is no measurable intermediate outcome; instead the output leads directly to the desired impact. This is the case in projects that distribute vitamin A capsules (impact), where improvements in micronutrient status result directly from taking the capsule (output) (See Annex 2, Table A2.1)[6].

***Impacts*** are the more meaningful changes in the condition of the target population and generally reflect the primary objectives of the project. For nutrition purposes, it is generally convenient to speak of impact in terms of change in nutrition status using anthropometric and micronutrient status indicators. How has the prevalence of iron deficiency anemia among women changed as a result of iron supplementation? What effect has the project (perhaps directly through on-site food supplementation- an output—or through nutrition counseling—an output, resulting in behavioral

6. In areas with high rates of infection, this model might also include an additional assumption to the right of the outcomes column (or an additional output assumption if there are no outcomes), that a food or nutrient consumed will be adequately absorbed by the body's intestinal tract. Where infection rates are particularly high, or where one nutrient's absorption is seriously inhibited by the deficiency of another, this problem may well have to be addressed in order for outputs or outcomes to translate into impacts.

change—an outcome) had on the incidence of wasting in targeted children under three years? What effect have food supplements to pregnant women with low body mass index (BMI) had on the incidence of low birth weight (LBW)?

Information on outcomes should be collected during the life of the project through ongoing monitoring or, alternatively, through special studies, making clear whether service delivery has had some first-level effect on the quality of life of the target population, such as increased food intake or improved caring practices. In cases where outcomes do not become apparent until the later stages of a project, they can be assessed along with impact indicators as part of an endpoint evaluation. Together, information on intermediate outcomes and final impacts is used to make decisions on the future of the project

Finally, ***benefits*** are the broader, sustainable changes in public health or economic status that a program seeks to achieve but which are inevitably influenced by a wide range of other factors. These benefits, such as decreased infant, child, and maternal mortality, improved economic conditions resulting from greater productivity, and increased lifespans, may not be seen until many years after the project is completed. Benefits usually are not included as indicators of project success unless there is a special interest (and corresponding funding) for such information, usually in a long-term research context. Fortunately, much has already been learned and documented about the associations between improvements in nutrition status (e.g. improved growth, higher birth weights and decreases in micronutrient deficiencies) and improvements in mortality, morbidity and productivity.[7] Accordingly, equipped with a particular set of evaluation-generated impact data, project personnel may be able to make projections on a range of benefits likely to accrue.

7. See, for example, tables 6A and 6B of the World Bank Toolkit #3 (Phillips and Sanghvi 1996).

Overall information on inputs and outputs should be collected regularly as part of a project's monitoring system and entered into a management information system. Input and output assumptions also should be monitored, although usually less frequently and often through special studies. Outcomes (often) and impacts (always) from participants and control groups, should be included in evaluations. This allows determination of the extent to which observed changes between baseline data and data collected subsequently among participants, can be attributed to the project. At the same time, it will almost always be useful to include "snapshots" of outcome and impact variables among project participants, even without control group data, as part of a monitoring system (rather than having to wait two and a half years for a midterm evaluation). In many nutrition projects impact "snapshots" can be taken using growth monitoring, pregnancy weight gain, or birth weight information regularly collected and comparing it both with other project areas and with data collected from the same area over the course of the preceding year (ideally with the same month in the previous year to assure seasonal consistency) . Outcome data, usually behavioral in nature, and often collected through special studies, is also important to monitor with some regularity. If a nutrition communication project designed to increase food consumption during pregnancy is not having this effect on participants, the project staff should know this and be acting on it well in advance of a formal evaluation.

It should be noted that this model assumes reasonably homogeneous projects from area to area. In a monitoring system, monitoring forms, data collection regimens and MIS indicator columns assume a discrete set of inputs, outputs and other categories of information which will be reasonably constant across project areas. Similarly a well functioning evaluation assumes that the indicators used in baseline and evaluation surveys will be the same in each area. What can be done, then, in projects, such as the Iringa Nutrition Project in Tanzania which place high premiums on local determination of project activities, and where, as a result, project activities vary from community to community?

In such cases, M&E options appear to be two-fold. First, if the locally determined activities or activity combinations fall into a small number of categories, and if the project as a whole is large enough so that evaluation sample size requirements will be met, each of these activity combinations can be considered a cohort, with particular sets of monitoring data collected in each, and with evaluation data analyzed separately for each (permitting also a comparison of the cohorts). Where activities from area to area are too heterogeneous, project monitoring will have to devolve largely to the local level, while project evaluation will necessarily be limited to assessing the composite impact of this heterogeneous set of activities on a pre-determined set of impact indicators.

# 2 Clarifying Project Goals, Objectives and Information Needs



To set the stage for the development of an effective monitoring and evaluation plan, it is necessary to determine what planners hope to achieve through the project, how resources and activities will be used to meet project goals and objectives, and the ways in which monitoring and evaluation will be used to enhance the project's capacity to accomplish these and other aims. A project task manager who sets in motion a well-organized M&E system will often find that system providing the multiplier benefit of sharpened objectives, better articulated assumptions, and greater clarity in project staff's understanding of the project.

### *Pre-Design Tasks*

The clarification of nutrition project goals and objectives presupposes a general understanding of malnutrition problems and their causes in the targeted geographic areas. This information, often collected through surveys, may be supplemented with secondary data collection, focus group sessions, and interviews designed to identify the constraints and resistance points that the project will need to overcome. Examples of positive deviance, i.e. households who have found means of improving nutritional levels despite prevailing social or economic constraints, may also be studied to inform project design. Although not technically part of the M&E process, this preparatory information-gathering and synthesis will lead to the development of appropriate inputs and activities as well as the establishment of realistic objectives. These steps, in turn, will permit identification of specific indicators of project efficiency or effectiveness that are essential for a properly functioning M&E system.

### *Clarifying Project Goals and Objectives*

Whether one is developing an M&E plan during the design phase of a project, establishing a system after implementation has begun, or evaluating an ongoing project, the clarification of goals and objectives is es-

sential. Monitoring and evaluation staff should never assume that project goals and objectives are clear or realistic, even if a project has been implemented for years. Additionally, while some goals and objectives remain constant throughout the life of a project, others may evolve or need to be redefined periodically. In practice, goals and objectives often need to be clarified or sharpened in response to the requirements of an M&E system.

To establish or clarify goals and objectives, which provide the primary focal points for monitoring and evaluation efforts, it is important to be able to differentiate between the two concepts and understand how they are used in project planning and M&E.

***Goals*** are the broad aims of the project, the significant, longer-term changes that planners expect to occur in people's lives. The reduction of severe protein-energy malnutrition, the improvement of childcare practices among young mothers, the enhancement of food security in single parent households, and the significant reduction of iodine deficiency disorders are all examples of project goals. Because goals do not specify concrete expectations for achievement or the criteria which will be used to measure project success, it is necessary to break down goals into objectives and, in turn, activities (outputs) which will contribute to achieving these objectives.

***Objectives*** are operationalized goals which specify the results and the level of change expected. Objectives allow a comparison of what is eventually accomplished in the project with what designers had originally set out to achieve. In order to be useful for M&E, objectives should adhere to the following key criteria:

These guidelines for developing clear, useful objectives have been applied to the following nutrition and nutrition-related project examples. As illustrated, it may be useful to specify not only impact objectives, but also the outcomes and/or outputs necessary to achieve them.

**Objectives Should Be:**

| *Measurable* | *Well-Defined* | *Realistic* |
|---|---|---|
| Objectives must specify the criterion for success in quantifiable terms, an expected magnitude of change, and a time frame in which the result should occur. It is important to note that even qualitative information can be translated into measurable objectives. | Each objective should be matched with a single, precisely defined indicator for success. There should be little room for individual interpretation of objectives. Each term in the objective should be examined carefully to be sure it is defined in a meaningful and operational way. | Objectives should be challenging but achievable within the specified time frame. Translating project goals into realistic objectives requires an understanding of the magnitude and causality of the problem within the target population and a review of relevant literature to determine what success rates have been achieved in similar interventions elsewhere.* |

*In project development one often finds a "chicken and egg" situation between specifying objectives and carrying out a baseline survey. In the absence of recent reliable survey data, it may be difficult to specify desired unit or percentage changes in the objectives and sub-objectives prior to the collection of baseline data. At the same time, development of a baseline survey is difficult in the absence of specified objectives. In practice, objectives often require at least some refinement once baseline survey results become available.

***GOAL 1: Reduce malnutrition in children under two years of age***

- Objective 1.1: Reduce the prevalence of wasting (WHZ< 2) in children under two years from 30% to 10% in five years
  - "Required Outcome" 1.1.1: Increase the percent of infants, aged 6 months to 9 months, who receive complementary foods (in addition to breast milk) from 20% to 80% in five years

- Objective 1.2: Reduce the prevalence of low birth weight infants (< 2,500 g) from 190 per 1,000 to 130 per 1,000 in five years
  - "Required Outcome" 1.2.1: Increase the average daily caloric intake of pregnant women from 1800 to 2100 calories within this time period
  - "Required Output" 1: Provide daily calorie-dense food supplement to 90% of pregnant women with BMI < 18.5
  - "Required Output" 2: Provide nutrition counseling messages on the importance of increased food consumption and rest to 90% of pregnant women

***GOAL 2: Significantly reduce vitamin A deficiencies in the target area***

- Objective 2.1: Reduce the prevalence of keratomalacia in the area covered by the project by fifty percent (from 20% to 10 %) by the end of the third project year
  - "Required Output" 2.1.1: Increase the coverage of the massive dose vitamin A supplement program from 25% to 75% by the end of the third project year

***Mapping the Project***

Once the project's goals and objectives have been established, the focus of M&E turns to the identification of project inputs and outputs and their conceptual linkages to desired outcomes and impacts. A critical step for project planners and M&E design staff is to organize this information logically into a model of the project, using the M&E framework discussed in the overview. While such a model can take the form of a matrix, graph, or set of mathematical equations, a conceptual framework in the form of a diagram or map works particularly well for nutrition projects.[8]

8. Such a conceptual framework is sometimes referred to as a dynamic model, a cause-and-effect model, or an input-output model.

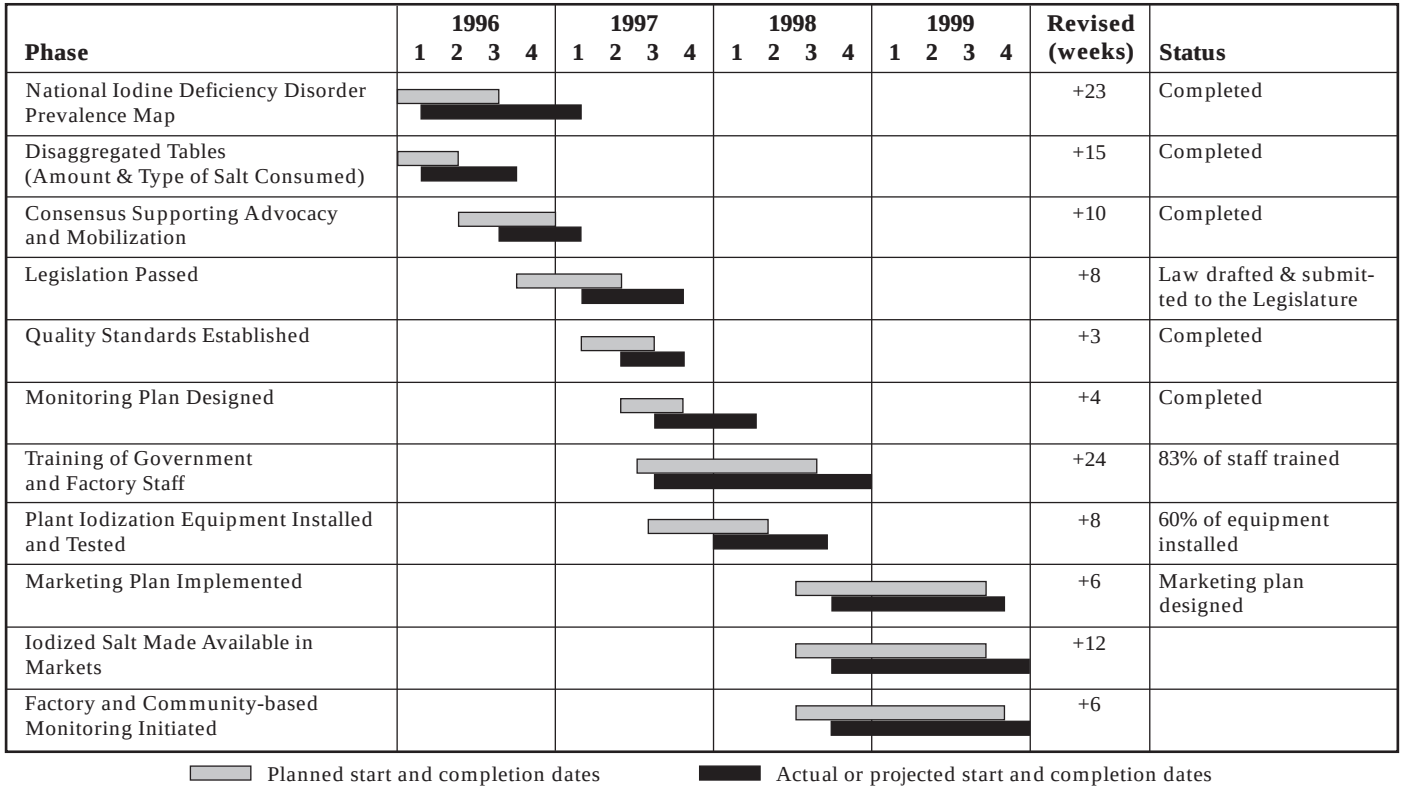
**Conceptual Framework of a Simplified Nutrition Counseling Project**

Inputs ⟶ Outputs ⟶ Outcomes ⟶ Impacts ⟶ Benefits

| | Assumptions | | Assumptions | | | |
|---|---|---|---|---|---|---|
| • In-service training of Community Nutrition Workers (CNWs)<br>• The design, production, and delivery of nutrition counseling materials | • CNWs understand concepts and are motivated to participate | • Provision of nutrition education to the target population | • Target population understands concepts, is able, willing, and motivated to participate<br>• Targeted behaviors are a constraint to child growth. | • Improved childcare practices (e.g. increased duration of breast feeding, age-appropriate introduction of complementary foods, increased food intake during | • Reduced malnutrition in children under 3 including reduced incidence of LBW, and reduced prevalence of wasting and/or stunting in children under 3 | • Reduced childhood mortality and morbidity<br>• Increased long-term productivity |

To initiate this process, a list of the project's main activities should be compiled. In the simplified nutrition counseling project that is modeled above, the provision of nutrition information to target families is the project's primary activity. All of the project resources required to generate this activity, i.e., the design, production and delivery of educational materials, and the in-service training of community workers, should be listed under the *input* column. It should be noted that these inputs may themselves be processes which in some projects should be carried out sequentially. Where this is the case, the conceptual framework should be accompanied by a flow chart indicating the sequence and timing of these input processes. Table 2.1 illustrates an input process which might

## *Table 2.1* Using a Two-Bar Gantt Chart to Plan and Monitor Main Phases in a Salt Iodization Program*




*Note:* Column headings indicate year and quarter.

*A monitoring system for a salt iodization project might include the following: internal (factory level) quality assurance; external physical inspection of salt using rapid test kit for presence or absence of iodine and rough levels or using titration method through a laboratory for accurate levels in parts per million; examination of factory production and distribution records; community level testing through schools; and examination of individuals in schools or health centers for clinical signs of IDD.

*Source:* Adapted from Valadez & Bamberger (1994).

take place in a salt iodization project. Next to the inputs one should specify any inherent assumptions or expectations about the quality and effectiveness of the inputs themselves that may influence the success of project activities. In the example used here, one such assumption is that community nutrition workers (CNWs) will be motivated to incorporate the specified messages into their work with the community. By identifying and then monitoring the assumptions on which project success rests (both for inputs and outputs), managers can more easily pinpoint and fix faulty links in the process.

Moving further to the right on the map, the project's services to be delivered are listed as *outputs* along with their assumptions. Ideally, output assumptions are derived from pre-project research[9] on local attitudes, and the population's motivation and ability to participate in the project and make behavioral changes. If such research has been well designed and implemented, it will inform project design and reduce the likelihood of unmet output assumptions.

The flow of outputs leads, in turn, to the intermediate *outcomes* and final *impacts* that the project is expected to achieve. Typically in nutrition projects, the outcomes will be stated as behavioral changes while impacts will be stated as changes in nutritional status. Often outcomes and impacts are derived directly from the objectives of the project. As indicated earlier, however, not all interventions will have intermediate outcomes. This is particularly true in the cases of Vitamin A supplementation and some cases of on-site feeding where service delivery—the output—should, assuming sound design, lead directly to the desired impact with no intermediary step necessary. In contrast, such intermediate actions or

9. Developing and pre-testing messages is sometimes referred to in communications projects as formative evaluation. The term "evaluation" used for such pre-project activity and implying assessment of customs, beliefs, and understandings of a population is sometimes confusing. In this manual we limit the use of the word "evaluation" to assessments of project effectiveness.

outcomes are necessary in an iron supplementation program for women (i.e. daily compliance) and in take home supplement activities (i.e. actual consumption by the target woman or child)[10].

Finally, long-term *benefits* should be listed in the right-hand column even though these are not usually assessed as part of project M&E.

This conceptual framework is fully compatible with the logical framework (log frame) model used by many development agencies. In that model, the goals to be achieved, the project purpose (i.e., steps to achieving the goals), outputs and inputs create the vertical structure of the log frame while a narrative summary, objectively verifiable indicators of whether the goals, purpose, outputs or inputs have been achieved, means of verifying the indicator and important assumptions provide the horizontal. This grid structure assists planners to design each step of the program model, beginning with the desired impact and then working backward through each step needed to achieve it.

While the log frame is a comprehensive tool for organizing project details, it is much more complicated to use and explain to stakeholders than the conceptual framework. The conceptual framework offers a useful and easily interpreted method for organizing a project.

Although this discussion has progressed through the framework from left to right, it should be clear from this discussion that the process of sound

10. This nutrition counseling example assumes a “behavioral change communications” (BCC) project which seeks to change a specific identified set of sub-optimal behaviors, as opposed to more generalized “nutrition education” which seeks to provide information to recipients, often on a broad array of subjects, but without specified behavioral change objectives. An M&E system for a BCC activity would identify as outcomes the specific behaviors to be changed, leading, in the impact column, to particular nutritional status improvement expected as a result. In the case of a broader based nutrition education project without specified behavioral change objectives, an M&E system would be limited to assessing inputs and outputs, and then assessing whether knowledge had increased as a result (an output assumption).

project design should move in the opposite direction, i.e., specification of objectives (impacts and sometimes outcomes), followed by specification of the goods or service delivery (outputs) necessary to achieve these objectives, followed by specification of the inputs necessary to bring about the outputs.

### *Determining Information Needs*

#### *What needs to be included in the M&E system?*

Now that each component of the project has been identified, the framework forms the basis for a project's M&E plan and can be used to ensure that all the essential components of the project are covered. (We shall see, in Section 9, that the conceptual framework can be revisited at the point of analysis to identify weak links inhibiting project effectiveness). It is generally not practical or feasible, however, to monitor and/or evaluate every detail that appears on the framework. Rather, it is necessary to prioritize information needs, decide which pieces of the framework should be studied on an ongoing basis, which can be left unchecked or checked only occasionally unless a problem arises, and what questions can be better answered through the use of special studies.

In designing a monitoring and evaluation system there are often difficult choices to make in deciding which indicators ought to be regularly monitored as opposed to being less frequently evaluated. Does it make sense, for example, to include changes in practices, an outcome variable in a nutrition education project, in an ongoing monitoring system or only in an evaluation? What about knowledge levels and attitudes that ultimately affect changes in practices?

In such cases there may be tradeoffs between a desire to have such information on an ongoing basis in order to make adjustments in the inputs as necessary, and the burden that such ongoing data collection may place on already overworked local level staff. Both ease of collection and ongoing need for the data are likely to be prime criteria in making

such decisions. Where data are both necessary and difficult to collect internally on a regular basis, a contract with a local consulting firm, a nutrition institute or a university may be the answer.

Once decisions are made on what information needs to be collected, both monitoring and evaluation plans need to be developed. Each should specify tasks to be undertaken, individuals responsible, and expected timing (including, where appropriate, duration and expected starting and completion dates). Annex 3 provides a list of tasks that might be included in such plans. The evaluation plan might well utilize the Gantt Chart discussed earlier in this section.

***What information is better collected and analyzed in special studies?***

Questions often arise in a project that are best answered through special studies (sometimes referred to in the M&E literature as diagnostic studies) rather than through routine monitoring and evaluation. Besides providing very specific information, special studies offer project staff several important advantages. First, it is possible to hire experts who specialize in the task at hand (e.g. cost-effectiveness analysis, assessment of nutrition communication, female education, or participatory assessment techniques). Second, by relegating sub-tasks to special studies, particularly the more difficult or time consuming ones, the M&E system can be a more manageable undertaking for project staff. Finally, if the collection and analysis of data are divided into smaller pieces, quality control is often easier to maintain.

Special studies can be closely linked with M&E and included as part of the overall assessment plan, or they can be separate. Below are various types of studies which have been employed in nutrition projects in conjunction with monitoring and evaluation.

- *Disaggregation studies*: As will be discussed in Section 3 (Designing a Monitoring System), disaggregated information on at-risk groups (e.g. girls, religious or ethnic minorities, or female-headed households) and non-participants can be extremely important for improving target-

ing and for tailoring project inputs and implementation to meet the needs of subsets of the population.

- *Longitudinal studies*: A second approach to studying special at-risk groups involves longitudinal data collection on a sample of individuals from particular population subgroups. This strategy was used successfully by the first Tamil Nadu Integrated Nutrition Project (TINP I) in India (see the Field Insight below). High-risk groups to be tracked may include groups with potentially high rates of non-participation (e.g. time-constrained female-headed households, those living at some distance from the service site, or low caste or minority religion households) or those who are less likely to be able to benefit from project inputs (e.g. extremely food insecure households).

- *Program constraints assessment*: At some point during project implementation, and particularly if M&E demonstrate that implementation is not running smoothly, it may be useful to carry out a participatory, qualitative stock-taking of the project. In this special study, constraints inhibiting project effectiveness are identified and the information organized to permit the identification of technical, policy, research, and training means of addressing them[11].

- *Operations Research*: In order to address a particular implementation problem, outside institutions are sometimes contracted to carry out relevant operations research in project areas. If, for example, project monitoring reveals unusual resistance among mothers with respect to caring practices, operations research might test a nutritional message targeting mother-in-laws and measure its effect on the motivation and ability of mothers to improve caring practices. While not formally a part of M&E, such operations research is often triggered by information from project monitoring and is closely associated.

---

11. Given the emerging importance of this methodological tool, the Program Constraints Assessment (PCA) is described further in Annex 1. See also Levinson et al., *Nutrition Program Constraints Assessment: Gauteng Province, South Africa* for more information on this type of study.

- *Qualitative studies*: There may be important aspects of a project which are not amenable to quantification and cannot be easily included into either ongoing monitoring or periodic evaluations. These could include the extent and quality of community mobilization and community participation, attitudes about complementary feeding practices, job satisfaction of field workers, and intra-household resource allocation. It may be easier to study such issues using qualitative research methods such as Participatory Rural Appraisal[12] (PRA), Rapid Rural Assessment (RRA)[13], focus groups, or Knowledge, Attitudes and Practice (KAP) surveys[14]. Conducting such studies require specialized training.

Once goals and objectives are clearly specified, a conceptual framework firmly established, and initial decisions made on those indicators which best lend themselves to ongoing monitoring, periodic evaluation or special studies, it is possible to begin designing each of these systems. Section 3 discusses design of the monitoring system.

---

12. PRA enlists people from the project area to appraise local conditions, identify development constraints and provide strategies to address them adequately. It stresses the important insights community members can bring to a project if given the opportunity.

13. RRA involves community members in a series of techniques including developing chronologies of local events; making case studies of people or situations; mapping of the area; and using ranking and scoring methods. These techniques can provide a wide range of results quickly and accurately. In nutrition programs, RRA can be useful in linking factors influencing nutritional status unique to the project area. Such qualitative data can usefully supplement and contextualize the quantitative data more regularly collected by M&E systems.

14. KAP surveys are administered to assess knowledge, attitudes and practices related to specific topics. Nutrition related topics include knowledge of which foods contain vitamin A, attitudes toward introduction of complementary foods, and feeding practices when children are ill. KAP surveys can provide valuable information in assessing the needs of communities and determining whether messages are being understood.

Experience from development projects has demonstrated that the early establishment of a monitoring system to collect and utilize information for purposes of project improvement is crucial for the successful achievement of project goals and objectives.

In addition to information on project inputs, outputs and results, it is necessary to monitor project costs for purposes of accountability and subse-

**Field Insight: Lack of monitoring renders a vitamin A fortification law ineffective**

In an effort to decrease vitamin A deficiency, legislation requiring vitamin A fortification of hydrogenated vegetable oil (vanaspati) at 25 IU retinol per gram of oil was passed by the Indian government. However, the program has been ineffective due to the absence of monitoring and therefore of enforcement of the fortification law.

Monitoring systems permit examination of the extent to which fortified products conform to national standards at points of production and consumption.

Another interesting point in this example is that the fortification targeted the wrong group. The government recognized later that vanaspati is mainly consumed by the upper and middle classes, and is too expensive for those with the lowest vitamin A status. Furthermore, the average consumption of vanaspati is reported to be 3 grams per capita per day, contributing only about 4% of the recommended daily allowance of vitamin A. Thus, failure to monitor the inputs and outputs of this fortification program resulted in management missing three essential issues. (1) Fortification levels were inadequate; (2) the fortified product did not reach the target populations; (3) even among consumers, level of consumption was too low for the program to be effective.

quent cost-effectiveness analysis. While it is possible for small projects to rely on a single management information system (MIS), large-scale projects may benefit from establishing parallel systems to manage the monitoring of costs, project implementation, and nutrition and health status.

This section provides an overview of monitoring by describing the step by step development of an MIS for project improvement, the use of ratios in monitoring, ways of using monitoring results, the planning the flow of information, and financial monitoring.

### *Monitoring for Project Improvement with a Management Information System*

As indicated, timely and accurate information on implementation is central to sound project management. This information is generally collected and maintained in a computerized[15] project management information system (MIS), by the project staff who implement project activities, and/or by their immediate supervisors. Some information, however, may be collected by higher level administrators with responsibility for quality control of data collection; supervisory checks on project sites, implementation and field staff; and data on delivery of resources to project headquarters.

Because much of the information needed for nutrition project monitoring, i.e., attendance records, feeding records, growth charts, medical records, and community mapping data, are routinely collected and compiled as part of project implementation, this information should be rela-

15. Decisions about the use of computers versus manual management and compilation of data are likely to depend on the availability and maintainability of computer hardware, the availability of project staff with computer operating skills, and the possibility of protection against heat, dust, humidity, power surges, and theft. Effective computerization includes programs that check data as it is entered, automate calculations (rather than having them made by field staff), utilize commercially available software, and automatically back up data frequently.

tively easy to collect, record and process at the field level as long as the quantity of required data is reasonable.

*Simple disaggregation* of information, by gender, ethnicity, religion, or type or location of household, can be extremely useful for monitoring the participation and effects of the project on various subgroups within the target population. Without such information, aggregate figures on the community may miss entirely the fact that, within a community, a particular subset of individuals are not participating in or benefiting from the project. Here, however, designers of a monitoring system must be particularly sensitive to the workloads of local staff. If such disaggregated information is desired, a special study might be commissioned using a consultant or institution hired specifically for that purpose, as outlined in Section 1. Alternatively, as suggested in Section 1, it may be appropriate in large projects to have an ongoing external (to the project) M&E entity responsible for the periodic disaggregation of monitoring data as well as quality checks on the data and special studies plus baseline and impact evaluation data collection and analysis.

As indicated, project monitoring and, in turn, the MIS, require information on project inputs, outputs, and “snapshots” of results. The challenge for those who design an MIS is to collect information on each of these project components (identified initially by development of the conceptual framework or project map) in a way that is not overly burdensome for field staff and is useful for project management.

When deciding which outcome and/or impact measures to include, those designing an MIS system need first to determine:

- what information already will be available through project implementation;
- what information will be most useful for project management; and
- which results can be feasibly monitored on an ongoing basis by field staff.

Keeping in mind these questions, and the primary requirement that this information be fully used for ongoing project improvement, the following subsection outlines one method for designing an MIS.

***Step 1: Assessing the Potential Value of Monitoring Information***

Selection of indicators and the collection and utilization of data on these indicators is the essence of project monitoring and, hence, an MIS[16]. However, it is important to decide whether collecting this information is worth the effort it requires, and it is also important to determine whether field staff can collect reliable information or if a special study would provide better information.

The assessment table provides a means to review potentially collectable information and to decide whether it ought, in fact, to be collected as part of a project's MIS system.

1. Using the Conceptual Framework as a guide, list the components which *could* be included in a monitoring system in column 1.
2. List the type of information required for each Conceptual Framework component in column 2.
3. Note whether field staff can reliably collect the information required or whether a special study is necessary in column 3.
4. Finally, it is important to consider the management/planning value of dedicating their time to this activity. If the management value is not immediately apparent, it may be best to exclude data collection. List whether collection of data will provide important management information in column 4.
5. If field staff can collect the information, and if the management value of collecting the information is great, enter the information to be collected (column 2) in the "Project Activity" column of the "Monitoring Record Keeping Chart."

16. Note that a fuller discussion of indicators is contained in Section 6.

## Assessing the Potential Value of Information Collection: Examples

| 1. Components of the conceptual framework being considered for inclusion in monitoring system* | 2. Information to be collected which will allow assessment of each component | 3. Possible for field staff to measure or requires a special study? | 4. Will the benefits of data collection during implementation offset the extra effort required? |
|---|---|---|---|
| Increased intake of major nutrients during pregnancy | Dietary intake surveys | Special study | No (the value of such detailed information may not justify the expense and time required to collect it on an ongoing basis. Alternatives may be collection of caloric intake data on a sub-sample or collecting qualitative data on intake changes resulting from the project) |
| Improvement in weight gain during pregnancy | Weight gain during pregnancy (from records of pre-pregnancyweight and last monthly weight prior to delivery) | Special study | Yes |
| Reduction in percentage of low birthweight births | Birthweight | Field Staff | Yes (food supplementation for women is very expensive; early indications that it is successful justify costs; early indications that it is not successful may prompt further investigation into service delivery) |

*This can include "snapshots" of impact indicators, e.g., assessments of child growth among participants although without reference to the control group, hence without assurance that any improvement seen is attributable to the project.

***Step 2 : Monitoring Record Keeping Chart***

The Monitoring Record Keeping Chart is used to assign monitoring record keeping duties to project staff.

1. To use the Chart, list all the individuals and groups who are involved directly or indirectly with project implementation in the columns under "Persons involved in project activity."

2. List each project activity (from the Conceptual Framework) and additional monitoring information to be collected (from step 1: Assessing the Potential Value of Monitoring Information) in the "Project Activity" column.

3. Determine the role of each involved individual and group and enter one or more of the codes provided.

4. For each activity, assign a project staff member the duty of keeping records in the "who will record" column. It is important to consider literacy/numeracy skills and ability to store record keeping books when assigning duties. More than one individual can be assigned record keeping duties if "double recording" could provide useful information—for example, if a community nutrition worker and a health worker both record medical referrals, it is possible to monitor both whether field staff make medical referrals and whether beneficiaries actually visit a health worker as a result.

5. Review the entire chart to assess the level of record keeping responsibility being given to any particular person or group. Consider this person/group's workload, the project's ability to supervise their work, and the total amount of record keeping duties being assigned. Too much record keeping may lead to poor quality information). If any individual or group has been assigned too many duties for collecting monitoring information, assign the duty to another appropriate individual/group

based on the codes given for their participation—for example, a supervisor may be able to collect some information to reduce the record keeping burden on the person/group conducting the activity.

***Step 3. Monitoring Information Summary***

The "Monitoring Information Summary" helps determine the kinds of registers that will be needed for each person who will collect information.

1. In the left-hand column, list each person who has been assigned record keeping duties from the "Monitoring Record Keeping Chart."

2. List the type of information they will collect in the next column (also from the "Monitoring Record Keeping Chart").

3. Considering the activities each person performs, list the events at which they will be able to collect the information.

4. Finally, summarize the types of information each person will collect in the right-hand column. The list in this final column will enable you to make specific forms/registers for each event at which an individual will collect information. It may be possible to combine several registers into one for each individual. Similarly, if one form is suggested for collecting several types of information on several types of participants, it may be convenient to create different forms for each type of beneficiary (e.g., a form for mothers and a separate form for children).

***Step 4: Creating the Record Forms***

Now that the types of forms required are known, it is important to consider any other additional information that will be necessary. For example, the weighing session form requires birth dates because they are necessary to determine weight for age (or height for age), which is needed for "status" determination. In addition, unique identifiers, such as

# Monitoring Record Keeping Chart

| Project Activity | Possible Information | Persons involved in project activity* | | | |
|---|---|---|---|---|---|
| | | CNP | Women's Group | Health System Field Workers | CNO |
| Growth Monitoring and Promotion | Weights Coverage; Status (growth, nutritional) | P | A | | S |
| Monthly Weighing | Weights Coverage of Pregnant Women | A | | P | |
| Child Feeding | Coverage; Individual attendance | P | A | S | S |
| Medical Referral | Coverage; Follow-up | P | | P | S |
| Iron Supplement Administration to Pregnant / Lactating Women | Coverage; Compliance | P | | S | |
| Food Preparation | Adequacy of supply; Quality standard; Costs; Profits | S | P | | I |
| Vitamin A capsule administration to newly-delivered women | Coverage | P | | S | |
| Village Nutrition Management Committee Meetings | Occurrence | A | | | |
| **Outcomes to be Monitored** | | | | | |
| Reduction in Birthweight | Weight; Measurement coverage | P | | | |

*Assignment Codes (P = Performs; A = Assists; S = Supervises; I = Informed). This personnel breakdown is taken from the Bangladesh Integrated Nutrition Project (BINP). "CNP" is the Community Nutrition Promoter at the village level. "CNO" is the Community Nutrition Officer who supervises several CNPs. "ATFPO" is the Assistant Thana Family Planning Officer responsible for particular management tasks at the thana or sub-district level. "DPD" is the Deputy Project Director. "NGO Admin" is the non-governmental organization administrator.

(Adapted from Hamilton, D. and U. Gaertner, *Goal Oriented Project Planning (GOPP): An Introduction to the Methodology*, GTZ-RPMAS, TG-PMC, UNDP / DTCP, 1992)

| Persons involved in project activity* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Village Nutrition Management Committee | ATFPO / Thana Manager | Thana Nutrition Management Committee | District Nutriton Management Committee | DPD (section) | Consultant (section) | NGO Admin. | Project Director | Who will record? |
| I | I | | | | | | | CNP |
| | | | | | | | | Health Worker can record, but how to report? |
| I | I | | | | | | | CNP |
| I | I | | | | | | | CNP; Health Worker |
| I | I | | | | | | | CNP |
| I | | | | | | | | CNO (CNP too busy; women's group members not literate) |
| I | I | | | | | | | CNP |
| | | | | | | | | CNP |
| | | | | | | | | CNP |

## Monitoring Information Summary

| Person to Collect Information | Activity | Location of the Activity | Information to Be Collected | Forms/ Registers Needed |
|---|---|---|---|---|
| CNP | Growth monitoring and promotion (weight from growth cards) | Monthly weighing | Weights; Coverage (number weighed as a % of total children); Status (growth faltering or severely malnourished) | Weighing session report |
| | Child Feeding | Daily feeding session | Coverage (number fed as % of number eligible); Individual attendance | Feeding register |
| | Medical Referral | Monthly weighing session | Coverage (number referred as % of number eligible; number treated as % of number referred) | Weighing session report |
| | Monthly weighing of pregnant women | Monthly weighing session | Weights; Coverage (number pregnant women weighted as % of total number pregnant women) | Pregnancy form |
| | Iron supplement to pregnant and lactating women | Monthly weighing session | Delivery coverage (number pregnant women given tablets as % of number weighed) | Pregnancy form |
| | Vitamin A capsule Administration to newly-delivered women | At birth | Coverage (number women receiving within 2 weeks as % of number of births) | Birth form |
| | Village Nutrition Management Committee (VNMC) Meetings | Monthly VNMC meeting | Occurrence | Monthly report |
| | Low Birthweight | At birth | Birthweight within 48 hours | Birth form |

individual identification numbers, names, and mother's name, are necessary so that follow-up is possible.

It is also important to consider how the forms can be used in the field. For example, a form with several months side-by-side may make it easier to spot recurring problems, such as a child who continually relapses into severe malnutrition.

Finally, it is important to test the forms in as realistic a situation as possible before putting them to use. Service record forms should follow standard guidelines for creation of data collection instruments.

Illustrative forms found in Annex 4, are based on those used by Community Nutrition Promoters (CNPs) in the Bangladesh Integrated Nutrition Project (BINP). BINP addresses two important causes of child malnutrition in Bangladesh: 1) mothers do not understand the relationship between adequate growth and a child's well-being, relying instead on developmental stages like ability to stand or walk,[17] and 2) mothers often do not realize the importance of providing complimentary food for their children at six months. Through monthly weighing it is possible to identify children whose weight is faltering, and those who do falter receive a small daily food supplement that demonstrates to mothers 1) that their child's weight or growth is not adequate, and 2) for mothers whose children should be receiving complimentary food but are not, that a small, affordable amount of solid food can significantly and quickly improve their child's growth. In combination with counseling and personal attention for each mother and child, this project has succeeded in reducing severe malnutrition (< 60% NCHS median weight for age) substantially in the first year of operation.

The "Weighing Session Form" was developed after using the process described for developing monitoring forms. CNPs conduct monthly weighing

17. Zeitlyn, Sushila. *Feeding Practices in Bangladesh with special reference to pregnant, postpartum and lactating women and infants and children: A review of the literature.* No date: UNICEF, Dhaka, Bangladesh.

over two days each month, and record children's weights, their nutritional and growth status, and medical referrals. Nutritional and growth status are very important to monitor because 1) children who are severely malnourished need medical referral, and it is important to track whether these referrals take place, and 2) faltering growth and severe malnutrition status are eligibility criteria for feeding, and the project needs to track these to ensure that children enter the feeding when needed and to ensure that ineligible children do not receive feeding (resulting in leakage and additional costs). The CNP is best suited to maintain these records, as she performs the weighing and thus is present at the weighing session, has at least some secondary school education, and a place to keep the record books. Her record keeping burden is not insignificant, but, after proper training, the CNPs do not report that the burden is excessive.

***Step 5: Using Monitoring Data for Project Management***

Once the data is collected on forms, the information is examined and used as fully as possible at the local level by local staff and village management committees, assessing, for example, how attendance at the child weighing, and food supplementation sessions has compared with that of previous months, and, where deficient, what steps are necessary to improve it.

The data is then sent to the next data assessment point, in Bangladesh the union level, where data from all of the villages in the union are examined and compared. Using the all important "management by exception" principle discussed later in this section, union level staff identify those villages where predetermined minimum achievement levels in coverage or even in results (using impact snapshots) have not been achieved, and initiate management action to address these shortfalls. (As will be indicated, such a review can also identify villages with exceptional performance and reward staff accordingly).

Data from all these villages is then aggregated into a union data set and sent on to next data assessment points (in Bangladesh, the thana and

central levels) where this process is repeated. In most cases this process up to the central level will be done by hand on paper, and then at the central level transferred to a computer program at which point a final "management by exception" review would take place followed by the preparation of a national summary report, possibly on a quarterly basis. (See Annex 4 for an exercise in "management by exception").

At each stage of the MIS process, most indicators being employed will be most easily used if they are presented as ratios. The following field insight presents some important ratios being used by BINP.

***Responding to Monitoring Results***

As previously stated, monitoring information is most beneficial to project staff when it is used to correct problems and improve implementation. One way in which project managers can maximize the effectiveness of the monitoring system is to include, for particular indicators, specified levels of substandard performance which would "trigger" an automatic management response.

The principle of management by exception can quickly demystify the sometimes paralyzing issue of MIS data utilization. The notorious under-utilization of nutrition project monitoring (and surveillance) data is at least in part the result of an absence of clarity on what to do with it.

The management by exception principle argues that at each level of review examining data from the level immediately below (e.g. in Bangladesh, Union level review examining village data or Thana level review examining Union level) no action need be taken in response to this data *unless* it indicates that particular units below (e.g. villages, unions, or thanas) have not met minimum achievement levels as measured by the trigger points. In these cases, the data should "trigger" management action to address the deficiency. Areas once "triggered" should then be examined carefully in subsequent reviews to assure that problems have been corrected and do not recur.

## Field Insight: The Use of Ratios in Project Monitoring

By using a few simple ratios to monitor key aspects of a project, a monitoring system can be considerably enhanced. Ratios are easy to record at the field level and can be converted later into percentages. In the Bangladesh Integrated Nutrition Project (BINP), for example, a simple, direct reporting system has been set up as part of the ongoing monitoring system. Each month, the Community Nutrition Officer (CNO) reports on a pre-printed postcard the following four ratios for each village under her supervision:

1) Growth Monitoring Coverage Ratio:

$$\frac{\text{Number of under-two children weighed}}{\text{Total number of registered under-two children}}$$

A drop in this ratio indicates a problem of outreach and coverage within that particular village. In response, the CNO will visit the village to determine, in conjunction with the Community Nutrition Promoter (CNP), the cause of the reduction. Possible constraints may be the CNP's failure to motivate mothers, the location of weighing sessions too far from particular households, inconvenient scheduling, shortfalls in equipment or materials (e.g. growth cards), or inadequate reporting.

2) Unsatisfactory Nutritional Status Ratio:

$$\frac{\text{Number of children eligible for feeding each month (due to growth faltering or severe malnutrition status)}}{\text{Number of under-two children weighed}}$$

This ratio measures the prevalence of growth faltering and / or severe malnutrition among children participating in the project. It enables the CNO (and higher levels of administration) to target project inputs, time, and inter-sectoral activities where problems are most severe. In addition, it provides higher levels of management with a constantly updated picture of nutritional status in project areas, particularly if the first ratio demonstrates that a high percentage of the target population is participating in the project.

3) Feeding Coverage Ratio:

$$\frac{\text{Actual number of child-days of feeding}}{\text{Number of expected child-days of feeding if all eligible children attend all feeding days}}$$

This ratio represents either the ability of mothers to participate in the project or the level of motivation to use project services. Low ratios may indicate that mothers

do not have time to bring their children to the center for feeding, that they are prohibited by other factors, that they are not convinced that their children are in trouble, that they do not have faith that the supplement will be effective in increasing their child's weight, or possibly that there is a failure in the supplement provision system. In any case, the CNO and the CNP can respond to low ratios by investigating the reasons and taking appropriate action.

4) Overall Coverage Ratio:

$$\frac{\text{Actual number of children registered}}{\text{Estimated total of age-eligible population}}$$

This ratio provides an estimate of the number of age-eligible children registered for growth monitoring. It allows the CNO to interpret the feeding coverage and nutritional status ratios. For example, if 50 children are weighed out of the 50 children registered, we calculate a growth monitoring coverage proportion of 1.0 or 100% coverage. We may conclude that growth monitoring coverage is outstanding in this area. However, perhaps only 50 of the 300 age-eligible children are registered. This reveals a less than desirable coverage rate.

By using these simple ratios, project staff are able to easily identify, at the village level, changes in effective project coverage and outreach over time (ratio 1), monthly variations in the nutritional status of each community's participating children (ratio 2)*, and trends in program participation of children with low or faltering growth (ratio 3). *These ratios provide rapid and continuous feedback to inform interested stakeholders whether the project is on track. If there are problems, adjustments can be made and specific problem areas can be further investigated and corrected without waiting for an evaluation.*

---

*In the case of such monthly prevalence data, sometimes represented as a "community growth chart", the most appropriate comparison in most countries would not be with the preceding or following month, which would require adjustments for normal seasonal variation, but with the same month in the preceding or following year.

Source: Adapted from World Bank. 1980. *Tamil Nadu Nutrition Project Implementation Volume.* Washington, DC: The World Bank, Population, Health and Nutrition Department (adapted in turn by BINP).

***"Trigger" Points***

Based on the conceptual framework of the project, it is possible to identify key inputs and outputs, and in some cases for outcomes and for impact "snapshots" for minimum achievement levels for purposes of "triggering" remedial management action when minimum levels are not achieved. Once these trigger points are identified, it is possible for planners to develop pre-determined action that can be implemented to correct poor performance. Upon receipt of a monitoring report indicating, for example, that the percentage of eligible children attending weighing sessions in a particular community fell below the trigger level of 80%, a pre-determined management response could be set in motion. Table 3.1 provides some examples of indicators suitable for trigger responses

## *Table 3.1* Sample Trigger Indicators for Automatic Response to Monitoring Results

| Project Component | Possible Trigger Indicators* |
|---|---|
| Growth Monitoring | • Percentage of eligible children attending weighing sessions |
| Feeding Supplementation | • Percentage of eligible children attending feeding sessions<br>• Percentage of children graduating from feeding<br>• Monthly availability of food supplements |
| Micronutrient Supplementation | • Supplement coverage ratios<br>• Monthly availability of micronutrient supplements |
| Household Food Security | • Percentage of identified food insecure households participating in the project's food security interventions<br>• Percentage increase in household real income, available food, or caloric intake in participating households** |
| General | • Number of training sessions held<br>• Home visit frequency<br>• Percent over or under annual budget<br>• Timeliness of salary payments<br>• Timeliness of delivery of monitoring reports |

*Specific trigger points for each of these indicators would be established by project management

**Likely to require a special study.

For trigger points to prompt effective action, there should be minimal delay between collection of the data and this trigger point analysis which should be carried out regularly at each level of review. It should be added that establishing meaningful trigger points for such indicators may be difficult to do until the project has been operating for some time.

### *Automatic Response to Exceptional Performance*

A progressive monitoring system might also identify trigger points which reflect exceptional implementation success (e.g., inputs are delivered ahead of schedule or under budget), efficient service delivery, and/or exceptional staff performance. This "positive deviance" approach at the project level offers management the opportunity to learn valuable lessons from successful performance. In addition, such identification provides important material for project reports and newsletters and an opportunity to give credit to staff who deserve it. Reporting of such successes can become another automatic management response to MIS data. Periodic recognition and rewards, for such success also can provide incentives and boost morale.

### *Monitoring the Budgetary Health of the Project*

Financial monitoring, or project bookkeeping, is essential to good project management. It is also generally required by donors for purposes of accountability. Even if it is not, project management must keep track of the project's budget. In addition, a well-managed financial monitoring system will provide cost data which will be needed for subsequent project evaluation, most specifically in determining the project's cost-effectiveness (discussed in Section 9). Any system designed to monitor project costs should serve to answer the questions:

- What are the project's actual costs and do they correspond to the budget plan?

- If not, which components of the project are over and under budget and why?

Because information on expenditures is usually managed by project administrators rather than implementers, financial monitoring is often carried out separately from the MIS, particularly in large projects. In donor-assisted projects, the details of budgetary monitoring will often be based on procedures specified during project negotiations. The best systems are those which distinguish between capital and recurring costs and which provide disaggregated expenditure groupings which can be reaggregated according to need[18]. Well functioning systems also permit estimates of annual expenditures which combine recurring costs and annualized capital costs. Any such system must list both the planned and actual costs.

It is usually desirable to be able to calculate the total annual cost of a project so that this can be compared with the cost of other national projects and with comparable projects elsewhere using annual cost per beneficiary figures. In terms of recurring costs, this involves nothing more than straightforward addition. With capital costs, however, the process is more complicated. Simply adding the capital expenditures made each year would make annual costs for the early years of a project—when most of the vehicles and equipment, to be used over the life of the project, are purchased—unfairly high. Accordingly, instead of adding the face value of capital expenditures each year, project accountants normally calculate the *annualized* capital cost each year for each capital expenditure. This annualized capital cost reflects the *depreciation* of the

18. While the terms are, in practice, often used interchangeably , “costs” in this discussion refers to budget line items (estimates or actual spending which has taken place), plus the estimated value of items which have not required full payment. “Expenditures,” by contrast, relates only to money spent. Accordingly, “costs” here includes not only budgetary expenditures, but also imputed costs (for example, of non-budgeted ministry staff—paid with funds outside of the project—who spend time working on the project) and “opportunity costs” (the real value of volunteer labor, rent-free buildings, or donated food). “Budgets” normally include only those cost estimates for which funds will have to be spent.

vehicle or equipment that has taken place during the course of a year, or the amount by which its market value has been decreased.

Depreciation tables, by type of capital good, are usually available from government or World Bank economists, and need not be calculated from scratch. Not surprisingly, depreciated values usually decrease from year to year. A vehicle will be worth substantially less in Year 2 than it was at the time of purchase, but the Year 3 value is not likely to drop by as much. By adding the total of these annualized capital costs to the recurring costs for a particular year, one can derive a reasonable and usable annual cost figure to use for comparability purposes, and ultimately for cost-effectiveness studies.

Below is a simplified, illustrative example of how a budgetary table might be designed using a spreadsheet package. Note that only a small fraction of total expenditure items have been listed here. Annual tables detailing actual monthly expenditures should be used and then fed into a more comprehensive project table (like the one below) for a longer-term view of budgetary compliance. Where components are notably different from the amount originally budgeted (either above or below projected costs), administrators will need to ascertain why this is the case, and adjust project spending and/or the budget accordingly. For monitoring purposes, a project which involves multiple sites must collect budget information from each site separately, so that sites can be identified as underspending, on target or overspending. This identification will assist managers in providing guidance to sites which exceed or fall below the target budget.

***Planning the Flow of Information***

The way in which monitoring information is used for decision-making varies at each level. At the implementation level, for example, decisions on logistics, time allocation, and individuals or groups in need of special attention need to be made by the community, project beneficiaries (or

## *Table 3.2* Sample Budgetary Table of a Three-Year Project

| Expenditures | Costs by Project Implementation Year | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Year 1 | | Year 2 | | Year 3 | | Project Total | |
| | Planned | Actual | Planned | Actual | Planned | Actual | Planned | Actual |
| ANNUALIZED CAPITAL COSTS | | | | | | | | |
| Vehicle, HQ* | | | | | | | | |
| Computer equipment, HQ | | | | | | | | |
| Office furniture, HQ | | | | | | | | |
| Weighing scales | | | | | | | | |
| Community center construction | | | | | | | | |
| RECURRING COSTS | | | | | | | | |
| Staff salaries and benefits | | | | | | | | |
| Telephone bills and other utilities | | | | | | | | |
| Office supplies | | | | | | | | |
| Training materials | | | | | | | | |
| Food supplements | | | | | | | | |
| Micronutrient supplements | | | | | | | | |
| Growth charts | | | | | | | | |
| Total Capital Costs | | | | | | | | |
| Total Recurring Costs | | | | | | | | |
| Subtotal | | | | | | | | |
| Price Escalation | | | | | | | | |
| GRAND TOTAL | | | | | | | | |

*Headquarters.

their parents), and the community level worker. Supervisory staff at higher levels need to ensure smooth logistics support, including the recruitment, placement and training of staff, the supply and maintenance of materials and equipment, and monitoring of the quantity and quality of service delivery. At each administrative level, relevant information should be used for project management and only the information needed by the next level of decision-makers should be compiled, aggregated and transmitted upward.

The frequency with which data are collected, compiled and analyzed will depend on information needs. Typically, field level reports are compiled on a monthly or quarterly basis in order to provide continuous and frequent feedback for project improvement. Quarterly, biannual and annual reports compiled at the headquarters level should be timed to coincide with other known deadlines such as due dates for project renewal proposals, interim reports, or budgetary requests.

The flow chart on the following page, which has been filled in with an illustrative monitoring example, can be used to help organize the flow of information.

## *Table 3.3* Information Flow for Monitoring

| | Information → | Information → | Information → | Information |
|---|---|---|---|---|
| **People/Staff Responsible** | **Community Nutrition Worker (CNW)** | **Union Nutrition Coordinator (UNC)** | **Statistics Officer** | **Project Director** |
| Information Level | Village | Union | District | National |
| Information-Related Duties | • Collect information on individual children<br>• Note any exogenous factors which may affect program results<br>• Summarize growth findings on a community growth chart<br>• Determine the need for supplies and materials<br>• Summarize findings and report to union | • Check each VQMR for errors<br>• Convert numbers into percentages<br>• Conduct spot checks of each village at least once every six months<br>• Summarize information and report to district | • Check each UQMR for errors<br>• Aggregate information<br>• Conduct spot checks of union offices at least once every six months<br>• Conduct periodic data quality control checks at all levels<br>• Summarize information and report to the national office each quarter | • Check each DBMR for errors<br>• Aggregate information<br>• Conduct spot checks of district offices at least once every six months<br>• Summarize information in an annual report |
| Information Uses | • Use growth information to counsel mothers whose children are malnourished or faltering<br>• Follow up on children who were not weighed<br>• Use information to keep community informed of overall progress | • Follow up on villages that have high percentages of malnutrition, growth faltering, or non-participation<br>• Work with the CNW to identify the cause(s) of the problem(s) and make necessary changes | • Follow up on unions that have high percentages of malnutrition, growth faltering, or non-participation<br>• Work with the UNC to identify the cause(s) of the problem(s) and make necessary changes<br>• Use information and follow-up to ascertain which unions are successful and why | • Use information to follow up on problems and successes |

| | | | | |
|---|---|---|---|---|
| Instruments Used | • Individual growth monitoring and promotion charts<br>• Community growth chart<br>• Community profile register<br>• VQMR | • UQMR | • DBMR | • Annual management report |
| Timeframe & Delivery | • Deliver VQMR to Union Office by the 1st of February, May, August and November | • Deliver UQMR to District Office by the 1st of March, June, September and December | • Deliver DBMR to National Office by the 1st of May and November | • Complete annual report by January 15th of each year |

Note: This simplified example focuses on the growth promotion component of a hypothetical nutrition project. It does not include information on other project components. "CNW" is community nutrition worker. "UNC" is union nutrition coordinator. "VQMR" is village quarterly management report. "UQMR" is union quarterly management report. "DBMR" is district biannual management report.

# 4 Selecting an Evaluation Design



While ongoing monitoring is essential for project management and improvement, periodic evaluations are necessary to draw conclusions about the effects of the project on target populations. Valid evaluations serve two important functions: they determine (a) *the extent to which desired changes have occurred*, and (b) *whether the project is responsible for such changes.* This information, in turn, allows those who plan, implement and fund nutrition projects to make sound judgments on the future of the project.

There are various ways in which evaluators can collect and analyze information, although most large scale projects require a formal evaluation plan involving initial collection of baseline data on a representative sample of the target population, one or two midterm evaluations and a final impact evaluation. This section seeks to explain the benefits and drawbacks of various evaluation designs so that managers, project planners, and M&E staff can make informed decisions concerning the type of evaluation best suited to their needs and resources. The section tackles questions which face evaluators again and again: Is a baseline survey necessary? Is a control group necessary? What do we sacrifice if one or both are absent? What fall back positions are available in their absence? As with other aspects of both monitoring and evaluation, choosing an evaluation design is often a case of finding a balance between the ideal and the practical.

This section and the next will discuss impact evaluation only, although it is understood, as indicated earlier, that the implementation process must also be reviewed to assure that implementation has proceeded properly. This guidebook argues strongly for an ongoing monitoring system which can serve this very purpose, requiring only assurance (provided through quality checks of the monitoring data) that the data is reliable. In those cases where a monitoring system is not in place, a “process evaluation” would be necessary to review those records which exist on the monitoring variables discussed in Section 3, collect single point in time data on variables, e.g., outcome variables and assumptions, where no records are

likely to exist and rely considerably on qualitative data. As should be clear, this is far less satisfactory than an ongoing monitoring system which collects such information regularly over the life of the project.

### *Designs for Evaluation*

#### *The experimental design*

An evaluation must employ data collection and analysis procedures that provide useful and valid information on the effects of a project. Various evaluation designs can be employed to accomplish these tasks. The most rigorous of designs, which, if properly conducted, will provide the most irrefutable conclusions, is called an **experimental design**. Although the term "experiment" may be an uncomfortable one when applied to nutrition projects, all projects are in a sense experiments because they test the hypothesis that the project intervention will improve the nutrition status of the target population.

A hypothesis of an experiment is a proposed explanation for particular occurrences. The key to experimental design is *elimination of competing hypotheses* or non-project explanations for change, and in turn, to isolate the core hypothesis that project activities are responsible for positive changes in nutrition status. For example, a nutrition project may be built upon the hypothesis that improved child caring practices lead to reduced childhood malnutrition. An evaluation may show that nutrition status indeed does improve in the project area after several years of project operation, but, without a proper evaluation design, there could be alternative reasons to account for the effect. One interpretation could be that nutrition status simply improved over time due to several good years of rainfall; another hypothesis could be that observed differences in nutrition status were due only to improved skill in weighing children among data collectors. With a good design, it is possible to eliminate these competing hypotheses and attribute observed changes to the project itself.

Three factors are necessary to establish an experimental design: *control*, *randomization*, and *pre-post analysis*. “Control” refers to the identification of a similar group of individuals to compare to those who will be participating in the project. “Randomization” is the assignment of people to the project or control group where each person has an equal chance of being assigned to either group. “Pre-post analysis” means that baseline data collected at the beginning are compared with data collected at some point after project implementation, often the mid-point or end of the project. Comparison with a randomized control group allows evaluators to measure net-change and attribute it to the project activities (assuming no biases). The pre-post analysis provides information on the magnitude of change that has occurred, and corrects for any time-related trends.

*Attributing change to the project: why a control group is necessary*

Creation of a control group for comparison with participants eliminates several different types of competing hypotheses resulting from forces outside the project. Because a group exists that shares the same characteristics of the participant group, *changes* that are observed in the project group and not in the control group *can be attributed to project participation* (with the exception of observed changes due to bias, to be discussed later). An evaluation may show, for example, that nutrition status among project participants has improved over the project period. It is possible, however, that the same change would have occurred had the project not taken place at all. Improved education or health care, introduction of deworming medications, increases in agricultural production, changes in food prices, or other factors totally independent of the project could have caused the improvement in nutrition well-being. The only way to determine whether the project itself was responsible for the observed changes is to compare changes observed in project participants with changes in a comparable group affected by the same external factors, but not receiving project services.

### *Table 4.1* How Control Groups Help to Eliminate Competing Hypotheses

| Sources of Competing Hypotheses | Competing Hypotheses | Benefits of Comparison with a Control Group |
|---|---|---|
| Historical or Secular Effects | Changes in nutrition status reflect general changes in the population which are due to exogenous factors (e.g., a drop in food prices, several good crop years). | Such nutrition status changes in the general population are reflected equally in participant and control groups. |
| Maturation | Changes in nutrition status reflect general changes related to the fact that individuals are different at the end of the project duration than they were at the beginning. | Natural changes occur equally in participant and control groups. |
| Testing | Changes in nutrition status reflect changes in the way nutrition status is measured. | Measurements are equal in participant and control groups. |
| Regression to the Mean | Changes in nutrition status are due to the tendency for improved nutrition status to improve among any group with poor nutrition status. | Regression to the mean is equal in participant and control groups. |

Additionally, control groups can demonstrate positive project effects in situations of deteriorating nutrition status. In the case of a severe drought, for example, an evaluation might conclude that nutrition status had actually deteriorated in the project areas. Comparison with a control group, however, may indicate that while nutrition status did decline in the project areas, it declined even more in the control areas. Without a control group, evaluators might be led to conclude that the project had harmful or negative effects on the target population when, in reality, it was successful in preventing more serious nutrition deterioration.

*The benefits of random selection*

In order to make valid comparisons, participant and comparison groups should be identical in all respects save one—project participation. The only way to ensure an identical control group is to randomly assign individuals to project and control groups. Random selection is an unrealistic (and unethical) choice for many field-based nutrition projects, although a *comparison group* can usually be found from a comparable area where the project has not yet begun activity. When a comparison group is not randomly selected, however, the possibility for error in the form of either *bias* or *confounding (*external factors which may influence findings and lead to erroneous conclusions) increases.[19]

This guidebook, designed primarily for task managers, will not seek to discuss the intricacies of either bias or confounding. At the task manager level it may be adequate to understand that data collection and analysis techniques can be used to control for possible bias and confounding, but that randomization is the only sure way to minimize these potential problems.

---

19. Although it is inappropriate to select *individuals* for a nutrition intervention through a randomized control design, it is both appropriate and sensible to use such a design in selecting districts, communities, schools, or health or daycare centers to be participants or controls (assuming normal programmatic context in which, for limited financial or human resources, it is not possible to include all such groups in the project at the outset), or when program services are to be phased in. In the latter case, random assignment determines which groups receive services first. In the case of district selection, where the total number may be small thereby creating statistical problems, a matched comparison can be introduced to strengthen the evaluation design. However, matching is difficult because it is seldom possible to match on more than one or two characteristics of the many which are possible confounders. Randomization controls for confounding factors that are not known, or not measurable, therefore, cannot be used as a basis for matching. For additional information on the use of randomized control designs in developing countries, see Newman et al., 1994.

*Measuring the magnitude of change: why a baseline is necessary*

As indicated earlier, pre-project information is usually necessary to establish a need for the project and to formulate specific objectives. It cannot be assumed, however, that a needs assessment can double as a baseline survey unless the data collected includes each of the relevant indicators, is geographically disaggregated (separating project and control areas), and is followed immediately by the initiation of project services (rarely possible).

For purposes of an evaluation, pre- and post-project information is necessary in order to demonstrate *if and to what extent change has occurred*. In a simple before-after (or pre-post) design, *baseline* measurements, which determine the pre-intervention status for the selected indicator(s), are compared with *follow-up* measurements taken either during project implementation (for a midterm evaluation) or upon project completion (for an endpoint evaluation).

Because decisions regarding the future of the project are often based on the degree to which project activities positively affect the lives of targeted individuals, it is important for stakeholders to have quantified information on project effects. Without baseline information, it is not possible to determine the magnitude of change that has occurred.

***Quasi-experimental evaluation designs***

As mentioned above, randomized control groups are rarely a feature of nutrition projects, or of nutrition project evaluations. Some nutrition projects are implemented without first collecting baseline data. In other cases, evaluations may focus on a project that has been in place for many years, conducting a baseline survey after several years of implementation would not provide pre-project information. In cases where no randomized controls exist or where a baseline survey has not been conducted, the challenge of evaluation design is the *approximation* of true

experimental conditions. Designs that eliminate competing explanations without the benefit of a true control group are called **quasi-experimental** designs, and examples are provided below, in declining order of rigor, or ability to eliminate competing hypotheses explaining observed nutritional effects. These are followed by two alternative quasi-experimental designs which, may also be used in nutrition projects.

*Pre-post design with a non-randomized control group*

Because a true experimental design is rarely possible, evaluations which (a) compare baseline and follow-up data and (b) use *non-randomized control groups* are considered acceptable for field-based nutrition projects.

When a randomized control group is not possible, a comparable group of individuals, from an area not served by the project can be used as a non-randomized control group as long as pre-project measurements are available for both participant and comparison groups. If properly executed, this design allows changes in nutrition status to be attributed to project effects and enables quantification of the magnitude of these effects.

Care should be taken to select comparison individuals who are as similar as possible to the participant group—a complicated procedure requiring careful selection. Consider, for example, a nutrition project attempting to reduce prevalence of childhood malnutrition by providing nutrition education to mothers who attend a health clinic for prenatal care. To evaluate project effects, it is necessary to determine whether the percentage of malnourished children in a population not receiving the nutrition education intervention is different from that in the population of mothers receiving the nutrition education. As indicated in the previous example, mothers from the general population will not serve as valid non-randomized controls because mothers who attend the clinic for prenatal care are different (e.g. they may be more motivated or have more free time) than

the general population of mothers. If the rate of malnutrition were lower among children whose mothers received the nutrition education intervention than among the general population of mothers, it would be impossible to be sure whether the low prevalence of malnutrition was due to the intervention or to other factors characteristic of women who seek prenatal care.

In this case, a population of mothers who attend the same clinic or a similar type of clinic for prenatal care is necessary for comparison. If a similar health clinic exists where the nutrition education intervention is not given, then children of mothers who come to that health clinic for prenatal care could provide a valid non-randomized control group. Another option might be to compare malnutrition rates among children previously born to participant mothers *before* the nutrition education intervention with malnutrition rates among children born *after* the intervention began. However, this "previous pregnancy" comparison group would not enable the evaluator to eliminate competing explanations related to history (malnutrition rates are different in children from more recent pregnancies), maturation (women change from one pregnancy to another, physiologically or otherwise), or other external factors.

*Pre-post design without a control group*

Without a valid control group for comparison, changes in nutrition status cannot be explained by project effects alone—history, maturation, dropout bias, and other factors could be responsible for observed changes. In this situation, it sometimes is possible to construct a reasonable basis for comparison using data collected periodically on another population for other purposes. Using outside data sources, however, requires the evaluator to be confident that the external data has come from a population equivalent to the participant population.

If external data are not available for valid comparison, the evaluator can use a pre-post design without a control group and carefully examine

factors external to the project to judge if these influences were likely to have had a notable positive or negative effect on nutrition status. Even if this task is undertaken conscientiously, however, conclusions always will be uncertain because of the difficulty in isolating the effects of these external factors, and the possibility of missing some altogether. Box 4.1 enumerates several ways in which a 1994 impact evaluation of TINP I dealt with this problem.

*Post-project design with a non-randomized control group*

When a baseline survey has not been conducted, an evaluation involving a non-randomized control group requires pre-project information to establish that both groups were the same before the project began. If pre-project information does not exist, the evaluator's subsequent assertions on project effects rest on the assumption that participants and controls were identical before the project began—often a dubious and indefensible assertion. The field insight below describes the problems which arose in seeking to evaluate the Nutrition Communication and Behavior Change component of the World Bank- assisted Indonesian Nutrition Development Program.

*Post-project design without any control group*

Without any pre-project information, and without a control group, an evaluation is simply a description of participant status after receiving services. While such descriptive analyses can be useful for project staff to better understand aspects of project implementation and client satisfaction, they do not provide conclusive evidence about the effects of the project.

### *Alternative quasi-experimental designs*

When an evaluation takes place without the benefit of a control group and/or pre-project information, the evaluator must look creatively for other

## Box 4.1 TINP I Impact Evaluation: Evaluating without a Control Group

Although the first Tamil Nadu Integrated Nutrition Project was originally designed with control areas, each of these areas was over time, incorporated into the project. While the phasing of areas into the project allowed for use of the "institutional cycle design" (discussed below under alternative quasi- experimental designs) the evaluation also relied heavily on data collected for other purposes. This included the following:

1. A comparison of yearly data for TINP and non-TINP areas (although the latter had not been incorporated as controls) using non-project data collected by the National Nutrition Monitoring Bureau (NNMB). The data found declines in the percentage of malnourished young children much sharper in the TINP areas.
2. An analysis of year by year NNMB data on the prevalence of underweight young children in the entire state. The analysis found that the significant decline in underweight children occurred after 1983, i.e., after TINP started, and that there had been practically no change in the years preceding TINP's activity.
3. A comparison for z scores for TINP areas with All India data collected from other sources. The latter indicated that the percentage of All India children more than two standard deviations below the reference mean had declined by 0.7 percentage points per year in the 1980's compared with 1.25 to 2.4 percentage points per year in TINP areas.

The evaluation also examined the likely effect of other possible explanations for observed improvements in nutritional status among project participants. The effect of income changes using per capita increases in GNP and elasticity data (finding that no more than 25 to 50 percent of the declines in malnutrition in TINP areas would likely have occurred in the absence of the project). The possible effect of the other large-scale nutrition project, the Midday Meal Program, was also considered and found to be negligible for children under age 2.

Although, in the absence of a control group, each of the comparisons and analytical procedures carried out is problematic, their combined effect becomes reasonably convincing.

---

*Source:* Adapted from World Bank. 1994 *Impact Evaluation Report : Tamil Nadu Integrated Nutrition Project*. Washington, DC: The World Bank, Operations Evaluation Department.

## Field Insight: Evaluation Problems Arising from the Absence of a Baseline Survey

The Nutrition Communication and Behavior Change component of the World Bank-assisted Indonesian Nutrition Development carried out in 1980, is rightly cited as a landmark nutrition education project. Evaluating the project has been exceedingly difficult, however, because of the absence of baseline data on participants and controls—and, accordingly, the inability to assure that the control group was comparable.

Project design staff originally hoped to compare the Indonesian community-based nutrition program, Family Nutrition Improvement Program (UPGK), with and without an innovative nutrition education campaign of counseling by "kaders" , village-selected volunteers trained to conduct weighing sessions and educate mothers, reinforced by radio spots. Without baseline data, however, it was difficult to assure that subsequent differences in nutritional status were attributable to the nutrition education campaign.

In addition, evaluators found, after the fact, service delivery differences between the participant (UPGK plus nutrition education) and control (UPGK only) groups above and beyond the nutrition education campaign:

- workers in participant villages were paid; in control villages they were unpaid
- workers in participant villages were more likely to be male, slightly older, had more program experience, reached larger numbers of women and children, and made more visits per household than their counterparts in control villages.

While the analysis as a whole leaves little doubt that the project was successful in improving caring practices and, in turn, nutritional status, the magnitude of these effects is difficult to measure in the absence of baseline data and comparable controls.

---

*Source:* Zeitlin, Marian F., et al., 1984. "Vol. IV, Household Evaluation, Nutrition Communication and Behavior Change Component, Indonesian Nutrition Development Program." Washington, DC: Report to the World Bank.

evaluation designs that can seek to isolate the role of project activity on observed changes in nutrition status. While such designs lacking control and/or baseline may be open to competing explanations for observed effects, they have the potential to eliminate most other explanations and may provide convincing evidence concerning project effects. Two of these alternative designs are presented below.

*Time series design*

When randomized controls do not exist, it may be possible to eliminate some competing hypotheses normally eliminated by control groups by using a time series design. In a time series design, a series of nutrition status measurements of participants before and after project implementation enables the evaluator to isolate project effects. Several measurements taken before the project begins establish trends related to history and maturation, then several measurements taken after the project commences establish the same trends after the intervention. If the only major discontinuity in the trend is found at project intervention, the assumption that the project is responsible for the observed effects is difficult to refute.

*Institutional cycle design*

For most nutrition programs, the decision to provide services is based on need, rather than considerations of randomization for experimental design. However, projects often implement service delivery in phases, making it possible to base random selection on *when* a group will receive services, instead of *who* will receive them. When randomized implementation phasing is possible, a series of before-implementation measurements as each phase begins can be combined to provide a randomized control group, and after-implementation measurements can be combined to create a randomized participant group. This was in fact done in the 1994 TINP impact evaluation. Because the project was introduced in different districts in a phased manner, “timed controls” could be used. The baseline data in each district added were used as an estimator of the nutrition situation without project inputs.

## *Table 4.2* Overview of Various Evaluation Designs

| Evaluation Design | Change Can Be Attributed to Project Activities | Magnitude of Change Can Be Determined | Comments |
|---|---|---|---|
| (1) Pre-post design with a randomized control group (experimental design) | Yes | Yes | The most rigorous evaluation design, yet often not feasible or ethical for field-based nutrition projects. |
| (2) Pre-post design with a non-randomized control group | Yes, if the control group is a valid comparison group. | Yes | Generally the most rigorous design possible for field-based nutrition projects. If the control group is valid and the evaluation is properly conducted, strong conclusions can be reached. |
| (3) Pre-post design without any control group | No | Yes | This design can be strengthened by examining exogenous factors which may have influenced the nutritional status of project participants. Another possible method is to establish a comparison group after the fact using secondary data sources. |
| (4) Post-project design with a non-randomized controlgroup | Possibly though conclusions are often weak. | No | Conclusions about the effects of the project can only be drawn if pre-project information demonstrates that project and control groups were the same in terms of relevant variables before the project began. |
| (5) Post-project design without any control group | No | No | Descriptive assessment of project. Can yield useful information but no conclusive results. |
| (6) Time series design | Possibly | Yes | This design can provide strong conclusions if properly planned and executed. |
| (7) Institutional cycle design | Yes, if implementation phasing is randomized rather than based on need. | Yes | This method is comparable to pre-post with a control. |

### *Summary of evaluation designs*

Experimental design allows evaluators to draw conclusions about the magnitude of a project's effect while eliminating other explanations that could account for changes in nutrition status. True experimental designs are rare in evaluation studies, but various forms of quasi-experimental designs allow confident conclusions without demanding the rigor one would expect in a clinical study. Table 4.2 provides an overview of the various designs discussed in this section.

When it comes to planning the technical aspects of an evaluation, project design and management staff, and even some professional evaluators, are often inclined to turn matters over to a statistician. While the services of a skilled statistician are usually necessary, particularly for determination of sample size and sampling strategy, M&E managers and staff should recognize that each of these technical decisions involves practical considerations and will have important effects on results obtained. Accordingly, the steps involved are outlined below in a highly abbreviated fashion.

The most important of these decisions involves specifying: the objectives of the evaluation, appropriate groups to sample, and the most practical methods for collecting the needed information.

The following discussion of evaluation planning assumes an ideal situation in which the evaluation design (a) is developed prior to project initiation, permitting collection of baseline data, and (b) includes the collection of information on both participant and control groups.

### *Planning an Evaluation*

#### *Principal Steps to Prepare for an Evaluation*

In addition to the broad steps outlined in this guidebook for all M&E, the following specific steps must be taken to plan and carry out a midterm or endpoint evaluation:

1) Determine the sample

2) Identify the control group

3) Choose the sampling frame

4) Collect pre-project information through a baseline study

Once these preparatory steps are accomplished, M&E staff can collect midterm and endpoint data which will be used to draw conclusions about the project's effectiveness.

### *Determine the Sample Size*

Most projects are too large to permit the evaluation of their effects on all participants. Therefore, a **sample**, or a subset of the population being assessed, is used to represent the entire group. Conclusions regarding project effects on the sample group are then assumed to be valid for the entire population of participants. A good sample, therefore, is one that accurately represents the entire group of project participants plus those eligible but not participating so that conclusions about project impact are valid.[20]

The size of the sample needed for evaluation depends on (a) the number of groups of interest being studied[21], (b) the amount of change expected in the indicator, (c) the level of confidence needed for final conclusions, and (d) the probability of detecting a difference in the indicator when one actually exists.[22] When many indicators will be used, it is general practice to base sample size decisions on those likely to be most critical to the evaluation.

---

20. While this discussion pertains specifically to evaluations, it should be noted that there also are cases in monitoring systems where sampling is appropriate and useful, e.g., in examining the effects of staff training on knowledge and motivation, in assessing changes in knowledge, attitudes, or behaviors where it may be unrealistic to collect information on all participants. Sampling will also be useful in quality checks on monitoring data. In such cases, these sampling guidelines will also be relevant.

21. While it sometimes is tempting to include an array of "cohorts" in an evaluation, (e.g., areas with particular sets of activities in a decentralized project such as Iringa; or project areas with different supervision ratios), one often finds that the resulting sample size of each becomes too small for meaningful analysis.

22. Indicators are discussed at length in section 6 (Selecting Indicators).

As a general principle, the smaller the amount of change necessary for meaningful conclusions, the larger the sample size must be. This is because changes in nutritional status may result from chance as well as from project activities, and a small change is more likely to be due to chance than a large one.

Sample size depends on the budget and desired confidence level for conclusions regarding the project. Ordinarily, there is some minimum sample size, below which the data are not sufficient to draw any conclusions. Similarly, there is a point of diminishing returns above which additional confidence is negligible. The purpose of sample size determination is to find the optimum point between these two extremes.

Simple statistical software packages and statistical formulas are available to calculate estimates of the sample size needed given different expected results and various evaluation designs. These computations usually require fairly accurate estimates of population variance which is seldom known in advance. Where extensive research has been conducted, valid estimates of variability can be calculated and accurate sample sizes calculated.

In computing sample size, it is always wise to add about 10% to the estimate to allow for the fact that some individuals included in the sample may be impossible to find, or may not cooperate, or may otherwise provide unusable data for a variety of reasons.

***Identify a Control Group***

As stated in the previous section, an evaluation that seeks to attribute change to project activities requires the use of a control (or comparison) group. Individuals (or communities) in a control group must be identical, or as similar as possible, to the group of project participants. To establish such a control group, evaluators must identify and control for confounding factors.

**Potential Confounding Factors in Nutrition Projects**

- Income and / or assets
- Land or livestock holdings
- Land type (e.g., suitability of soils for crop production or fish cultivation)
- Geographic location (may affect food prices, types of foods available, and market access)
- Local health care infrastructure (primary health care facilities, providers per unit population)
- Family size and composition (e.g., female-headed households may be at greater risk)
- Individual health status
- Pregnancy status
- Education
- Ethnicity
- Religion
- Season
- Time availability (e.g., households with full time workers may have less time for project participation)
- Areas with more employment opportunities may have more workers.

A *confounding factor* is any factor that is associated both with the result being measured and with the delivery of inputs responsible for the result. If not identically represented in participant and comparison groups, confounding factors can lead to misinterpretation of project effects. Below is

a non-exhaustive list of confounders frequently encountered in field-based nutrition projects.

In an evaluation, the method of selecting participants and controls can control or eliminate confounding. Randomization, if sample size is sufficient, is the most effective method to control for all confounders, known *and unknown*. However, while randomization is feasible in selection of geographic areas or centers, it is less feasible at the individual level for evaluation studies of field-based projects. When random assignment to project and control groups is not feasible, *restriction* (e.g., comparing only low income participants with a comparison group of only low income families or individuals); *matching* (selecting participants and control groups so that potential confounding factors are equally distributed in both groups); and control during *statistical analysis* (requiring data to be collected on potential confounders) can be examined as alternatives.

Selection and utilization of a control population in a nutrition project must be guided by the same ethical considerations that guide the evaluation of any social service project intervention. In a health or nutrition project, a control population normally will be one receiving some level of pre-existing government services. In addition, efforts must be made to assist control group individuals identified in the course of data collection as having a life threatening condition. In many cases, a control population will be brought under project intervention in a subsequent iteration.

***Choose a Sampling Frame***

Once the sample size is determined, evaluators must decide the specific group of individuals (or other unit of analysis) that will be used for the evaluation. In some cases, data will be collected on the entire population. More commonly a subset of the population, will be used. There are many ways to select a sample, but four are the most common: simple random sampling (SRS), stratified sampling (StS), systematic sampling (SS), and cluster sampling (CS). Both SRS and StS require a comprehensive list of

***Table 5.1*** **Guidelines for Choosing a Sampling Plan**

| Method | Advantages | Disadvantages |
|---|---|---|
| Simple Random Sampling (SRS) | • Simple procedures<br>• Simple analysis | • Logistical disadvantages for sample selection, especially for large and dispersed populations<br>• Usually requires a list of the population to be sampled |
| Stratified Sampling (StS) | • Often provides a more representative sample than simple random sampling<br>• Permits adequate representation in the sample of sub-groups of particular interest<br>• Permits a smaller total sample size | • Logistical disadvantages for listing and sample selection<br>• Requires the use of sampling weights in analysis<br>• Requires the evaluators to know in advance the relative sizes of the different strata |
| Systematic Sampling (SS) | • Simple procedures<br>• Enables identification of a representative sample without an exhaustive list | • Logisitical disadvantages if population is dispersed in time or place. |
| Cluster Sampling (CS) | • Logistical advantages for listing and sample selection | • Requires adjustments in analysis<br>• Requires larger sample size |

the population, while SS and CS provide alternatives when a list is difficult to obtain.[23]

***Which sampling method is best?***

The best sampling design for an evaluation will depend on the needs of the study (sample size, variability of indicators, subgroups of interest), on the method of data collection, and on the available resources. Table 5.1 provides some guidelines for choosing a sampling plan.

23. The interested reader can refer to textbooks by Valadez and Bamberger (1994) or Rossi and Freeman (1993) for a more exhaustive discussion of their theory and use. Annex 5 provides tips on how to draw a sample.

**Minimal Information to Gather at Baseline for Participants and Controls**

- Unique identification codes for individuals
- Subgroup identification codes
- Sample cluster identification
- Interviewer identification codes
- Location identification codes
- Date of data collection
- Nutrition status (according to the selected indicator)
- Other outcome indicators (e.g. dietary adequacy, specific caring practices)
- Information on confounding factors

***Collect Pre-Project Information through a Baseline Survey***

After calculating the necessary sample size, selecting a sampling plan, and establishing a control group, the final step in preparing for an evaluation is to collect pre-project information on both the participant and control groups. As stated earlier, baseline information is essential for measuring with confidence the magnitude of change that occurs during the implementation of a project. The box above offers minimum baseline information needs.

# 6 Selecting Indicators



By the end of the activities described in Section 5, consensus has been reached on the goals and objectives of the project, the information which needs to be collected and analyzed through monitoring and evaluation, and the evaluation design to be used. The next step is to find measures that adequately capture each element being assessed. Finding indicators that will be both valid and feasible to measure given logistical and resource constraints is often the most challenging design issue in a monitoring system or evaluation.

Because the choice of key indicators can make a critical difference in the results of an evaluation (see field insight) it may be important for project task managers to be cognizant of discussions being made in this regard.

***What Is an Indicator?***

An indicator is an

- *objectively verifiable measurement* which
- *reflects the activity, assumption, or effect being measured* and
- *allows for comparisons* both between different populations or individuals (e.g. the infant mortality rates of two different countries) and between measures of the same population or individual at different points in time (e.g. changes in one country's infant mortality rate from decade to decade).

By collecting information on selected indicators, the efficacy of the project can be tracked. Furthermore, by comparing the same indicators over time, it is possible to monitor changes and predict effects. Such insights permit the refinement and improvement of projects.

***Characteristics of a Good Indicator***

We all use indicators in our daily lives—we check the temperature before deciding what to wear. We expect the current temperature to be a good

indicator of how cold or hot it will be during the day. Clearly, this indicator is indirect and of variable validity but is still useful because it is accessible and easily interpreted. In addition, it possess the following characteristics:

Good Indicators Are . . .

- simple
- valid
- clearly defined
- reliable
- measurable
- quantifiable
- variable

### *Simple*

Indicators should not be more complex than they need to be. Because the collection, management, and analysis of data is costly both in human and financial terms, indicators should be as ***simple*** as possible without compromising the essence of the variable.

In field settings, direct measures of some variables are often impossible or impractical to gather. In such cases it is necessary to rely on indirect, *proxy indicators.* The ideal way to test for vitamin A deficiency, for example, is to measure the retinol present in blood samples. This biochemical measurement is not feasible, however, on a large scale for the majority of field-based projects. Instead, researchers and project staff typically rely on less sensitive indicators such as clinical signs of vitamin A deficiency (e.g. night blindness or Bitot's spots) or such proxy indicators as dietary intake information. Proxy measures are appropriate if they meet the other criteria of a good indicator, particularly if they are both valid and reliable[24].

24. A risk with such substitute measures is that ease of measurement may tempt evaluators to select a poor indicator. For example, Bitot's spots, although far less complicated and less costly than analysis of blood samples, has few other character-

### *Clearly and precisely defined*

The measures used in M&E must be ***clearly and precisely defined***. It is not sufficient, for instance, to use "percent of underweight children" as an M&E indicator. What does "underweight" mean in this case? Which children are being measured? A better indicator would be the following ratio:

$$\frac{\text{the number of underweight (WAZ} < -2\text{) children aged 6–24 months}}{\text{the total number of children aged 6–24 months who were weighed}}$$

### *Measurable*

Indicators should be **measurable**, whether they are quantitative or qualitative in nature. Height and weight are directly measurable; access to piped water can be measured simply by observation once "access" is defined (e.g., available inside the household; available within 250 yards of the house). Often, a scale or index needs to be created to measure a qualitative variable in quantitative terms. For example, knowledge of correct breast feeding practices might be measured by a respondent's ability to give the correct answers to a set of objective questions.

### *Variable*

To be useful, indicators must show **variation** between subjects and over time. If the indicator does not vary, then even if it is valid, it will not dis-

istics of a good indicator. Bitots spots are rare and, in field settings, difficult to distinguish from other ocular manifestations. As a measure of vitamin A adequacy they lack both specificity (the ability to distinguish current vitamin A deficiency from other deficiencies which affect ocular normality) and sensitivity (the ability to identify all or most of the people who suffer from the condition). Similarly, land ownership, though easier to measure than household income, may be a poor proxy for economic status if the landless have good access to employment opportunities.

criminate between those who have benefited from the program and those who have not. For example, height is a variable indicator for young children, and we can expect well-nourished preschoolers to show more rapid rates of growth in height than malnourished ones. Among adults, height does not vary over time nor with changes in nutritional status; therefore it is not of interest for ongoing tracking of program impact.

Some indicators are useful in one setting but not in another. For example, the materials used in house construction may be a good indicator of economic status in rural areas, where houses may be made of mud, sticks, or cement, but not in urban areas where even the poorest households live in cement structures.

### *Valid*

It is important that an indicator be ***valid***, that it accurately reflect the concept it is supposed to measure. The percentage of admissions in a hospital pediatric ward who are below 60% weight-for-age, for example, would not be a valid indicator of the prevalence of severe malnutrition in the area as a whole.

Similarly, a single day's intake of vitamin A is not a valid indicator of individual vitamin A status, because the vitamin is stored in the body, and day-to-day variation in intake can be substantial. A more valid indicator of overall vitamin A adequacy in the diet might be a food frequency that asked the intake of foods contributing to vitamin A intake over time.

It is important to note an indicator which is valid in one context may be less so in another. This means that it may be inappropriate to transfer indicators from region to region or project to project. In the case of indicators for vitamin A deficiency, for example, the validity of dietary intake depends on the conversion process of beta-carotene in foods to retinol in

the blood, a conversion, in turn, which can be affected by fat and zinc intake and infection levels. Thus dietary beta-carotene may reflect vitamin A status where fat intake is adequate, but not where it is extremely low. Similarly, a tin roof might reflect relatively high economic status in a rural village but low status in an urban area.

### *Reliable*

Indicators must be ***reliable*** so that regardless of who collects the data, the results will be nearly identical. One aspect of reliability has to do with the selection of indicators themselves. The presence of a tin roof, for instance, may be a more reliable indicator (though perhaps a less valid one) of economic status in areas of high underemployment than weekly or even monthly income figures; because income may vary widely from month-to-month or season-to-season while overall economic status tends to be relatively constant.

Note that reliable indicators do change over time if the variable being measured changes. If consumption of vitamin-A-rich foods was found to be low in the winter, and high six months later, this could be an accurate reflection of seasonal change in diet. A food frequency that asked about annual consumption patterns, though, if it were reliable, would not change from one season to the other, as it is supposed to reflect usual intake throughout the year.

Reliability is not the same as accuracy. Accuracy refers to the precision with which a measurement is taken. For example, height should be measured to the nearest centimeter; weight to the nearest 50 grams. Evaluators should be careful to train data collectors carefully, so that their measurements are all accurate to the same standard. This refers both to the accuracy of repeated measures taken by one person (intra-individual reliability), and to the accuracy of the same measurement taken by several different individuals (inter-individual reliability).

### *Quantifiable*

Finally, indicators should be **quantifiable**, and where appropriate, presented as ratios. Actual numbers are often meaningless unless they are converted into some type of proportion. A community nutrition center implementing a growth monitoring project, for example, might report that thirty children failed to grow sufficiently in a particular month. If only sixty children were being served by the center, the proportion of children failing to thrive (50%) would seem quite high. If, however, the center weighed three hundred children each month, that figure would drop to 10%.

But while percentages and ratios can make indicators more useful, it is also important that *actual numbers* be collected, recorded and maintained, so they can then be used in various ways. For example, if individual anthropometric status is collected, it can be aggregated into community, project, or country-wide indicators. Later on, if it is decided to look at the breakdown of anthropometric status by some other variable, such as child's gender, the individual information can be re-aggregated using the new criterion. Information should always be recorded and preserved in the form and at the level of disaggregation at which it was collected.

### *Types of Indicators for Measuring Program Effectiveness*

It is important to have a balanced set of indicators that will measure the combination of *inputs*, *outputs*, *outcomes* and *impacts plus assumptions* that is most appropriate for a given project. This should not be difficult to do if information needs have already been identified based on project objectives, available M&E resources, and potential information collection and analysis constraints.

The following are examples of possible indicators that could be used for three different types of projects. Please note that these indicators are

## *Table 6.1* Illustrative Indicators for Nutrition-related Projects

| Type of Project | Type of Indicator | Indicators |
|---|---|---|
| Vitamin A supplementation (for children aged 6 months to 5 years) | Input | • Quantity of supplements delivered on time to the project site / Total quantity expected to be delivered to the project site |
| | | • Number of quality checks* conducted per year |
| | | • Quantity of supplements found to be unusable / Total quantity of supplements checked |
| | Output | • Number of targeted children who received the supplement / Total number of targeted children |
| | Impact | • Number of children with clinical signs of night blindness* (proxy) / Total number of targeted children |
| Growth Promotion and Nutrition Counseling | Input | • Number of community nutrition workers who received training* / Total number of community nutrition workers slated for training |
| | Output | • Number of women who were counseled* by nutrition educator / Total number of targeted women |
| | | • Number of children aged 6–24 months who were weighed / Number of children 6–24 months in project area |
| | | • Average "knowledge and attitudes" score* of targeted women |
| | Outcome | • Average number of months of exclusive breastfeeding* of targeted women |
| | | • Number of mothers who exclusively breastfed for at least 4 months / Total number of mothers interviewed with children aged ≥ 4 months |
| | | • Number of mothers who introduced complementary foods* to their infants before nine months of age. / Total number of mothers interviewed with children aged > 6 months |
| | Impact | • Number of underweight (WAZ < –2) children aged 6–24 months / Total number of children aged 6-24 months who were weighed |
| School feeding | Input | • Quantity of commodities delivered on time to the project site / Total quantity expected to be delivered to the project site |
| | | • Number of days that school was actually in session / Number of days that school was scheduled to be in session |
| | Output | • Number and % of children who actually received meals each day |
| | | • Number of meals received per child per school year |
| | Outcome | • Number of student absences* (proxy) / Total number of student school days |
| | | • Number of drop-outs* (proxy) / Total number of school children |
| | Impact | • Literacy rate* |
| | | • Primary school completion rate* |

*Terms need to be more clearly defined by M&E staff. For example, literacy would be linked to a test of ability to read and comprehend (i.e., standardized test scores measuring school performance). Primary school completion rates might be percent of children enrolled in first grade six years ago who completed primary school.

illustrative only and are by no means exhaustive. Because these are general examples, no time frames are included. Rather, the assumption is that these indicators are collected on a regular basis (e.g. monthly, quarterly or annually) and compared over time.

***How to Select Indicators***

Using the criteria for a good indicator, those responsible for designing a monitoring and evaluation system may find it helpful to involve stakeholders in selection of appropriate measures for at least outcome and impact indicators. Such efforts at the design stage may well reduce confusion and second guessing later on. The selection of indicators may be critical in subsequent perceptions of whether a project has been successful (see field insight).

A participatory process of selecting indicators should involve stakeholders who are directly involved with project implementation, ideally together with a professional experienced in M&E. One approach is to generate, from scratch, a list of desired indicators from the stakeholders themselves. These may include some indicators that do not necessarily reflect the project objectives. A second is to present to stakeholders a menu of possible indicators listing the advantages and disadvantages of each. A third, less participatory, approach is to have a small team, consisting of outside professionals and key project staff, develop the indicators. Regardless of the technique used, the indicators generated for project M&E should be reviewed to make certain that they conform to the above-mentioned criteria before being incorporated into a data collection system.

This process can offer a range of options that M&E staff alone may overlook. However, it is likely to generate more indicators than is feasible to collect and manage. M&E staff need to select a set of indicators which, when taken as a whole, provide enough information to assess implementation or the effect of the program. This generally requires finding a balance between the *ideal* and the *practical* and collecting only what is needed rather than what is possible or interesting.

## Field Insight: The Cost of Selecting an Inappropriate Indicator

An International NGO operating a major take-home food distribution program through health clinics in Africa in the 1980's made a headquarters decision to begin moving out of this type of program. Looking for an appropriate justification for program termination in one Southern African country, the NGO commissioned an evaluation from a local contractor. Since the program had been justified as a means of improving child growth, the indicator selected by the donor was anthropometric status.

Despite the objections of other agencies working in the country, the evaluator readily agreed to the use of this indicator, and proceeded, in the absence of baseline data or control groups, to compare growth records over time in clinics with and without food distribution. Not surprisingly, the "evaluation" found no significant improvement in child growth resulting from the intervention, and the program was terminated.

Other agencies objected that child growth was an inadequate indicator for such a program and at a minimum, should have been used together with other outcome or impact measures. Critics noted that in a situation characterized by household food insecurity, it is highly unlikely to expect food in a take home feeding program to go solely or even disproportionately to a single child. Additionally they argued that it was inappropriate to use anthropometric measures to assess the effectiveness of a program which addressed the determinants of child growth so inadequately (e.g., beyond the uncertainties of additive calories for the child, there was an absence of nutrition counseling in the program which might have addressed child feeding and home care practices).

By contrast, critics suggested that the program may, in fact, have had important benefits, and that a responsible evaluation should have considered them regardless of whether they were part of the original objectives. The self targeting of recipients, resulting from the long distances between villages and health clinics, and the provision of food supplements to these households may well have led to important reductions in household food insecurity and, in turn, increases in real income and household caloric intake. The program also clearly increased clinic attendance. Finally, the small payments made by recipients for the food covered half of the operating budgets of the health clinics themselves.

Ultimately, the choice of an inappropriate indicator led to an unfairly negative evaluation which terminated a program that was, in fact, making several important contributions to the community. Looking only at child growth overlooked other benefits and, in turn, sealed the fate of the program.

The process of selecting practical indicators also implies that the frequency of collection is manageable. In large-scale nutrition projects, most collection of monitoring data is carried out on a monthly basis or tabulated as monthly averages. Child growth monitoring and pregnancy weight gain data is generally collected on a monthly basis. Coverage of on-site supplementary feeding, although carried out and recorded daily, are usually presented on a monthly basis as the proportion of eligible women or children coming daily or on a majority of days, or the average percentage of eligible recipients appearing on an average day.

Data seeking to measure behavioral change, e.g. the proportion of infants aged 6–9 months who are receiving complementary food, would likely be difficult to collect on a monthly basis, but might be appropriate for quarterly or semi-annual collection and analysis. Data on coverage and massive dose supplementation for children are usually presented semi-annually.

# 7 Deciding Data Collection Strategies



There is a variety of ways in which information can be collected for monitoring and evaluation purposes. Some techniques, such as focus groups and key informant interviews, offer rich, qualitative information about the social and cultural context of an intervention and the problem it is designed to address. Other strategies, including the use of surveys and direct measurements, provide more quantitative data on the implementation and effects of the project. Each method offers certain advantages as well as disadvantages.

***Qualitative vs. Quantitative Data Collection Methods***

There is an ongoing debate among those responsible for monitoring and evaluating development projects as to whether collecting quantitative data is worth the effort, given the time and expense involved, the problems with standardized methods and inaccurate measurement, and the frequently ambiguous results. Skeptics of quantitative data collection often suggest that well-executed qualitative data can provide reliable information at a fraction of the cost.

Qualitative data can indeed provide many insights into project *appropriateness*, that is, the degree to which the project is perceived by beneficiaries and staff to be addressing priority needs in optimal fashion. In addition, qualitative methods can enable the evaluator to understand better what is actually happening in a project, by exploring the *reasons* for particular behaviors and responses. Accordingly, qualitative data, though subjective, is likely to shed important light on processes at work in a project. When responses by a sizable number of project participants are consistent with project outputs and results, such responses may qualify as reasonably valid evaluation measures. Replicable results, based on a representative sample of participants, provide firmer conclusions than do isolated impressions.

At the same time, because qualitative methods are, by their nature, subjective, particular efforts must be made to increase the likelihood of reli-

ability. One way to reduce the subjectivity of qualitative information is to provide rigorous training in the data gathering techniques, to ensure high inter-observer consistency.

Accurate assessment of project impact requires quantitative information; e.g., *numbers* of beneficiaries reached; *proportions* of the target population with improved outcomes; *increments* in height-for-age or birth weight. Note, though, that quantitative information may in some instances, be collected using qualitative instruments. For example, in one project the result might be the increase in the proportion of mothers who believe breast milk is the best food for infants. Beliefs might be measured in a qualitatively-oriented questionnaire which seeks the opinions of mothers through open-ended questions, but the measure (proportion of mothers) is quantitative.

Most large scale projects require some quantitative assessment of change attributable to project activities, while qualitative data provides context and a basis for interpretation.

The purpose of this section is to discuss the most common qualitative and quantitative data collection techniques and to describe when and how each method is generally used for the monitoring and evaluation of field-based nutrition projects. In addition, there is a brief discussion on when to use open-ended and closed-ended questions for data collection and how bias can be minimized through particular data collection strategies.

For the task manager, the choice of data collection strategies will have implications for the nature of M&E information collected and, in turn, the types of personnel needed.

The methods most frequently used to collect data for monitoring and evaluation include:

- observations;
- key informant interviews and focus groups;
- surveys; and
- direct measurements.

### *Observations*

Much useful information can be procured when a trained observer visits a community and collects project-relevant information on community conditions and household/individual activities. Observations are commonly used in project M&E to assess time-use patterns or behaviors that are relevant to achieving desired outcomes and impacts. How much time is spent preparing complementary foods? Are families using iodized salt in their homes? Are women eating more during pregnancy? Observation can also be used for classification purposes. Proxy measures for socioeconomic status, for example, are often derived from observing the type of house in which a family lives, the material used for roofing, or the number of large and small livestock owned.

Observations offer valuable insights into the social and physical context of a problem, but they also introduce opportunities for bias. Most often, observations are conducted by trained specialists outside the project, who are considered to be more objective than project staff. However, outsiders can also bias observations through the selection of non-representative sites, subjects, and seasons. It is helpful if observers can *structure* their observations using checklists of specific behaviors or environmental characteristics, as a way of making observations more objective. Subjects introduce their own source of error if they change their behavior when in the presence of the observer. It is, therefore, often wise to let beneficiaries know that observations will be taking place, but not let them know when they will occur or exactly what will be observed.

### *Key informant interviews and focus groups*

Interviews and focus groups with beneficiaries or local staff are among the fastest and least expensive of data collection tools. If informants and focus group participants are chosen carefully to represent the entire range of stakeholders, and if the interviews are careful to be both open-minded and neutral, these methods can yield clear and candid insights. Key informant interviews and focus groups are useful when one needs to learn more about knowledge, attitudes and practices, of staff, participants or important community members. These techniques are particularly effective for projects that attempt to change behaviors, such as nutrition communication campaigns, and can be used for project planning, material development, and the pre-testing of instruments as well as for M&E.

Key informant interviews involve a face-to-face meeting between a trained interviewer and a person selected to represent a certain group whose knowledge, attitudes or practices are being monitored or evaluated (e.g. project staff, mothers, school children, or mothers-in-law).[25]

Focus groups are small group discussions led by a trained moderator who introduces a topic and facilitates participation by all members of the group. Participants are encouraged to talk among themselves, expressing opinions, relating experience, and providing constructive feedback on the topic, while an observer takes detailed notes on the discussion. The group dynamic allows participants to respond to one another's per-

25. These key informant interviews differ from but may overlap with advocacy efforts designed to gain the support of local political and religious leaders often necessary for project acceptance and success. The same leaders may offer valuable insights about attitudes and constraints inhibiting attitude and behavioral change and, subsequently, about public perceptions of the project.

ceptions, generating new ideas and highlighting conflicting attitudes that would otherwise be inaccessible to an outsider.

Creating the right environment for group interaction requires special attention to the design of the focus groups. Settings, group composition, and the moderator's skill at guiding conversation all influence the participants' willingness to divulge conflicting opinions and discuss sensitive topics with candor. Focus groups generally should be comprised of similar individuals, e.g., mothers of small children; fathers; or parents of teenagers. Focus groups cannot be conducted in the presence of staff or administrators of the program being evaluated as this often will bias or limit the responses.

Either of these techniques can be used to generate information on such topics as the level of motivation, workplace satisfaction, and knowledge among community nutrition workers; mothers' constraints to improving child care practices; cultural beliefs about diet and food preparation; and the attitudes of community leaders regarding the usefulness and appropriateness of the project.

Though key informant interviews and focus groups can provide important contextual information, certain difficulties should be anticipated. These difficulties relate in part to open-ended questions which are used for these purposes. Such questions can provide important insights on perceptions, sometimes reveal project-related issues not originally considered, and may provide the basis for subsequent close-ended questions. At the same time, open-ended questions themselves are often difficult to code and analyze. Because there are no standard response categories, in such open-ended questions, it is usually not possible to compare the information statistically within or between projects (although qualitative comparisons can be made, for example regarding perceived effectiveness and likely sustainability; and one project's experience may shed light on another's). Interviews and focus groups require experienced staff

capable of probing, with follow-up questions, for answers and eliciting and recording adequately detailed information.

***Surveys***

Structured surveys in nutrition projects, usually use closed-ended questionnaires to capture information on inputs, outputs, outcomes and impacts. Surveys are useful because (a) information can be generated on a wide range of factors (e.g., service delivery, attitudes, knowledge, and practices), (b) closed-ended questions can be easily analyzed and compared, (c) findings from surveys can be generalized to the population of interest if the sample is representative, and (d) large numbers of people can be interviewed.

As a general rule, questionnaires need to be administered by thoroughly trained interviewers. While surveys conducted in industrialized countries can be delivered through the mail and self-administered or conducted by phone, these are generally not feasible options for field-based nutrition projects in developing countries.

Some disadvantages of surveys are that (1) they can result in oversimplified information, particularly when closed-ended questions are used; (2) interview bias is often introduced; and (3) instrument development and sampling are complicated procedures.

The first of these disadvantages can be largely overcome by careful preparation of the questionnaire. Closed ended questions must always be based on sound and complete background information about what the response categories should be. Many closed-ended questionnaires include a few open ended questions, or include the response category “other”, to permit respondents to give additional or unanticipated information. The second disadvantage, interviewer bias is not unique to formal surveys; any technique in which the data collector interacts with respondents is subject to some degree of bias, which can be overcome

## Open or Closed: When to Use Which Type of Question

There are two basic kinds of questions that evaluators use in interviews, focus groups and surveys: *open-ended* and *closed-ended*. Respondents answer open-ended questions in their own words, and the interviewer is expected to record the answer just as it is given. Closed questions, on the other hand, require specific replies, or include a set of possible answers from which the respondent is asked to choose. Because each type of question has certain advantages and disadvantages, M&E staff must find an appropriate balance based on the purposes of the information gathering.

| | **Open-Ended Questions** | **Closed-Ended Questions** |
|---|---|---|
| Advantages | • promote spontaneity.<br>• encourage people to answer using their own words and concepts.<br>• help generate additional information and insightful details that would be missed by asking closed questions.<br>• elicit more accurate information about sensitive or taboo behavior. | • allow for comparisons within and between programs.<br>• are easy to ask and record and do not require highly skilled interviewers.<br>• can help people recall information.<br>• are often easier for people to answer.<br>• have simpler, less error-prone coding, particularly if the format is well designed. |
| Disadvantages | • are difficult and more costly to code therefore, it is impractical to have a large sample.<br>• are prone to bias because answers can be interpreted in different ways.<br>• take more time for respondents to understand, process, and answer.<br>• require highly skilled and disciplined interviewers who will take word-for-word notes and probe for details. | • hinder spontaneity by forcing people to choose among offered response categories instead of answering in their own words.<br>• do not give people enough time to reflect and remember.<br>• may result in oversimplification.<br>• may limit the range of responses. |

As a rule of thumb, one should rely primarily on closed questions with *well written, comprehensive** response options and use open-ended questions primarily when:

- conducting focus group sessions and key informant interviews;
- there isn't enough information to write appropriate response categories;
- trying to answer questions of "why" and "how";
- promoting awareness of and participation in the program is one objective of the data collection effort.

---

*Response categories should be exhaustive, that is, cover all possible answers. To accommodate unforeseen responses, an open-ended other category may be used. If only one response is permitted, categories should be mutually exclusive, that is, categories should not overlap. In many cases, though, respondents may be allowed to give more than one response. For example, "what difficulties do you have in attending the clinics ?" might permit both "distance" and "hours of operation" as responses.

to some extent (though never perfectly) by careful training. The third disadvantage, complex procedures, can be decreased through utilization of guidelines such as those included in Section 8 and, in the case of sampling, assistance from specialists outside the project.

***Direct measurement***

Direct measurements are usually necessary to ascertain changes in nutrition status resulting from a project. The specific methods employed most often for nutrition projects include the use of anthropometry through height-for-age, weight-for-age, weight-for-height, and body mass index measurements; analysis of blood, urine, and breast milk; the assessment of clinical signs of micronutrient deficiencies such as goiter and night blindness. Other direct measurements include weighing of food served to household members (often before and after meals, to capture the amount left over), but this method of measuring food consumption is extremely time consuming and intrusive, and therefore tends to alter the behavior being measured. Food consumption for most purposes is better measured by recall methods.

While many direct measurements are not technically complex, all require precision and practice. They also require familiarity with the local context. In Astonia for example, errors in head circumference measurements were traced to the practice of including braids rather than measuring under them. Errors also occur when measuring tools are not properly calibrated and when data collectors are not well trained to use them. Data collectors should be trained to a specified standard of accuracy (so that, for example, repeated measures are accurate within a specific range). If the data collection period is lengthy (several months) periodic retraining is wise.

Challenging problems of direct measurement are likely to arise in collecting particular types of evaluation data in a control area. One such case is the collection of control group pregnancy weight gain data. Another is control group birth weight data. Neither lends itself to the normal proce-

dure in a programmatic evaluation of collecting specified control group data at single points in time. Yet both are likely to be important in projects seeking to improve pregnancy outcomes and wishing to ascribe improvement to project interventions.

In these cases, there may be no alternative to locating evaluation team personnel in a control area for the lengthier periods of time necessary to collect monthly pregnancy weights (over, say, a three month period) on a subset of pregnant women at varying stages of pregnancy. During that same time period, the evaluation team, necessarily relying heavily on the cooperation of local midwives or clinic personnel, would collect birth weight data on all births taking place until sample size requirements were met.

### *Maximizing the Coordination of Data*

Information needs for project M&E are often extensive. As a result, it is advisable to maximize the efficiency of data collection by:

- maintaining and compiling information from project records, and
- using already existing information (i.e. secondary data sources) when appropriate.

#### *Maintaining and compiling information from project records*

Monitoring and evaluation should, to the extent possible, use information that is already being collected for program purposes. In addition to management-related information, such as data on actual costs and deliveries and quality checks on inputs, many projects collect considerable pre-project data on communities and individuals. Anthropometric data, information on the incidence and severity of micronutrient deficiencies, and project participation rates are all examples of information that often is collected routinely for nutrition projects and often form part of the ongoing monitoring system. Such information, however, should be checked periodically for quality.

### *Secondary data*

What can be done in a mid-project evaluation in cases where no baseline information was collected and/or no control group was established? As indicated in Section 4, one answer is to seek secondary sources of data. Secondary data consist of statistics and other information that were originally collected for purposes other than the project. The most useful types of secondary data for nutrition projects include national or regional surveys of nutritional status, dietary intake surveys, micronutrient deficiency prevalence surveys, and consumer expenditure and consumption surveys. Such information will usually come from government offices, donor agencies, NGOs, or research institutions.

While using secondary data can be a fast, inexpensive, and convenient way to obtain information, it requires a detailed inspection of the original collection process, keeping in mind that the validity and reliability of the present findings will rest upon the quality of another's collection methods.

### *Controlling for Bias through Data Collection Methods*

There are various types of bias, or systematic error, which can affect the results of an evaluation and lead to erroneous conclusions. While some types of bias can be controlled through sampling or data analysis techniques, much can be done to minimize such errors by using proper data collection methods.

While it usually is impossible to eliminate bias altogether, bias can be reduced when information is collected in exactly the same way for participants and controls, both before and after project activity. To do so, data collectors must assure that the *administration* of the data collection instrument is well thought out and implemented. The following table summarizes specific measures that can be taken to decrease the likelihood of bias.

## *Table 7.1* Minimizing Bias through Data Collection Methods

| Type of Bias | Description of Bias | How to Minimize Bias through Data Collection Procedures |
|---|---|---|
| Interviewer bias | Baseline and post-project data or information about participants and controls is collected by interviewers in different ways. Observed differences are thus due to differential measurement and recording and not to project effects. | • Train interviewers thoroughly.<br>• Standardize the interview protocol.<br>• Use highly objective, closed-ended questions.<br>• If there is more than one data collector (or team of data collectors), each collector or team should gather information from different areas, both in the baseline and in the subsequent evaluation surveys, to minimize the effects of individual differences in skill or method. If possible, the same data collectors should conduct both pre- and post -surveys.*<br>• For longitudinal surveys, the same data collector (or team) should collect information for the same individuals throughout the duration of the evaluation. if this is not possible, carefully train replacements to continue data collection is an identical manner. |
| Instrument or measurement bias | Measurement errors are due to instruments that are not identical for participants and controls; thus differences are due to differing measurement and not to project effects. | • Standardize measurement instruments and procedures.<br>• Calibrate instruments (such as weighing scales) frequently. |
| Recall bias | The participants and controls remember and report information in different ways due to their different exposures either to the intervention or to a particular results. | • Train interviewers thoroughly on how to probe for information and how to help respondents remember past events.<br>• Use specific and meaningful reference / recall period. |
| Time or seasonal bias | Data on participants and controls are collected at different times of the day or different seasons of the year; thus differences are due to time or seasonal effects rather than the project activities. | • Standardize the time of day or season of data collection so information on treatment groups and controls is collected during the same time period. |

*This has been particularly problematic in several large World Bank-assisted projects in which separate contracts to different contractors were awarded for baseline data collection (in one case two separate contracts for successive baseline surveys as new areas were brought under project coverage) and for subsequent midterm and end-point evaluations. Beyond problems of inconsistent data collection, evaluators find themselves in a near impossible situation of seeking to develop evaluation instruments which will be consistent and comparable with differently designed baseline and midterm survey instruments. A single contract for ongoing external evaluation, discussed in section 1, is one way to avoid such problems.

# 8 Developing Data Collection Instruments



Monitoring and evaluation rely on data collection instruments to elicit and record information. There are various types of instruments available for M&E including *summary report forms* used to aggregate project data, *questionnaires*, more informal *interview guides*, *inventory forms*, and *observation checklists*. Several formats can be combined on one form. For example, a household survey may include questions for the mother about her family's health as well as a checklist for observations about the size, construction, and cleanliness of the house.

The development of clear, easy to read instruments is a vital component of both monitoring and evaluation. Knowing what to include and how to organize the questions and answers requires skill, common sense, and practice. Considering how the collected data will be analyzed and used before designing the data collection instrument, creates a focused, efficient instrument. The following section outlines considerations to be taken when designing data collection instruments. (Clear language and an easy to follow format are essential to collect accurate, useful information).

### *Steps in Designing M&E Instruments*[26]

In designing a questionnaire, report form, or other type of instrument, M&E staff should follow the following concrete steps:

1) Draft the *content* of the instrument based on pre-determined information needs.
2) Pay attention to *language*.
3) Craft *questions* carefully.
4) Optimize the *sequence* of the questions.

26. These guidelines have been formatted into an easy-to-use checklist that can be found in Annex 6.

5) Make the *format* easy to follow.

6) Check for *consistency* between instruments.

7) *Pre-test* the instrument.

8) *Revise* the instrument.

Once these tasks are accomplished, the instruments can then be printed, distributed, and put to use in the field.

### *Draft the Content*

***Keep it short and simple.*** Information should be collected and recorded in the simplest way possible while maintaining reliability. An M&E system should collect only what is needed. Those designing the data collection system should determine in advance how *each piece* of information collected will be used. It is useful to keep in mind that the collection, processing, and analysis of information is costly in terms of financial and human resources.

Some project evaluation teams have found it particularly useful to construct, at the point of survey questionnaire design, dummy tables corresponding to the filled in tables anticipated in the evaluation report. The very process of constructing these tables often clarifies the type and form of data required and can significantly reduce the proportion of information collected but unutilized—often over half of the data in an evaluation survey.

***Identify subjects and sources.*** The instrument must provide enumerators with an easy way to identify the child, household, clinic, or community being studied. Codes are often used to identify the physical location of the subject (e.g. a province, state, district, county, village, cluster, house) and the particular unit of analysis (e.g. a clinic, a household, an individual). In addition, the persons filling out the forms must be identified on the form.

### Field Insight: Identification Codes that Have Worked

A household survey conducted in Nicaragua that focused on under-twos, combined various codes to form one long identification number for each child.

| ___ | ___ | ___ | ___ |
|---|---|---|---|
| Province | Town | Household | Child |

Several monthly reports for the Bangladesh Integrated Nutrition Program have separate lines to identify the *thanas,* unions, and community nutrition centers (CNC). The numbers are then combined into a single code when entered into a computer.

Thana: _____ Union: _____ CNC: _____

***Use filter questions*** to avoid asking inappropriate questions. For example, ask a woman if she has children before asking the age of her youngest child.

***Calculate later.*** Forms to be filled out in the field should not require calculations. Data collectors should record information as provided to them. Any manipulations (e.g. ranks, percentages, etc.) should be done later. In many cases, easy-to-use computer programs, such as EpiInfo[27] or

27. EpiInfo is a word-processing, database and statistical software package provided free by the Centers for Disease Control and Prevention. Dean AG, Dean JA, Coulombier D, Brendel KA, Smith DC, Burton AH, Dicker RC, Sullivan K, Fagan RF, Arner TG. EpiInfo, Version 6: A Word-Processing, Database, and Statistics program for Public Health on IBM-compatible computers. Centers for Disease Control and Prevention, Atlanta, GA, USA. 1995.

**Field Insight: Defining Terms to Avoid Misinterpretation**

Sometimes what is clear to those designing instruments may be incomprehensible or misinterpreted by others. This is particularly true when translating instruments into foreign languages. Even in the same language, however, there may be different understandings of concepts. In Ghana, for example, the term "short-term hunger", which was introduced by a workshop facilitator to discuss the consequences of skipping one or two meals, was interpreted by workshop participants as the seasonal hunger that local people experience six months of the year.

basic spreadsheet or statistical packages, can be used to make these calculations.

***Pay Attention to Language***

***Communicate appropriately.*** When designing monitoring and evaluation instruments, it is important to use the local *spoken* language and in ways which capture local concepts. Sometimes pictures or symbols that are easily recognized by respondents are useful to convey ideas, particularly where level of literacy is low. When discussing past events or seeking to estimate the age of a person, it may be helpful to use local holidays, seasons (e.g., the harvest or the rainy season), and special occurrences (e.g., a flood or construction of a school) to help respondents with recall.

***Do not compromise on translations***. When working in foreign languages, a good translation is essential. If the original instrument is to be written in English, it is useful to translate the English version into the local language and then have someone else translate it back into English. Not only will this prevent translation mistakes, but it will also allow investiga-

### *Table 8.1* General vs. Specific Language for Instruments

| General Language | vs. | Specific Language |
|---|---|---|
| Does the school have adequate sanitation?<br>_____ yes _____ no | | Number of functioning latrines _____<br>Number of people who use latrine (teachers, students, and others) _____<br>[Note: these figures can be used to later decide if there are adequate sanitation facilities given the ratio of users to latrines.] |
| Is the house clean?<br>_____ yes _____ no | | Inside the house, there is evidence of:<br>Animal feces _____<br>Rotten food _____<br>Garbage on the floor _____ |
| How often does your child have diarrhea? _____ | | How many loose or watery stools has your child had in the past five days? _____ |

tors to understand which concepts do not translate well in the local language and vice-versa. In Bangladesh, for example, an ORS campaign recommending use in all cases of diarrhea, mistakenly used the Bangla word for watery stools, understood locally as a small subset of diarrheal infection.

***Use simple language***. Avoid using double negatives such as, “Do you believe that a mother should not have to attend feeding sessions if her child is not ‘at risk’?”

***Be specific***. Language used, whether in interviews or on reporting forms, must be specific to prevent confusion. Instruments should avoid general terms such as “small”, “big”, “frequently”, or “often”. The examples provided in Table 8.1 may be useful in judging whether the language in an instrument is sufficiently specific or too general.

### *Craft Questions Carefully*

***Carefully craft questions.*** Inappropriate wording of questions can influence responses and may lead to information which do not accurately reflect the situation. Employing the following strategies when crafting questions can prevent this phenomenon.

***Avoid "double barreled" questions*** that present two ideas simultaneously such as, "Do you think children should receive deworming medicine and micronutrient supplements?. Make sure that each question covers a single idea.

***Keep questions neutral*** and refrain from using words that could influence answers. For example, the question, "Do you think it is healthy to take a daily vitamin supplement" is a slanted question, because it presents the idea of "healthy" without the opposite concept. A better question would be "Do you think taking a daily vitamin supplement is good for your health, bad for your health, or makes no difference?" While this may sound cumbersome, it will assure that the use of a positive word in the question doesn't bias the answer. Another way of asking the question would be "What do you think about taking a daily vitamin supplement?" and recording the answer in pre-coded categories.

***Do not reference authorities.*** Questions can also be slanted by using references to authorities: "Do you agree or disagree with the nurse's advice to breast-feed exclusively for four months?" appears balanced, but reference to the nurse suggests that one answer is better than the other. Respondents may give an answer which reflects what they think is expected rather than what they truly believe.

***Avoid hypothetical questions***. All questions should refer to actual events and circumstances. Hypothetical questions frequently are difficult to conceptualize, and often generate confusion.

***Ask for facts, not judgments***. Ask for facts or descriptions rather than judgments on the adequacy of the service. For example, ask, “How many meals did each household member consume each day during the past week?” rather than “Is your family short of food?” The definition of “short of food” may vary between participants and therefore, be difficult to interpret during analysis of the data. Subjective questions can be useful, as long as it is recognized that they measure opinions and attitudes, not facts.

***For closed-ended questions***, which have response options listed, the *options given must* cover each of the major possible answers. (It is unrealistic to seek to be fully exhaustive. In those cases where it is anticipated that responses may fall outside the options listed, an “other” response category may be included. However, inclusion of “other “ should rest on the assumption that it will be minimally used). When the question refers to specific facts, the response categories should be mutually exclusive. Examples include income and education levels, where a person can only belong to one category. Response categories should not overlap. For example, categories should be “some elementary school, but did not finish”, and “completed elementary school.” If the first category were “some elementary school” a person could misread the question and check both the first and second responses. In such cases, it should be clear that the boundaries of each category are not identical, e.g., “less than 6th grade”; “6th grade or more, but less than 12th grade,” “completed high school; no further education” and so on. It is easy for the respondent to be confused if the categories are “0–6th grade” “6th to 12th grade” “12th grade and above.”

In questions concerning beliefs, attitudes, or behavior, you can choose to make the ***response categories*** mutually exclusive, but in many cases, multiple answers should be permitted. For example, a question concerning use of the clinic might include “hours are inconvenient” and “medicines are not available” as reasons for failing to attend. It is quite

possible that both factors influenced the respondent's decision, so both should be permitted answers to the same question.

### *Optimize the Sequence*

***Begin with a non-threatening question.*** Always begin an instrument with easy, non-threatening questions which will build rapport with the respondent. Questions should be grouped by topic and generally within each topic by response type (e.g., yes/no, agree/disagree, and multiple choice). Place open-ended questions and questions about sensitive topics, such as income, near the end. Some surveys end with an easy to answer question; in an effort to leave participants feeling positive about responding.

### *Make an Easy to Follow Format*

***Pay attention to layout.*** The *format* of a data collection instrument, whether for a one-time survey or for ongoing monitoring, is often as important as the content itself. A poorly laid out instrument will result in errors, and information will be unreliable. It is important to keep the instrument uncluttered and with plenty of white space between questions. Questions should be clearly numbered and ample room provided for answers. Boxes or lines should be used to contain responses. Each form should be clearly labeled at the top with both the name of the form and the level of data collection (e.g., health clinic, district office, etc.) with a space to indicate the date of data collection. If the instrument is longer than one page, the pages should be stapled together with identification numbers on all pages.

In addition, leave some space at the end of forms for additional comments. Interviewers should be trained to recognize and explain external factors likely to influence the project (e.g., drought, conflict, or temporary interruption of supply delivery) and unintended effects and note them accordingly.

***Pre-code if possible.*** For questions that are asked with response options, it is often desirable to include codes on the instruments themselves as long as the format remains uncluttered. The codes should be placed next to the answers, as the following example demonstrates, so that computer entry can be done easily and quickly.

| |
|---|
| 7 Type of roof: 1. _____tin 2. _____thatch / grass 3. _____other |

When the same prewritten response options are provided in more than one question, the same code should be assigned to a particular response throughout. For example, no will always be coded as 1, yes as 2, don't know as 3 or for question 7 above, tin = 1, thatch/grass = 2. Wherever tin is found in the questionnaire it should always be coded as 1. However, the same codes can't always have the same meaning, 1 = no in some questions and 1 = tin in question 7.

If it is not possible or appropriate to pre-code the entire instrument, leave spaces for enumerators or coders to write in codes as well as make any necessary observations, calculations or verifications after the data have been collected. This can be done by providing and extra column on the right side of the form.

For example, in food consumption questionnaires, it is impossible to record the codes for all possible foods that may be reported. The form should have space for the interviewer to write the name of the food, and a space next to it so the numeric code for the food can be put down later. The codes, though, should be determined in advance and available on a list.

### *Check for Consistency*

***Match up the forms.*** After designing the instruments, including the computer data entry program, the entire set of instruments should be re-

viewed together to make sure there is consistency among them in terms of format and coding. In addition, comparison of instruments allows the investigators to determine if all needed information has been collected and if identical information is obtained from more than one source. If critical information has been forgotten, having all forms in view makes it possible to determine the most efficient way to collect data.

***Compare evaluation questionnaires across time.*** In order to compare baseline data to data collected at another point in time, often at the midterm or endpoint of a project, questionnaires need to be as similar as possible. While information needs may change slightly, adding a question or series of questions to gather the additional information will not affect the validity of the questionnaire. However, the wording of a question or the order of questions may influence the answers received. Therefore, it is important to maintain consistency between baseline, midterm and endpoint questionnaires. It is also important to note that if a question is found to provide incorrect or non-informative data on the baseline, it is likely to give the same results if presented in the same way and should be changed on subsequent questionnaires.

### *Pre-Test*

***Pre-test all instruments.*** Part of the process of designing data collection instruments is to pretest them. Pre-testing allows the data collectors to determine if there is sufficient variation in the responses; the language and concepts used are clearly understood; the response categories are appropriate; participants are able to answer the questions easily; the format is easy to follow; the flow of information is logical; and the instrument can be completed in a reasonable amount of time.

Even when questions have been used before, they must be pre-tested if they are going to be used in a new context or area. Pretests should be administered to people who have backgrounds and experience similar to the intended respondents. Standard report forms used for monitoring

**Field Insight: The Value of Pre-testing**

A doctoral student conducted household surveys in remote areas of Mali to examine food security. Unfortunately, he didn't take the time to pretest his instrument. During the first few interviews, he found that people didn't understand the supposedly "simple" question, "How many people live here?" Did he mean how many people slept here at night or those who gathered here for meals? Should they include family members who had temporarily migrated in search of dry-season labor? Because he had not taken the time to pretest, he had to redesign and reprint the instrument.

In another survey, the designer of an agricultural questionnaire asked farmers, "Do you receive from farming less than ¼ of your income, ¼ to ½, more than ½, or almost all of your income?" Field pre-testing showed that farmers could easily report absolute amounts produced and money earned, but that they did not understand the concept of fractions. The questionnaire was redesigned to collect the information in absolute terms.

should also be pre-tested. The people responsible for regularly filling out these forms (e.g., clinic staff, trainers, and project managers) can offer invaluable feedback on how easy forms are to complete and may suggest ways to improve the format.

### *Revise*

***Revise all instruments.*** Making appropriate changes to the original data collection instrument based on the information revealed during pre-testing leads to a more effective questionnaire. Pre-testing and subsequent revision invariably results in an instrument that produces fewer mistakes in answering the questionnaire or form.

# 9 Analyzing the Data



This section discusses the analysis of collected data. It pertains primarily but not exclusively to analysis of evaluation data. Technical issues of less immediate interest to project managers are briefly summarized. More attention is given to cost-effectiveness and cost-delivery calculations ideally at the core of evaluation results. A generalized sequence of steps for evaluation analysis is presented in Box 9.1.

### *Analyzing Quantitative Information*

Once the data have been cleaned and rounded or collapsed where appropriate (see Box 9.2), two types of analysis follow: analysis of the key project effects through comparisons planned before the project began, and analysis suggested by the results themselves. Key principles are provided below.

#### *Significance levels and trends*

Analyses look for statistically significant results, evidence that the observed effect did not occur due to chance alone. The level of statistical significance is usually set at 0.05, that is, an effect is presumed to be "real," if it has less than 5% chance of having occurred due to random variation. In small samples, sizable effects can be statistically insignificant. In an evaluation, it makes sense to look at results that are statistically insignificant, especially if several effects seem to indicate the same thing.

It could be that these results would be significant if better data collection methods had been used or if more time had been permitted for project activity to show results. If the latter is perceived as a distinct possibility, this supposition might itself be taken into account in considering the future of the project, including future evaluations.

#### *Pre-planned comparisons*

*Analysis of "before and after" data*

To analyze an evaluation based on "before and after" data collection, the key analysis issue is a difference in nutrition status between participant

**Box 9.1 A Sequence of Steps in Evaluation Analysis**

The following steps represent one logical sequence for evaluation analysis in a situation where a well functioning monitoring system also exists:

1. Review project monitoring data to assess the extent to which the project was implemented according to plan, note locations where it was not, and, in the case of shortcomings in outputs, move backwards (to the left on the conceptual framework) in an effort to understand the reason for the shortcoming
2. Clean and review the evaluation data
3. Compare "pre" and "post" data on impacts for participants and controls
4. Whether or not the results of "3" above indicate ascribable project impact, do the same for outcomes in projects with behavioral change components
5. Move backward (to the left) through the conceptual framework in an effort to understand impact and outcome results which have fallen short of expectations
6. Investigate unplanned results
7. Incorporate project cost data to permit cost-effectiveness analysis and, as relevant, analysis of costs per output, outcome or assumption (e.g., change in knowledge)
8. Consider sustainability issues

During steps 3 through 5, the relevant qualitative data which has been collected should be reviewed to provide context and facilitate interpretation.

and control groups *after project services have been functioning for some period of time.* To determine this "after project" difference, it is necessary either (a) to establish retrospectively that nutrition status among participants and controls was not different before the project and then to examine after-project differences, or (b) to establish baseline mean values for participants and controls and then to compare after-project *change* among participants to after-project change among controls. Of these two options, the second is preferable because it is more precise.[28]

*Multiple comparisons*

The level of a statistical test defines the probability that a decision about project effects is wrong. That is, if the confidence level of a statistical test is 0.05, then the chance of being wrong is 5% when concluding, based on the data, that the project had an effect. As investigation leads to more and more conclusions about relationships found in the data, it becomes possible to draw many conclusions as part of a single evaluation. While the probability of error for each individual decision is set by the level of the statistical test, the probability of *any* incorrect decision being made as part of the overall evaluation increases as more decisions are made.

*Increasing complexity*

As complexity of analysis mounts, the possibility of making invalid comparisons and drawing inaccurate conclusions increases. Consider, for example, an evaluation of a program providing nutrition services to women and adolescent girls. If the statistical results show no effect of project participation on nutritional status, no effect of age on nutritional status, and no interaction between the two, there may still be a tempta-

28. More generally, these pre-planned quantitative comparisons should, where possible, be coupled with qualitative data, e.g. interviews with service providers and/or beneficiaries on their perceptions of changes attributable to the project.

## Box 9.2 Cleaning the Data

Cleaning the data means identifying and correcting errors in data. This can be done by checking for valid values for each individual variable; checking for valid values for specific combinations of variables; and screening for outliers.*

*Checking valid values for each variable.* One way frequency distributions of data items will identify those which fall outside of acceptable ranges. For example, adult intakes below 100 calories per day or infant weights over 10 kg.

*Checking valid values of one variable in comparison with another*. For example, Seventy year old women are unlikely to be pregnant; cement roofs are unlikely to be supported by bamboo foundations.

*Screening for outliers.* Graphs of correlations between two factors provide another way to check for values that may be plausible but seem uncharacteristic or questionable for the study population. These values are called "outliers" because they lie outside the general grouping of data points seen on a graph. They require decisions concerning their inclusion for analysis.

When outliers exist, three possibilities for dealing with them are:

1. The outlier may referent data entry or recording errors. Reviewing the original data forms or checking directly with the participant may identify the source of error.
2. Where the graphed data point combines two pieces of information, e.g., height and age, only one of the two may be incorrect. A check on other information collected from that individual or household may identify the one which is correct and only the incorrect data point need be excluded from further analyses.
3. Both values may be incorrect, or enough uncertainty exists to justify discarding both data points.

*It should also be noted that some data are categorical, i.e. they have discrete and often a limited number of values while other data are continuous, i.e., they take on a wide range of values, e.g., height, weight, BMI. For analytical purposes it may not be useful to create frequency distributions of continuous variables unless they are rounded or collapsed into categories.

tion to search for effects of project participation on nutritional status at different ages. It would be tempting to test for project effects among a particular age group—say, 15 to 18 year-olds—and such a test might appear to show a positive project effect. In this case, however, the difference would be spurious, or due simply to chance, since the analysis already had shown that age and project participation have no interaction effect on nutritional status. The axiom “keep it simple” helps to prevent these kinds of errors in interpretation.

***Investigating unplanned results***

Projects frequently have unplanned results that may or may not be picked up in the statistical analysis. Measuring effects outside the scope of the original project may reveal benefits which were not considered in the project objectives.

For example, a food supplement intended to improve weight gain in pregnant women might have been diverted to the woman’s other children; while her own weight gain may not have been affected, the diet adequacy of the household might have been. Alternatively, the food supplement may not have resulted in any measurable improvements in nutritional status, but may have increased the frequency of prenatal visits and therefore reduced the rate of complications in childbirth. Analysis may also reveal unintended negative effects. One project in South Asia, for example, investigated the effect of caloric supplements on the growth of adolescent girls. While the caloric supplements improved growth in adolescent girls, it was also associated with earlier onset of menarche, likely to result in first pregnancy at a younger age.

*Post-hoc analysis*

As evaluators work with the data, they may observe interesting relationships that were not anticipated. For example, it may be found that women in households with another adult woman present are more likely to attend the clinic than women in nuclear family households. Or an unexpected

### Field Insight: Positive Results Shadow Underlying Detrimental Effects

Consider, for example, a project designed to increase food security through commercial fish cultivation. An evaluation based solely on the project design might show successful food security results in terms of increased numbers of households cultivating fish, and increased incomes generated among these households from selling their yields. However, an unanticipated effect of project activity might be that small fish normally consumed by these households are killed when preparing family ponds for commercial fish cultivation, thus removing these small fish from the diet with a resulting decline in vitamin A and calcium intake (effects that would show up if micronutrient status were included as an impact measurement). This example also demonstrates the importance of collecting qualitative information. Had the evaluators not spoken directly with the participants, and in the absence of data on micronutrient status, the improvements in economic status would have been the sole consideration in subsequent decision-making

seasonal pattern of clinic attendance may emerge that is worth exploring. However, evaluators must be cautious not to over-interpret every two-variable correlation, nor to explore every possible relationship to see if any are significant, as these “fishing” expeditions can result in misleading conclusions; they may be a random occurrence unique to this data set rather than an actual association between two factors. Relationships to be explored should be chosen based on a plausible underlying rationale for believing that they might be important to the project being evaluated. This rationale may stem from the findings of other similarly conducted projects or feasible biochemical mechanisms.

### *Cost-Effectiveness, Cost-Delivery, and Sensitivity Analysis*[29]

*Cost-Effectiveness Analysis*

A useful format for presenting quantitative analysis results from project evaluation relates impact or project effectiveness to the cost of producing that impact. Effectiveness data in isolation may not be meaningful if the cost of producing the impact is prohibitive (as in early examples of nutrition rehabilitation centers). Similarly, cost data alone, often presented in cost per beneficiary per year, says nothing about impact achieved for such an expenditure.

Cost-effectiveness calculations relate the total cost of a project (see Section 3) to particular impact indicators, and are often presented in terms of an individual participant. Pre-post comparisons of participants and controls are used, for example, to produce total *numbers* of infants with birth weights above 2.5 kgs, as a result of the project, or, similarly, children elevated through growth from third degree malnutrition (and who would not have been elevated—as seen from control group comparisons—without the project intervention). Dividing total project cost by these numbers provides cost-effectiveness estimates with considerable meaning, e.g.:

Cost per child removed from third degree malnutrition (through improved nutritional status);

Cost per case of child stunting averted;

Cost per child death averted;

Cost per 0.1 kg increase in birth weight.

29. Also see the World Bank Toolkit #3 (Phillips and Sanghvi 1996).

These cost-effectiveness figures can be compared with those collected in projects with similar objectives and in areas with comparable cost structures.

The generalized formula for a cost-effectiveness ratio ( R ) is R = C/U, where C = the total cost, and U = the number of units of change resulting from the project. The following example, using calculations made by T.J. Ho (1985) from first year data of TINP, illustrates the use of this tool.

1. Total cost—U.S. $28 million

2. Reduction in Numbers of Malnourished Children:
   Reduction in number of children moderately and severely malnourished (II–IV degree) due to project = 554,000. (This number was calculated by adding the reduction of those in the second degree category, 325,000, and the reduction of those in the third and fourth degree categories, 229,000).

3. Annual cost per child removed from moderately and severely malnourished categories:
   $28 million (given in step 1) divided by 554,000 children (calculated in step 2) = $51 per child per year

***Table 9.1*** **Reduction in Numbers of Malnourished Children**

| | Initial Distribution | | Final Distribution | | | | Difference Due to Project |
|---|---|---|---|---|---|---|---|
| | Project and Control | | Project | | Control | | |
| | % | No. (,000) | % | No.(,000) | % | No.(,000) | No. (,000) |
| Normal | 21.1 | 1,025 | 18.5 | 898 | 15.5 | 753 | +145 |
| I | 31.8 | 1,544 | 40.2 | 1,952 | 31.8 | 1,544 | +408 |
| II | 31.0 | 1,505 | 29.1 | 1,413 | 35.8 | 1,738 | –325 |
| III & IV | 16.1 | 782 | 12.2 | 592 | 16.9 | 821 | –229 |

## *Table 9.2* Comparison of Three Vitamin A Distribution Programs

| | Type of Program | | |
|---|---|---|---|
| | **Fortification** | **Capsules** | **Gardening** |
| Beneficiaries* | 5,500,000 | 38,000 | 47,000 |
| High risk beneficiaries * | 2,400,000 | 38,000 | 21,000 |
| Annual cost** | 2,380,000 | 71,556 | 82,284 |
| Cost** per beneficiary | 0.43 | 1.86 | 1.83 |
| Cost ** per high risk beneficiary | 0.98 | 1.86 | 4.16 |

*Number of beneficiaries

**$U.S. 1991

*Cost-delivery Analysis*

Often the term “cost-effectiveness” is used erroneously in assessing the cost of achieving outputs rather than impacts. Such calculations, more accurately referred to as “cost-delivery” ratios, are, however, important in projects which may wish to compare alternative delivery systems, e.g., the relative cost per 100 calories delivered of food supplement production by a factory and by community women’s groups.

Phillips, et al. (1996) compare the relative cost-delivery ratios of three means of protecting mothers and children from Vitamin A deficiency in Guatemala: sugar fortification, high dose Vitamin A capsules, and nutrition education designed to encourage the production of Vitamin A-rich foods through home gardens. This comparison revealed the following:

Note that the comparisons do not take into account differences in effectiveness. Since the impact of these programs, involving assessment of Vitamin A status through blood tests, is expensive and difficult to carry out in most developing countries, this type of cost-delivery information is often used and assumes that nutrients in these programs which are delivered will lead to impact (a more valid assumption for capsules and fortifi-

cation than for gardens because of absorption issues)[30]. Cost-delivery calculations, however, are valuable in their own right to determine which of the approaches is most cost-efficient in delivering services intended to improve vitamin A status.

These data also demonstrate the importance of the potential scale of an intervention's coverage. Even if the capsule or gardening activities were found to be as efficient as fortification per unit of vitamin A delivered, it might be difficult for these approaches to achieve the same coverage in the short and medium term as food fortification.

It is important to note that the value of this analysis may extend beyond Guatemala; other countries unable to perform such an analysis can make more informed decisions on vitamin A distribution based on the Guatemala results. Of course each country, or even regions within a country will need to consider the relevant differences between the study area and their own.

In addition to cost-delivery analysis, which considers that cost of project outputs, it may be useful to assess the cost of achieving particular outcomes or behavioral changes, and, in some projects, the cost of bringing about particular changes in knowledge (an output assumption.)

*Sensitivity Analysis*

The comparison of three vitamin A distribution programs presented above can be used to illustrate the value of sensitivity analysis in project evaluation. Sensitivity analysis involves using a range of different values for an indicator whose actual value is uncertain and for which major as-

30. In this instance, it may be possible to relate cost-delivery data with impact by using scientific studies examining the effect of consumption of vitamin A in a particular form with improvements in blood retinol levels.

sumptions have been made. In cost-effectiveness analysis, for example, the specific assumptions made about costs (e.g., future foreign exchange rates or the opportunity costs of government staff) could significantly affect the calculations. In such cases a range of plausible values can be inserted to assess the degree to which results are influenced by the selection of one value as opposed to another. Even where values are known, it is often useful to insert more optimistic or pessimistic values to assess possible future scenarios. In assessing alternative vitamin A delivery systems**,** Phillips et al. (1996) introduced into the capsule and gardening calculations more optimistic assumptions on coverage rates, but found that they did not significantly affect the bottom line ratios in the comparison. On the other hand, the results were highly sensitive to the level of vitamin A found in fortified sugar samples in retail shops. At lower levels of concentration, fortification no longer has the best cost-delivery ratio, underlying the importance of effective monitoring and of maintaining adequate levels of Vitamin A in the sugar at the point of consumption.

*Disability-Adjusted Life Years*

As indicated in Section 3, it normally is unnecessary to calculate cost benefit ratios for individual nutrition projects. Where indications of “benefits” (e.g., mortality or morbidity decreases or productivity increases) are considered useful, it usually makes more sense to estimate them from the research literature, e.g., as summarized in Toolkit # 3. It should be noted, however, that health economists now regularly use Disability-Adjusted Life Years (DALYs) to compare the relative “benefits” of health-related interventions. The DALY expresses years of life lost to premature death and years lived with a disability of specified severity and duration. In cases such as vitamin A capsule distribution or iron/folate distribution to pregnant women, where both mortality and morbidity consequences are considered significant, and where the extent of each can be quantified, and in situations where choices (i.e., in a subsequent national health project) may have to be made among various health and nutrition interventions, calculations of DALYs for nutrition interventions may be useful.

# An Index for Assessing Project Sustainability: Application to Integrated Child Development Services (ICDS)

| Indicator | | Sustainability Status | Ranking on a 5-point scale* |
|---|---|---|---|
| **A. Continued delivery of services and benefits** | | | |
| A.1 | Volume of services delivered | Covers 75% of the development blocks in India. | 5 |
| | Stability of services delivered | Food supply in program is irregular. | 3 |
| A.2 | Efficiency of service delivery | With only one worker per 1000 population to provide services for both children (0–6 years) and pregnant and nursing women service delivery has not been efficient. | 2 |
| A.3 | Quality of service delivery | At present, the Indian Government places a higher premium on increasing coverage to the entire country rather than improving the quality of existing services. | 2 |
| A.4 | Satisfaction of beneficiaries | Although the food has not been well used for its intended purposes (i.e., nutritional improvement of the most vulnerable family members) it has augmented household food supply among largely food insecure households. In addition, a component of early childhood education is child care. The food and child care services have been greatly appreciated by most recipients. | 4 |
| A.5 | Distribution of benefits among economic and social groups | Projects are generally located in areas of particular need. | 4 |
| *Category Average* | | | 3.3 |
| **B. Program infrastructure and support** | | | |
| B.1 | Condition of physical infrastructure | Anganwadi centers are often inadequate and unhygenic. | 2 |
| B.2 | Condition of plant and equipment | no information available | |
| B.3 | Adequacy of maintenance procedures | no information available | |
| B.4 | Efficiency of cost-recovery and adequacy of operating budget | The budget is inadequate to meet the goals of the project. At present there is no cost-recovery. | 2 |
| B.5 | Beneficiary involvement in maintenance of procedures | no information available | |
| *Category Average* | | | 2.0 |

*Rating scale : 1 = very poor; 2 = poor; 3 = average; 4 = good; 5 = very good.

*Source:* Adapted from Valadez and Bamberger (1994), p. 194.

| Indicator | | Sustainability Status | Ranking on a 5-point scale* |
|---|---|---|---|
| **C. Long-term institutional capacity** | | | |
| C.1 | Capacity and mandate of the principal operating agencies | The national government perceives a clear mandate but has often been regimented and inflexible in its policies. State governments responsible for operation have varied enormously in terms of capacity. | 3 |
| C.2 | Stability of staff and budget of operating agency | Government funding for ICDS has been increasing regularly. While Anganwadi workers (AWW) are not considered government staff, turnover of these workers has been much lower than other parts of the world. | 4 |
| C.3 | Adequacy of interagency coordination | Interdepartmental coordination, particularly with health infrastructure, has been inadequate. | 2 |
| C.4 | Adequacy of coordination with community organizations and beneficiaries | While beneficiaries are called upon to provide land, buildings, fuel and labor, their involvement in ICDS activities has been marginal. Beneficiaries consider ICDS an externally provided government program. | 2 |
| C.5 | Flexibility and capacity to adapt project to changing circumstances | While there is little flexibility at the national government level, some states have initiated creative means of addressing identified problems. | 3 |
| *Category Average* | | | 2.8 |
| **D. Support from key stakeholders** | | | |
| D.1 | Stability and strength of support from international agencies | Actively supported by NGOs and international and bilateral assistance agencies. ICDS has been the primary focus of World Bank nutrition activity in India. | 5 |
| D.2 | Stability and strength of support from national government | The budget has been increased regularly over the past five years. Support for existing institutional arrangements is strong. | 5 |
| D.3 | Stability and strength of support from provincial and local government | Government support has varied enormously among the states. | 3 |
| D.4 | Stability and strength of support at the community level | There is little sense of community ownership at present (see C.4). | 2 |
| D.5 | Ability of program to be politically controversial | The only point of political controversy is the continuing demand of AWW for government worker status. | 5 |
| *Category Average* | | | 4.0 |
| *Overall Average for All Categories* | | | 3.0 |

*Rating scale : 1 = very poor; 2 = poor; 3 = average; 4 = good; 5 = very good.

*Source:* Adapted from Valadez and Bamberger (1994), p. 194.

The procedure (not detailed in this manual) for calculating the DALY of an intervention addressing a particular condition in a particular population involves the estimation and summing of the years of life saved by the intervention and the years of disability not suffered as a result of the intervention.

***Sustainability***

Even if the project is found to be cost-effective, based on the life of the project to date, stakeholders often are interested in assessing whether the project is likely to be "sustainable." By this they are asking whether the activity is likely to generate adequate commitment, and, in some cases, enough direct financial returns, and be resilient enough to withstand disturbances, so that, once the specific "project" being evaluated is completed, the activity will continue to be supported from some source, or be economically self-sustaining. Measuring such sustainability, given the range of factors which affect it, however, can be difficult.

In an effort to address the need for quantification, Valadez and Bamberger (1994) devised a sustainability index. To illustrate the value of the index for nutrition purposes, we use it above to assess the sustainability of the Indian Government's Integrated Child Development Services (ICDS). This program, introduced in the mid-1970's and now covering 75% of the rural development blocks in the country, seeks to combine in one program nutrition and health improvement for pregnant and nursing women and children 0–3 years of age, plus early childhood education for children 3–6 years of age.

The Valadez-Bamberger index is divided into four subsections: (A) Continued delivery of services and benefits, (B) Program infrastructure and support, (C) Long-term institutional capacity, and (D) Support from key stakeholders. The ICDS project is strongest in category "D", and weakest in category "B". If the four categories are weighted equally, ICDS emerges with a sustainability average of 3 or "average" on a 5 point

scale. If, however, it were decided that stakeholder support is relatively more important in the Indian nutrition context while infrastructure is relatively less important, so that the “D” category were double weighted while the “B” category were given a 0.5 weighting, the project would emerge with a score of 3.4, relatively more sustainable.

Clearly the index is subjective as are the weights assigned. Nonetheless the exercise is likely to provide insight into the question of longer term sustainability, and provide an additional dimension to a project evaluation. Additionally, if the index is adopted by an organization and applied with a standardized format, it may be useful in comparing similar projects.

***Analyzing Qualitative Information***

As indicated earlier, qualitative methods have an important place in most M&E systems but by their very nature less objective than quantitative data. Consequently, qualitative results need to be carefully verified. One way of doing this is to cross-check different sources of information. In addition to cross-checking data, M&E staff who are using qualitative techniques should beware of the following.

Particular care should be exercised when using “expert’ judgment. Although relevant experience may lead to important insights regarding project activity, experts with significant experience in a particular type of or intervention may allow pre-determined positions on issues to influence unduly their judgments in evaluating particular projects and also may overlook characteristics unique to the project being evaluated. This argues for carefully considering information provided by project managers, staff and beneficiaries.

When interpreting qualitative data, it is also important to determine how much of observed effects could be due to individuals seeking to please the evaluator. Individuals thinking they know what the evaluator wants to

### Analyzing M&E Results: Involving Stakeholders

Many creative methods exist for analysis of monitoring and evaluation results which permit the involvement of project stakeholders. Some examples include...

- regular stakeholder meetings to discuss monitoring information as it is gathered. Careful listening to the questions raised will improve the utility of data collected.
- focus group discussions (or questionnaires) on the decision-making needs of stakeholders and the range of options available.
- tabulation of results by community volunteers, who may also wish to participate in presentations.
- facilitation of "analysis meetings" with project staff and community participants to identify important patterns in the data suggesting, for example, patterns of malnutrition causality as particular predictors of project success or failure.

hear may provide the evaluator with less than fully objective responses. Similarly, project staff who know that an evaluator will visit may make a special effort to make things go well during the visit. Unannounced visits can reduce these effects.

In other cases, both beneficiaries and project staff may believe they can benefit from poor evaluation reports. Beneficiaries may anticipate increases in project services if they can convince an evaluator that current service levels are not providing satisfactory results. Project staff might

find it in their interests to try and convince an evaluator that they cannot reach beneficiaries or staff they supervise—and therefore require a new vehicle for the project. Comparison of responses/observations from many sources and comparison with quantitative results can help to identify and correct for this phenomenon.

Involving stakeholders (management, staff, beneficiaries, community members) at an early stage of the M&E process has considerable benefit. It permits M&E staff to consider unanticipated benefits and negative effects and allows them to respond early to concerns about results. But it is important that external M&E staff retain their own commitment to objectivity. They should carefully consider the concerns of project stakeholders, but also be willing to stand by unpopular results if these findings are substantiated by the evidence.

***Returning to the Conceptual Framework***

In addition to analyzing the numerical data, the conceptual framework or project map, first introduced in Section 2, should be revisited and used to draw conclusions from the M&E system as a whole, particularly in cases where some or all of the project objectives were not achieved.

As with the ideal process of project design, this analysis of results is best done by starting with impacts at the right side of the framework and moving to the left. If a particular impact has not been achieved, one can work backwards through the framework to identify the point or points where the process was derailed. An example of this backwards mapping process, using TINP I, can be found in Annex 7.

# 10 Maximizing the Usefulness of Results



If monitoring and evaluation have been conducted properly, results should permit reasonable conclusions about the effectiveness of a project and insights into its strengths and weaknesses. A review of evaluations over time, however, indicates that efforts to ensure the accuracy of results does not translate automatically into policy or program changes. Instead, evaluation results usually represent just one input into the very complex process of decision-making. In addition to having accurate and relevant information, other considerations determine the extent to which monitoring and evaluation findings influence decisions regarding a project's future. This section addresses the critical process of how to maximize the likelihood that decision-makers will actually use the information provided.

***Ownership of the M&E Process***

Experience suggests that the likelihood of results being used is often directly proportionate to the sense of "ownership" that decision-makers feel in regard to the process. If decision-makers have a strong sense of "ownership" or involvement, they are more likely to incorporate the results into ongoing or future activities. This is a perfectly human reaction. Yet many well executed monitoring and evaluation systems fail to affect decision-making because the ownership factor is not taken into account.

Those professionals who view monitoring and evaluation as technical tasks to perform rather than a mutual process to engage usually neglect ownership issues. As a result, evaluations are conducted, results analyzed, and reports presented, a normal sequence for research in general. But decision-makers presented with final reports into which they have had little input are unlikely to accept the results**,** *particularly* if the results are not consistent with their own preconceptions or assumptions.

Even the findings of an internal monitoring system may not be adequately used for project improvement purposes if decision-makers are inadequately involved in the monitoring process as it evolves.

The following field insight, “Absence of Ownership”, provides an example where a well executed evaluation but an inadequate evaluation *process* meant that important results were shelved and ignored. In this instance, the results were at odds with a working assumption, convincing even to legislative critics, that food-for-work projects serve the best interests of the poorest households. The results were also likely to upset the domestic wheat lobby in the donor country.

Yet, as indicated, outright rejection of the evaluation might not have resulted if the evaluation process had been more inclusive, if results and drafts had been shared, if alternative explanations had been explored, and if technical means had been considered to address the problems identified. Such dialogue might well have allowed decision-makers to be more comfortable with evaluation results that were at odds with program assumptions.

Evaluators, of course, need to be careful not to be persuaded by project stakeholders to alter valid conclusions. But if stakeholders are included in every step of the evaluation process, they are more likely to accept the results, knowing that their alternative explanations were considered.

Furthermore, the purpose of an evaluation should never be presented as simply determining if the program was “good or bad”, or if it “worked or didn’t work.” Rather, the evaluation should concentrate on the extent to which project objectives were accomplished and on means by which particular project components could be changed or improved. Evaluations should tailor their recommendations to the range of realistic policy alternatives available, where possible. If a project is not adequately reaching the target group, for example, it might be possible to recommend adding new locations or hours, or changing the outreach process, rather than concluding that the project “didn’t work” and should be scrapped.

## Field Insight: Absence of "Ownership"

Informal assessment of a food-for-work program in Bangladesh turned up enough doubts about the program that the country office of the food aid donor requested a more formal evaluation. Although the home office, under pressure from its domestic wheat lobby to keep this outlet open, did not endorse the evaluation, it could not oppose it altogether.

The evaluation, organized by the country office, and carried out by a local contractor, found that the decentralized decision-making mode of the project had the real effect of putting decision-making on project selection and location into the hands of large local land owners who used the program to their own advantage. The result was an increasingly skewed income distribution, an increase in land sales, and, ironically, an increase in the numbers of workers showing up for food-for-work activities.

The evaluation results were carefully tabulated and sent to the home office where they were immediately shelved, never to resurface.

In retrospect, the field office and the evaluator might have been more pro-active in seeking to assure some translation of their results into policy measures. Below are some possible steps that could have been taken to prevent such discord between the field and home offices.

- At the outset, the home office might have been asked to collaborate in the evaluation, perhaps through a representative who might physically participate.
- Drafts of the evaluation method could have been sent to the home office for comment.
- Initial results could have been shared with the home office for comment as they became available.
- Efforts could have been made to tailor the report, at least in part, to the range of policy choices actually open to the decision-makers. Therefore, instead of being read as an indictment of the program as a whole, the evaluation could have underlined specific steps in selection or approval processes likely to reduce or eliminate the abuses.

### *Effective Presentation of Findings and Recommendations*

The likelihood of results being used also increases if findings are presented in an effective manner. Usually, evaluation results are presented both in oral and written form. Most of the following comments relate to both. A brief discussion specific to oral presentations is offered at the conclusion of this section.

#### *Format*

The format of oral presentations and written reports is equally as important as the results themselves in affecting decision-making. A report which includes a succinct executive summary, clearly presented tables and charts which highlight findings plus real examples and concrete recommendations can be readily used for subsequent decisions. Monitoring reports can be used for management purposes, utilizing the principle of “management by exception” (see Section 3), and to compare outputs, outcomes, assumptions and impact “snapshots” from one period to another. Endpoint evaluation reports used together with baseline data and midterm evaluations permit the examination of changes which have resulted from the project overtime. Evaluation reports can also be used in making further determinations and assist in decision-making on a particular project, and can even be referred to subsequently in the planning of future activities or projects.

#### *Challenging assumptions*

But what of the results themselves? Beyond the challenges of presentation, many evaluators face an inherent dilemma with the content itself. It is possible that results may be viewed as wholly consistent with accepted understandings and expectations, and thus, the report itself may receive little attention. Conversely, if the results are seriously at odds with prevailing assumptions, they may be similarly dismissed.

In the field insight, “Absence of Ownership”, the results were clearly counter to the donors’ assumptions and expectations. In this instance, a better skilled presenter would have made it a point to state clearly those assumptions and expectations *before* leading the reader or the listener to results that call them into question.

### *Oral presentations*

While only one written evaluation report is likely to be prepared, in a large project, it may be useful to arrange a series of oral presentations of results, for particular groups of stakeholders. These will necessarily vary according to the particular interests of each group. All groups are likely to be interested in overall results, i.e., did the inputs achieve the desired results? Beyond that, however, project participants are likely to be concerned primarily with service delivery, while district level managers may be particularly interested in the effectiveness of information flows, and planners and donors in the cost effectiveness and structural soundness of the intervention.

In each case, a primary focus of the presentation, as with the M&E system as a whole, should be on lessons learned which can improve this project or others like it in the future.

## *The “Usability” of Results*

In addition to issues of format, results must be credible (see table 10.1), delivered in a timely fashion, and, perhaps most importantly, consistent with implementation realities.

### *Timeliness*

Results must be available at the time when decisions are to be made. External events may dictate the timelines of evaluations (e.g., legislative debates, project review deadlines, budgetary decisions); an evaluation

### *Table 10.1* How to Improve the Use of M&E Findings

| Format | Relevance to the Needs of Decision-Makers |
|---|---|
| • Begin with an executive summary, and where appropriate, a table of contents and a list of acronyms.<br>• Keep the report concise.<br>• Use simple, clear language.<br>• Use tables, charts and graphs to summarize results.<br>• Synthesize available findings.<br>• Use real examples.<br>• Make concrete recommendations.<br>• List "lessons learned"<br>• Provide references for additional sources of information. | • Meet the timing of the project design and review cycles.<br>• Prepare and disseminate the reports quickly.<br>• Ensure that the recommendations are relevant and implementable.<br>• Present alternative recommendations so decision-makers can choose the most appropriate action given resource and other constraints. |
| **Credibility** | **Dissemination** |
| • Develop direct relationships with staff early and maintain throughout the course of an evaluation.<br>• Develop relationships with project participants and community leaders.<br>• Use reputable information sources.<br>• Assure that the information is accurate and unbiased.<br>• Discuss the findings informally with the Project Director and other staff before finalizing the report. Include their comments and suggestions to the extent possible. | • Involve stakeholder groups in discussions of preliminary results as they are produced.<br>• Clear the report with key parties before it is formally presented (while being careful to maintain the integrity of the report).<br>• Decide who should receive reports in advance.<br>• Share findings with project participants and community leaders. |

*Sources:* Adapted from Florio et al. (1979) as cited in Rossi and Freeman (1993), p. 451, and Valadez and Bamberger (1994), p. 437

team should be knowledgeable about the timing of such decision-making cycles. Often valuable evaluation results have been unusable, at least in the short run, because they arrived too late. For monitoring systems timeliness is defined by internal needs; the program manager has more dis-

cretion over when results will be used. However, the periodic compilation of data should be set when creating a monitoring system and followed through out the course of the project.

***Feasibility potential***

The “usability” of results and recommendations also depends importantly on the extent to which they can be put into practice. Efforts to tailor the report, at least in part, to the range of policy choices actually open to the decision-makers, and concrete and realistic suggestions to address project-specific problems increases its usefulness. If the evaluation suggests concrete and realistic steps to address project-specific problems, it is more likely to be implemented, instead of being disregarded.

Recommending biochemical anemia screening for pregnant women, for example, will not be helpful if the project has no capacity to collect blood or carry out field-based analyses. Likewise, results suggesting an increased reliance on caring practice messages for low income women will not be usable if these women already face untenable time constraints.

In the above cases, a broad understanding of implementation realities will usually be sufficient. But the problem can be more complicated. In the following field insight , “A Locked-In Implementation System”, the project at hand was so efficiently organized, with such a tight, interlocking management system in place, that changing even a single piece of the matrix was exceedingly difficult. In this instance, even a decision-maker who was receptive to a particular evaluation result would be hard pressed to know how to put it into action.

Being cognizant of the importance of “ownership,” of prevailing assumptions and understandings among stakeholders, of feasibility, issues facing implementers, of the presentation process and of the timeliness of reports and recommendations, will improve the likelihood that M&E will translate into desired policy change and/or project improvement. Atten-

## Field Insight: A Locked-In Implementation System

In the previously discussed first Tamil Nadu Integrated Nutrition Project (TINP I), the final evaluation and project completion report identified a series of major problem areas which were systematically, and, in retrospect, effectively addressed in the design of the continuation project, TINP II. Interim evaluations of TINP I, however, proved much less effective in terms of their effects of project operations.

In retrospect the problem was related, ironically, to the tightly organized and administered nature of the project itself, in other respect undeniable assets in a project. After the first highly experimental year of the project, tight interlocking systems were put in place that proved difficult to adjust without making changes in the system as a whole, a daunting task. The result was that interim evaluation reports with suggestion, for example, on multiple growth monitoring session, and revised entrance and exit criteria, were read and digested but did not lead to significant alteration of project implementation.

If, for example, project managers had chosen to implement a recommendation calling for the monthly collection of pregnancy weight gain data, this would have required not only the purchase of new weighing scales with effects on the project's budget cycle, but also revised job descriptions for community workers, revised personnel evaluation forms, revisions in the project's training manual, revised supervisory guidelines, revised home visitation forms, revised monitoring guidelines, and revisions in the project's management information system. In such a tightly managed project, the necessity of such extensive revisions clearly represented a disincentive to implement midterm recommendations.

tion to these often undervalued issues may well make the difference in assuring that a well designed and executed M&E system, and the often considerable effort devoted to it, are well used.

In sum, it should be clear that a well monitored and evaluated nutrition project is likely to emerge as a better designed project and a better managed one, that stakeholders will be better informed and quite likely more involved as a result, and that rational decisions about the project are likely to be made. The manual hopefully has conveyed as well that, with a few exceptions (e.g., sampling) requiring specialized assistance, M&E systems are feasible and can be tailored to the needs of individual projects. Meanwhile a broader understanding of M&E concepts and the value of such systems is likely to increase their use and, as a result, improve the quality of nutrition projects themselves.

# Annex 1: Program Constraints Assessment (PCA)



A useful exercise to accompany monitoring and evaluation is the “program constraints assessment” (PCA). This tool, originally developed to facilitate the development of program driven nutrition research and training agendas, is now also used more broadly for periodic stocktaking in a project. A PCA, carried out at the time of a mid-term evaluation, could provide a valuable qualitative complement to the data normally collected, and a useful agenda for corrective action.

Essentially the tool involves information collection from project staff and beneficiaries (individually or in groups) to identify constraints inhibiting project effectiveness. These constraints may be grouped in particular categories, such as the following which were used in the PCA in Gauteng Province of South Africa:

- Political viability and vested interests
- Problem definition and program thrust
- Target group definition and targeting
- Program implementation
- Monitoring and evaluation

Once the constraints are identified, the means to address them are then identified in conjunction with staff and project beneficiaries. The necessary actions for project improvement are then grouped into the following categories:

1) Technical adjustment
2) Policy change
3) Operations research
4) Training

***Table A1.1*** **Illustrative Example of a PCA Table Entry**

| Constraint | Technical Adjustment | Policy Change | Operations Research | Training |
|---|---|---|---|---|
| Inadequate identification (in a community-based project) of pregnant women who are nutritionally "at risk" | Provision of equipment to collect anthropometric measurements on women | Policy directive explaining the importance of maternal nutrition and the need for intensified action | Study to identify appropriate body mass index (BMI) cut-off points for the population | Training of staff in the use of BMI and appropriate follow-up action |

An illustration of the Program Constraints Assessment approach is provided in Table A1.1.

A complete PCA would provide adequate detail for these needed actions, ideally with time and cost estimates.

## *Table A2.1* Conceptual Framework of a Simplified Vitamin A Supplementation Project for Children

Inputs → Outputs → Outcomes → Impacts → Benefits

| | Assumptions | | Assumptions | | | |
|---|---|---|---|---|---|---|
| Supplies of vitamin A oral doses | Timeliness and quality of supply<br><br>Adequate storage provision | Targeted children receive vitamin A supplements | Proper record-keeping on growth charts assures coverage and avoids risk of double supplementation | | Increase in vitamin A stores<br><br>Reduced kerato-malacia and night blindness in target population | Reduced morbidity, mortality, and preventable blindness |
| Organization of vitamin A supplementation days or incorporation into National Immunization Campaigns | Information is properly disseminated<br><br>The population is motivated to attend | | | | | |
| Staff training | Staff are well trained & motivated | | | | | |
| Development, production, and delivery of IEC materials | Materials adequately explain value of vitamin A consumption and address resistance points | | | | | |

## *Table A2.2* Conceptual Framework of a Simplified Targeted Food Supplementation Project for Children

Inputs → Outputs → Outcomes → Impacts → Benefits

| | Assumptions | | Assumptions | | | |
|---|---|---|---|---|---|---|
| Adequate supplies of food supplement | Timeliness and quality of supply | On-site daily food provision for children moderately or severely malnourished or experiencing growth faltering | Identified children are brought daily for feeding (and counseling for parents) | Children's daily caloric and micronutrient intake increases | Reduced prevalence of moderate and severe malnutrition | Reduced morbidity and mortality |
| | Adequate storage provision | | | | | Increased learning capacity |
| Staff training | Staff are well trained & motivated | | | | | Higher productivity in the long-term |
| Regular growth monitoring | High attendance at growth monitoring sessions | | | | | |

## *Table A2.3* Conceptual Framework of a Simplified Iron Fortification of Flour Project

Inputs ⟶ Outputs ⟶ Outcomes ⟶ Impacts ⟶ Benefits

| | Assumptions | | Assumptions | | | |
|---|---|---|---|---|---|---|
| Necessary national legislation (including enforcement mechanisms) | Legislation is binding | Wheat flour is fortified | | | Reduced iron deficiency anemia | Reduced morbidity<br><br>Increased learning capacity |
| Agreements with mills to fortify flour | Compliance with agreements | | | | | Higher productivity in the long-term |
| Equipment and training for mills and mill staff | | | | | | |
| Supply of appropriate iron compound | Timeliness and quality of supply | | | | | |
| Development, production, and delivery of IEC materials<br><br>Provision of subsidy (or an understanding that mills or population will absorb the added cost) | Materials adequately explain value of fortification and address resistance points | Target population consumes fortified flour | | Population chooses fortified flour (if there is a choice) | | |

# Annex 3: General M&E Checklist

_____ Clarify project goals and objectives
_____ Develop a conceptual framework of the project
_____ Determine information needs
_____ Hire staff or consultants (if necessary)
_____ Establish a multi-disciplinary team for M&E
_____ Select an M&E coordinator
_____ Compile and disseminate background information
_____ Determine training needs
_____ Plan and implement training
_____ Determine the level of computerization
_____ Purchase necessary equipment (including computer hardware and software)
_____ Devise a timeline for remaining tasks
_____ Select an evaluation design
_____ Determine sample size
_____ Identify a control group
_____ Choose a sampling plan
_____ Select and review indicators
_____ Decide data collection methods
_____ Develop the data collection instruments
_____ Pretest the instruments
_____ Collect pre-project information on participants and controls through a baseline survey
_____ Collect project information for ongoing monitoring and evaluations
_____ Compile data
_____ Clean data
_____ Analyze data
_____ Prepare the report(s) of the findings
_____ Disseminate the information
_____ Use the information to improve the project
_____ Use the information to inform decision-making on the future of the project
_____ Use the information to inform decision-making on related policies

# Annex 4: Illustrative Data Collection Forms Which Facilitate "Management by Exception"

**Based on those used in the Bangladesh Integrated Nutrition Project (BINP)**

***Project Description***

This exercise, potentially valuable for training purposes, utilizes data collection forms based on those used in the Bangladesh Integrated Nutrition Project. The project's premise is that modifications of mother's feeding behaviors, chiefly the more timely introduction of complementary food, can eliminate the bulk of nutritional problems in this age group. To motivate mothers to introduce complementary foods, the project combines growth monitoring (to identify children at risk) with a small feeding supplement (to demonstrate the benefits of complementary feeding) as demonstration tools for behavioral modification.

Children whose growth is failing enter a three-month feeding program in which they receive a small daily supplement. The supplement is small enough (150 kCal) that most mothers can easily afford to prepare it, or a similar type of food, at home. Although growth faltering children are the primary target, severely malnourished children also are included in the feeding.

For this exercise, the project is assumed to be in it's seventh month of field implementation. Prior to the project's initial field work, intense communication efforts informed communities about the project, about who would be eligible for services, and about the services available. Women's groups were formed to increase community involvement. Two months before the first weighing session, project staff conducted a household census throughout the project area to identify children under 2 and to gather basic household information. Mothers of children under 2 received a growth chart marked with the child's weight at the time of the census. All women of childbearing age (15–44 years of age) were given an identification number that could be linked to their records from the household census.

Project Field Staff, the Community Nutrition Promoters (CNPs), serve a population of 1,500 people, approximately equivalent to one village.

Each month, CNPs, who are volunteers that live in the village, conduct two days of growth monitoring at the local Community Nutrition Center (CNC) for children under 2 and record each child's weight on the growth chart kept by the mother and on the "Weighing Session Report." (It is the CNP's responsibility to register newborns for project services as births occur). Mothers of children who are not present for weighing on the first day are visited by members of the women's group to encourage them to attend on the second day. Any child six months or older whose growth is faltering (failure to gain 300g during the previous month under the age of one year; failure to gain 150g when older than one year), or who is classified as severely malnourished (< 60% of NCHS median weight for age) is eligible to receive food supplements. CNPs also refer severely malnourished children to health services for care from a trained provider. The CNP provides the mother with a referral form to take to the Health Centre, but it is the mother's responsibility to take the child. Health Services are provided free of charge at the Government Health Centre, but mothers may take referred children to any provider they choose.

The CNC is open for feeding six mornings each week, Monday through Saturday. The project's food supplement is a packaged product containing dried, ground cereal, dried ground pulse, and molasses. At the time of feeding the members of the women's group, under CNP supervision, mix the packaged product with a pre-measured micronutrient powder (a sachet-packaged formulation of vitamin A, iron, vitamin C, and B vitamins procured from a national pharmaceutical company), soybean oil, and water from a well at the CNC. Growth-faltering children receive one supplement, approximately 150 kCal, and severely malnourished children receive two.

The CNP keeps feeding attendance records in the "Feeding Register." No record is kept of the amount of supplement consumed. Once they enter the feeding program, children receive daily food supplements at the center for three months, "graduating" from feeding if they gain adequate weight during the final month of feeding. If a child does not

graduate from feeding after three months, CNPs enroll them for a second three-month period and refer them to health services. If a child does not graduate after a second three months of feeding, CNP and her supervisor together visit the child's home to assess whether a third round of feeding is appropriate.

### *Monitoring Tools*

#### *The Weighing Session Report (WSR)*

The CNP keeps the Weighing Session Report at the CNC. As a child becomes a project participant, either through the initial household census or as a new birth, CNPs enter her or his name on the report along with an identification number, her/his date of birth, and the name of her/his mother. Each month the CNP records the child's weight and determines whether her/his growth is faltering and whether she/he is severely malnourished, and she notes whether she has referred the child to health services. Each month a total is entered on the report for number of children weighed, identified as growth faltering, severely malnourished, and children referred.

#### *The Feeding Register*

The CNP records in the Feeding Register attendance for each eligible child. At the end of the month, she calculates her CNC's feeding coverage percentage as described on the register.

### *Project Monitoring System*

Each month the CNP meets with her supervisor to review records from the CNC, and together they complete the "CNC Monthly Progress Report" (MPR) using the data from the current month's Weighing Session Report and the recently completed month's Feeding Register. During this meeting they check data quality and look for specific information requir-

ing additional intervention, like a home visit or a referral to the Health Center. After meeting with each of the 10 CNPs she supervises, the Supervisor tabulates the MPR results from all 10 CNPs she supervises to create a "Supervisor's MPR" (SMPR), which she sends to the District Manager (DM). The DM tabulates all SMPRs from the district and sends a District Report to the Project Office.

### *Training Exercise*[31]

#### *CNC Review*

Using the forms provided, review the CNC records for completeness, and then complete the following:

1. Calculate the monthly totals for the CNC, on the Weighing Report, the Feeding Register, and the Monthly Progress Report.

2. Review the data on all the forms, checking especially for any problems—for example, children who fail to graduate from feeding, inconsistencies in weighing or record keeping, and chronically poor attendance at growth monitoring or feeding sessions.

3. Write a brief analysis of progress at your particular CNP, including recommendations for action. Most important to identify are specific steps to be taken by the CNP or members of the women's group in response to data found on the reporting forms.

---

31. In the charts (e.g., monthly progress reports) provided here, the numbers have been calculated. In the actual exercise, blank forms of the monthly progress reports and the supervisor's monthly progress report should be presented to participants for completion. Note that to facilitate this exercise anthropometric reference data is provided following the forms.

***Data Aggregation***

Using the totals from your CNC review, along with your CNC analysis, join with others in your same Supervisory area (see reference number on the CNC forms).

1. Create a single, aggregated SMPR for your working area.

2. Review the data from the various CNCs, looking for cases of exemplary performance (the reasons for which might be important to share with other CNPs) or inadequate performance (which, utilizing the principle of management by exception, might "trigger" a supervisory visit to review CNP skills, observe food supplement preparation, check the condition of equipment, and/or referral services being delivered by health clinics).

3. Write a brief analysis of progress in your working area, including recommendations for action.

**CNC 01**
**Total Children < 2 yrs: 10**

1997 Weighing Session Report

| Child's ID | Birth Date | Sex | Mother's ID | Jan Weight | Jan Status | Jan Falter | Jan Referral | Feb Weight | Feb Status | Feb Falter | Feb Referral | Mar Weight | Mar Status | Mar Falter | Mar Referral |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 101 | 24-Jun-95 | M | M207 | 8.05 | | | | 9.15 | | | | 9.55 | | | |
| 102 | 15-Jul-95 | M | M101 | 8.35 | | | | 8.45 | | Yes | | 8.65 | | | |
| 103 | 25-Sep-95 | F | M46 | 6.00 | Sev | | Y | 6.35 | | | | 6.55 | | | |
| 104 | 30-Oct-95 | F | M65 | 9.35 | | | | 9.50 | | | | 9.75 | | | |
| 105 | 14-Jan-96 | F | M44 | 8.55 | | | | 8.95 | | | | 9.30 | | | |
| 106 | 13-Feb-96 | F | M48 | 5.35 | | | | 5.70 | | | | 5.75 | | Yes | |
| 107 | 21-Mar-96 | M | M47 | 5.60 | | | | 5.85 | | Yes | | 6.00 | | Yes | |
| 108 | 22-Apr-96 | M | M224 | 6.15 | | | | 6.70 | | | | 7.05 | | | |
| 109 | 22-May-96 | F | M93 | 4.95 | | | | 5.25 | | | | 5.10 | Sev | Yes | Y |
| 110 | 17-Dec-96 | M | M45 | 2.60 | | | | 3.25 | | | | 3.85 | | | |
| | | | | | 1 | 0 | 1 | | 0 | 2 | 0 | | 1 | 3 | 1 |

| | | | | Apr | | | | May | | | | Jun | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Child's ID | Birth Date | Sex | Mother's ID | Weight | Status | Falter | Referral | Weight | Status | Falter | Referral | Weight | Status | Falter | Referral |
| 101 | 24-Jun-95 | M | M207 | 9.60 | | Yes | | 9.80 | | | | 10.00 | | | |
| 102 | 15-Jul-95 | M | M101 | 8.75 | | Yes | | 8.95 | | | | 9.25 | | | |
| 103 | 25-Sep-95 | F | M46 | 6.70 | | | | 6.90 | | | | 6.65 | Sev | Yes | Y |
| 104 | 30-Oct-95 | F | M65 | 9.95 | | | | 10.10 | | | | 10.00 | | Yes | |
| 105 | 14-Jan-96 | F | M44 | 9.55 | | | | 9.85 | | | | 10.05 | | | |
| 106 | 13-Feb-96 | F | M48 | 5.90 | | | | 6.15 | | | | 6.10 | Sev | Yes | Y |
| 107 | 21-Mar-96 | M | M47 | 6.00 | Sev | Yes | | 6.15 | Sev | | | 6.35 | Sev | | Y |
| 108 | 22-Apr-96 | M | M224 | 7.65 | | | | 7.80 | | Yes | | 8.00 | | | |
| 109 | 22-May-96 | F | M93 | 5.40 | | | | 5.70 | | | | 6.05 | | | |
| 110 | 17-Dec-96 | M | M45 | 4.35 | | | | 4.80 | | | | 4.65 | Sev | Yes | Y |
| | | | | | 1 | 3 | 0 | | 1 | 1 | 0 | | 4 | 4 | 4 |

Sev = severely malnourished.

**CNC 02**
**Total Children < 2 yrs: 10**

| 1997 Weighing Session Report | | | | Jan | | | | Feb | | | | Mar | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Child's ID | Birth Date | Sex | Mother's ID | Weight | Status | Falter | Referral | Weight | Status | Falter | Referral | Weight | Status | Falter | Referral |
| 201 | 12-Jun-95 | F | M111 | 9.45 | | | | 9.75 | | | | 9.85 | | Yes | |
| 202 | 15-Jul-95 | M | M95 | 7.85 | | | | 8.05 | | | | 8.15 | | Yes | |
| 203 | 4-Nov-95 | M | M42 | 6.35 | | | | 6.35 | Sev | Yes | Y | 6.70 | | | |
| 204 | 18-Dec-95 | M | M216 | 9.90 | | | | 10.35 | | | | 10.50 | | | |
| 205 | 28-Mar-96 | F | M225 | 5.45 | | | | 5.85 | | | | 6.15 | | | |
| 206 | 8-Apr-96 | M | M240 | 5.20 | Sev | | Y | 5.55 | | | | 5.85 | | | |
| 207 | 15-May-96 | F | M218 | 5.05 | | | | 5.40 | | | | 5.70 | | | |
| 208 | 21-May-96 | F | M262 | 5.05 | | | | 5.45 | | | | 5.75 | | | |
| 209 | 13-Jul-96 | F | M189 | 5.55 | | | | 6.10 | | | | 6.55 | | | |
| 210 | 25-Nov-96 | F | M63 | 2.85 | | | | 3.45 | | | | 4.15 | | | |
| | | | | | 0 | 0 | 1 | | 0 | 1 | 1 | | 0 | 2 | 0 |

| | | | | Apr | | | | May | | | | Jun | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Child's ID | Birth Date | Sex | Mother's ID | Weight | Status | Falter | Referral | Weight | Status | Falter | Referral | Weight | Status | Falter | Referral |
| 201 | 12-Jun-95 | F | M111 | 10.00 | | | | 10.35 | | | | 10.50 | | | |
| 202 | 15-Jul-95 | M | M95 | 8.55 | | | | 8.65 | | Yes | | 10.05 | | | |
| 203 | 4-Nov-95 | M | M42 | 6.85 | | Yes | | 6.90 | | Yes | | 7.25 | | | |
| 204 | 18-Dec-95 | M | M216 | 10.80 | | | | 11.00 | | | | 11.15 | | | |
| 205 | 28-Mar-96 | F | M225 | 6.55 | | | | 6.85 | | | | 7.00 | | | |
| 206 | 8-Apr-96 | M | M240 | 6.15 | | | | 6.25 | | Yes | | | | | |
| 207 | 15-May-96 | F | M218 | 5.95 | | Yes | | 6.25 | | | | 6.65 | | | |
| 208 | 21-May-96 | F | M262 | 6.15 | | | | 6.45 | | | | 6.75 | | | |
| 209 | 13-Jul-96 | F | M189 | 7.15 | | | | 7.55 | | | | 8.05 | | | |
| 210 | 25-Nov-96 | F | M63 | 4.95 | | | | 5.25 | | | | 5.45 | | Yes | |
| | | | | | 0 | 2 | 0 | | 0 | 3 | 0 | | 0 | 1 | 0 |

Sev = severely malnourished.

**CNC 03**
**Total Children < 2 yrs: 10**

1997 Weighing Session Report

| Child's ID | Birth Date | Sex | Mother's ID | Jan Weight | Jan Status | Jan Falter | Jan Referral | Feb Weight | Feb Status | Feb Falter | Feb Referral | Mar Weight | Mar Status | Mar Falter | Mar Referral |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 301 | 18-May-95 | M | M21 | 8.00 | | | | 8.00 | | Yes | | 8.00 | | Yes | |
| 302 | 31-May-95 | M | M92 | 8.00 | | | | 9.00 | | | | 9.00 | | Yes | |
| 303 | 12-Jun-95 | F | M07 | 9.00 | | | | 9.00 | | Yes | | 9.00 | | Yes | |
| 304 | 15-Sep-95 | M | M57 | 8.00 | | | | 9.00 | | | | 9.00 | | Yes | |
| 305 | 14-Jan-96 | F | M243 | 6.00 | | | | 6.00 | | Yes | | 6.00 | | Yes | |
| 306 | 21-Jan-96 | M | M39 | 6.00 | | | | 7.00 | | | | 7.00 | | Yes | |
| 307 | 17-May-96 | M | M131 | 5.00 | | | | 6.00 | | | | 6.00 | | Yes | |
| 308 | 13-Jun-96 | M | M328 | 5.00 | | | | 5.00 | | Yes | | | | | |
| 309 | 26-Nov-96 | M | M25 | 3.00 | | | | 3.00 | | Yes | | 4.00 | | | |
| 310 | 19-Feb-97 | M | M268 | | | | | | | | | | | | |
| | | | | | 0 | 0 | 0 | | 0 | 5 | 0 | | 0 | 7 | 0 |

| | | | | Apr | | | | May | | | | Jun | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Child's ID | Birth Date | Sex | Mother's ID | Weight | Status | Falter | Referral | Weight | Status | Falter | Referral | Weight | Status | Falter | Referral |
| 301 | 18-May-95 | M | M21 | 9.00 | | | | 10.00 | | | | 10.00 | | Yes | |
| 302 | 31-May-95 | M | M92 | 10.00 | | | | 10.00 | | Yes | | 11.00 | | | |
| 303 | 12-Jun-95 | F | M07 | 10.00 | | | | 10.00 | | | | 10.00 | | Yes | |
| 304 | 15-Sep-95 | M | M57 | 10.00 | | | | 10.00 | | | | 11.00 | | | |
| 305 | 14-Jan-96 | F | M243 | | | | | 7.00 | | | | 8.00 | | | |
| 306 | 21-Jan-96 | M | M39 | 6.00 | Sev | Yes | | 7.00 | | | | 7.00 | | Yes | |
| 307 | 17-May-96 | M | M131 | 7.00 | | Yes | | 7.00 | | Yes | | 8.00 | | | |
| 308 | 13-Jun-96 | M | M328 | 6.00 | | Yes | | 7.00 | | | | 7.00 | | Yes | |
| 309 | 26-Nov-96 | M | M25 | 7.00 | | | | 8.00 | | | | 8.00 | | Yes | |
| 310 | 19-Feb-97 | M | M268 | 4.00 | Sev | Yes | | | | | | 6.00 | | | |
| | | | | | 2 | 4 | 0 | | 0 | 2 | 0 | | 0 | 5 | 0 |

Sev = severely malnourished.

**Feeding Register** **CNC 01** **Month/Year: May/97**

| May 97 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Child's ID | Mother's ID | | | | | | | Weighing | Weighing | | | | | | | | | | | | | | | | | | | | | | | | | | | Weighing | | | | | |
| 101 | M207 | | | | | | | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | | | | | 24 |
| 102 | M101 | | | | | | | x | x | x | x | | x | x | x | x | x | | | x | x | | x | | x | | | x | x | x | x | | | | | | | | | 17 |
| 108 | M48 | | | | | | | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | | | | | 24 |
| 107 | M47 | | | | | | | x | x | | | | | | x | x | | | | | | x | x | | | | | | x | x | | | | | | | | | | 8 |
| 109 | M93 | | | | | | | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | | | | | 24 |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

(1) CNC Total Child-Days Feeding 97

(2) Feeding Days [# days CNC open for feeding] 24

(3) Number Children in Feeding 5

(4) Total Possible Person-Days Feeding [(2) x (3)] 120

(5) Percent Feeding Coverage [(1)/(4) x 100] 80.8

## Feeding Register

**CNC 02**

**Month/Year: May/97**

| May 97 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Child's ID | Mother's ID | Weighing | Weighing | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | Weighing | | | |
| 201 | M111 | | | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | | | | 28 |
| 202 | M95 | | | x | | | | x | x | x | x | | x | x | x | | | | | x | | x | x | x | | | | | x | x | x | | | | x | x | | | | 17 |
| 203 | M42 | | | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | | | | 28 |
| 206 | M240 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 0 |
| 207 | M218 | | | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | | | | 28 |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

(1) CNC Total Child-Days Feeding 101

(2) Feeding Days [# days CNC open for feeding] 28

(3) Number Children in Feeding 5

(4) Total Possible Person-Days Feeding [(2) x (3)] 140

(5) Percent Feeding Coverage [(1)/(4) x 100] 72.1

## Feeding Register

**CNC 03**

**Month/Year: May/97**

| May 97 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Child's ID | Mother's ID | | Weighing | Weighing | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | Weighing | | |
| 301 | M21 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| 302 | M92 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| 303 | M07 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| 304 | M57 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| 305 | M243 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| 306 | M39 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| 307 | M131 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| 308 | M328 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| 310 | M268 | | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | x | x | | x | x | x | x | | | 30 |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

(1) CNC Total Child-Days Feeding 270

(2) Feeding Days [# days CNC open for feeding] 30

(3) Number Children in Feeding 9

(4) Total Possible Person-Days Feeding [(2) x (3)] 270

(5) Percent Feeding Coverage [(1)/(4) x 100] 100.0

## Monthly Progress Report

**CNC No.: 01**
**Month/Year: June/97**

| | |
|---|---|
| (1) Total Children < 2 years | 10 |
| (2) Total Children Weighed | 10 |
| (3) Total Children Growth Faltering | 4 |
| (4) Total Children Severely Malnourished | 4 |
| (5) Total Children Referred | 4 |
| (6) Total Children in Feeding | 5 |
| (7) Total Child-Days Feeding | 97 |
| (8) Total Possible Feeding Days | 120 |
| (9) Percent Weighing Coverage [(2)/(1) x 100] | 100.0 |
| (10) Percent Growth Faltering [(3)/(2) x 100] | 40.0 |
| (11) Percent Severely Malnourished [(4)/(2) x 100] | 40.0 |
| (12) Percent Children Referred [(5)/(2) x 100] | 40.0 |
| (13) Percent Feeding Coverage [(7)/(8) x 100] | 80.8 |

Follow-up necessary for . . .
Percent Weighing Coverage < 60%
Percent Severe Malnourished > 15%
Percent Feeding Coverage < 75%

## Monthly Progress Report

**CNC No.: 02**
**Month/Year: June/97**

| | |
|---|---|
| (1) Total Children < 2 years | 10 |
| (2) Total Children Weighed | 9 |
| (3) Total Children Growth Faltering | 1 |
| (4) Total Children Severely Malnourished | 0 |
| (5) Total Children Referred | 0 |
| (6) Total Children in Feeding | 5 |
| (7) Total Child-Days Feeding | 101 |
| (8) Total Possible Feeding Days | 140 |
| (9) Percent Weighing Coverage [(2)/(1) x 100] | 90.0 |
| (10) Percent Growth Faltering [(3)/(2) x 100] | 11.1 |
| (11) Percent Severely Malnourished [(4)/(2) x 100] | 0.0 |
| (12) Percent Children Referred [(5)/(2) x 100] | 0.0 |
| (13) Percent Feeding Coverage [(7)/(8) x 100] | 72.1 |

Follow-up necessary for . . .
Percent Weighing Coverage < 60%
Percent Severe Malnourished > 15%
Percent Feeding Coverage < 75%

## Monthly Progress Report

**CNC No.: 03**
**Month/Year: June/97**

| | |
|---|---|
| (1) Total Children < 2 years | 10 |
| (2) Total Children Weighed | 10 |
| (3) Total Children Growth Faltering | 5 |
| (4) Total Children Severely Malnourished | 0 |
| (5) Total Children Referred | 0 |
| (6) Total Children in Feeding | 9 |
| (7) Total Child-Days Feeding | 270 |
| (8) Total Possible Feeding Days | 270 |
| (9) Percent Weighing Coverage [(2)/(1) x 100] | 100.0 |
| (10) Percent Growth Faltering [(3)/(2) x 100] | 50.0 |
| (11) Percent Severely Malnourished [(4)/(2) x 100] | 0.0 |
| (12) Percent Children Referred [(5)/(2) x 100] | 0.0 |
| (13) Percent Feeding Coverage [(7)/(8) x 100] | 100.0 |

Follow-up necessary for . . .
Percent Weighing Coverage < 60%
Percent Severe Malnourished > 15%
Percent Feeding Coverage < 75%

## Supervisor's Monthly Progress Report

**Sup. No.: 07**
**Month/Year: June/97**

| | |
|---|---|
| (1) Total Children < 2 years | 30 |
| (2) Total Children Weighed | 29 |
| (3) Total Children Growth Faltering | 10 |
| (4) Total Children Severely Malnourished | 4 |
| (5) Total Children Referred | 4 |
| (6) Total Children in Feeding | 19 |
| (7) Total Child-Days Feeding | 468 |
| (8) Total Possible Feeding Days | 530 |
| (9) Percent Weighing Coverage [(2)/(1) x 100] | 96.7 |
| (10) Percent Growth Faltering [(3)/(2) x 100] | 34.5 |
| (11) Percent Severely Malnourished [(4)/(2) x 100] | 13.8 |
| (12) Percent Children Referred [(5)/(2) x 100] | 13.8 |
| (13) Percent Feeding Coverage [(7)/(8) x 100] | 88.3 |

# Anthropometric Reference Data

## Males

| Age (month) | Median Weight (kg) | 60% of Median |
|---|---|---|
| 0 | 3.3 | 2.0 |
| 1 | 4.3 | 2.6 |
| 2 | 5.2 | 3.1 |
| 3 | 6 | 3.6 |
| 4 | 6.7 | 4.0 |
| 5 | 7.3 | 4.4 |
| 6 | 7.8 | 4.7 |
| 7 | 8.3 | 5.0 |
| 8 | 8.8 | 5.3 |
| 9 | 9.2 | 5.5 |
| 10 | 9.5 | 5.7 |
| 11 | 9.9 | 5.9 |
| 12 | 10.2 | 6.1 |
| 13 | 10.4 | 6.2 |
| 14 | 10.7 | 6.4 |
| 15 | 10.9 | 6.5 |
| 16 | 11.1 | 6.7 |
| 17 | 11.3 | 6.8 |
| 18 | 11.5 | 6.9 |
| 19 | 11.7 | 7.0 |
| 20 | 11.8 | 7.1 |
| 21 | 12 | 7.2 |
| 22 | 12.2 | 7.3 |
| 23 | 12.4 | 7.4 |
| 24 | 12.6 | 7.6 |

## Females

| Age (month) | Median Weight (kg) | 60% of Median |
|---|---|---|
| 0 | 3.2 | 1.9 |
| 1 | 4 | 2.4 |
| 2 | 4.7 | 2.8 |
| 3 | 5.4 | 3.2 |
| 4 | 6 | 3.6 |
| 5 | 6.7 | 4.0 |
| 6 | 7.2 | 4.3 |
| 7 | 7.7 | 4.6 |
| 8 | 8.2 | 4.9 |
| 9 | 8.6 | 5.2 |
| 10 | 8.9 | 5.3 |
| 11 | 9.2 | 5.5 |
| 12 | 9.5 | 5.7 |
| 13 | 9.8 | 5.9 |
| 14 | 10 | 6.0 |
| 15 | 10.2 | 6.1 |
| 16 | 10.4 | 6.2 |
| 17 | 10.6 | 6.4 |
| 18 | 10.8 | 6.5 |
| 19 | 11 | 6.6 |
| 20 | 11.2 | 6.7 |
| 21 | 11.4 | 6.8 |
| 22 | 11.5 | 6.9 |
| 23 | 11.7 | 7.0 |
| 24 | 11.9 | 7.1 |

***What are the problems?***

While problem-identification is an important aspect of data review, participants' suggestions for remedial action are at least as or even more important. The best steps to take will depend on the situation, the resources available, and the field workers' knowledge of the individual situations: there are no "right" answers for actions. The suggestions below highlight a few of the possible ways to respond to these data, but many other actions may be appropriate.

***CNC01***

In comparison with previous months, the number/percentage of severely malnourished children has suddenly increased in the June 1997.

Suggestion: CNO discusses with CNP possible reasons for this shift. Possible causes could be new diarrhea cases (check condition of tubewells in the area—a broken tubewell might be resulting in several families drinking pond water), or infection/disease (refer to health services). The CNO can advise the CNP to visit these homes in hopes of discovering a link/cause, or they can visit the homes together.

One child, ID 107, growth faltered for three consecutive months (Feb through April) and was severely malnourished in April. No referral is recorded in conjunction with the severely malnourished status in April. In May this child's growth did not falter, according to project criteria, but he was still severely malnourished. At the current month's weighing, this child's growth again did not falter, according to project criteria, but he remains severely malnourished (and received a referral). According to the May feeding record, this child's attendance at feeding was poor (note that the mother brings the child only two days each week, and always on the same two days of the week).

Suggestion: The CNP can visit the home and help the mother identify constraints on bringing the child to the feeding, focusing especially on the reason that the mother only can come two days per week. If family members are preventing the mother from bringing the child, the CNP can try to convince them of the need for feeding, or she can involve the Village Nutrition Management Committee in discussions with the family members. If the mother has other constraints (e.g., day labor), the CNP can suggest another family member who could bring the child.

Close inspection will show that the mothers' IDs, which are linked to household numbers and thus indicate location in the village, reveal that all the severely malnourished children in June live near to each other. This may strengthen suspicions that there is a single cause like a broken tubewell. Note that the high percentage of severes may not be apparent until after completion of the SMPR.

***CNC02***

Child 202 showed poor attendance at feeding (17/28 possible days) but showed amazing growth during the last month.

Suggestions: the CNO may wish to conduct a visit to this home along with the CNP and to re-weigh the child. If the child is severely malnourished, the CNO can review weighing procedures with the CNP and remind her of the importance of careful weighing so no child who needs help is overlooked. If the child's weight is correct, the CNO can review weighing procedures and emphasize the need to feed only those who are eligible in order to keep costs under control While in the home, the CNO and the CNP can discuss with the mother to determine reasons why she is not bringing the child to feeding, taking appropriate steps or giving appropriate counseling to remove any constraints.

One child, ID 206, was severely malnourished at the project outset in January, when he received a referral and, presumably, entered feeding. His nutritional status and growth trajectory improved while in feeding, but he faltered the next month (May 1997) after graduating from feeding. He did not attend the weighing session in June.

Suggestion: the CNP should regularly conducts home visits for any children who do not attend feeding. The CNO should discuss with the CNP to determine that a home visit has taken place (she may want to visit the home herself) and to find out what counseling the CNP provided to encourage the mother to bring this child to feeding. In addition, the CNP should intensify feeding advice for this mother since the child has relapses after having once undergone demonstrative feeding. If the family suffers from a severe food insecurity problem, the CNP may wish to involve the VNMC to seek ways that the village can help this family.

One child, ID 210, growth faltered in June 1997 at age six months. Close inspection also reveals that her monthly growth, which had been quite good, was not as good from her fourth to fifth month.

Suggestion: CNP discusses complementary feeding with the mother.

***CNC03***

There is a serious growth faltering problem in this CNC, but there doesn't seem to be any pattern to it and not a single child has even been recorded as severely malnourished. Close inspection will show that all weights are given in whole kilograms—perhaps even rounded up.

Suggestion: CNO reviews weighing technique with the CNP and check the scale at the CNC.

Two children have reached their second birthday and thus should no longer continue project participation.

Suggestion: CNO reviews eligibility criteria with CNP and investigates why children over 2 years of age are in the program.

***SMPR***

When aggregating the SMPR, it should be evident that some CNCs have individual problems (e.g., CNC01 has too many growth faltering and too many severely malnourished; CNC03 has a very high percentage growth faltering). These problems become easy to detect when comparing CNC records but are masked once the aggregate report is complete.

***How to Take a Simple Random Sample***

- Make a list of all elements (e.g., participants, beneficiaries or communities).
- Number the list.
- If the list contains 25 or fewer individuals/communities/elements:

  1a) Number a chit, or a small piece of paper for each item in the list. Each piece should be from the same kind and color of paper, and they all should be the same size.

  2a) Fold all the chits in half, put them into a bowl, or hat, and mix them well.

  3a) Select the number of chits to be included in the sample. For instance, if the sample will contain 5 individuals, select 5 chits.

  4a) Mark the list according to the numbers on the selected chits—these individuals constitute the sample.

- If the list contains 25 or more individuals/communities/elements:

  1b) Get someone to mark the list as the sample is selected.

  2b) Place a random number table in front of you.

  3b) Take a pencil in your hand, close your eyes and raise the pencil over the page, then drop the pencil to the page.

  4b) Open your eyes, and read to the right of the number the pencil is pointing to, taking as many digits as necessary to count to your entire list. That is, if there are between 10 and 99 items on your list, each number is a two digit number, so read two digits; if there are between 100 and 999 items, each number in the list is a 3 digit number, so read three digits.

  5b) Read out loud, in order going down the page, a random number for each element in your sample. The person assisting you

should mark the list at each number you call out, as this will be an element in your sample. If you reach the bottom of a column, move to the right according to the number of digits you are selecting and start reading downward again from the top of the column (i.e., if you are taking two digits, move two columns to the right and start from the top; if you are taking three digits, move three columns to the right). After moving to the right, the table may not contain enough digits for your list—for example, only 2 columns of digits may be left in the table and you need 3 for a list of 100 to 999 items. If this happens, start again at #3b.

***How to Take a Stratified Sample***

- Make a list of all eligible participants.
- Arrange the list according to the factor for stratification, from smallest values to highest values of the stratifying factor. The values chosen to group the stratifying factor must not overlap, and the levels must include all possible values of the stratifying factor in the population.
- Decide how many individuals will be taken from each stratum. As long as there is a good likelihood that individuals in the sample will, in fact, participate in the information gathering, it is easiest to take the same proportion of individuals from each stratum. There is no rule to decide this, but some guidelines include the following:
  - Sample enough persons from each stratum to have meaningful and representative estimates of means and standard deviations.
  - Consider the cost of using a particular proportion—taking a high proportion to get enough individuals from a small stratum will require taking a large and possibly expensive number of individuals from a large stratum if the same proportion is used for both.
- Use the methods for simple random sampling within each stratum to

obtain the actual sample.

***How to Take aSystematic Sample with a Random Start***

- Obtain an estimate of the total population size.
- Determine the necessary sample size (or the number represented by the sampling fraction), and divide it by the total population size. This will provide the sampling interval.
- Choose a random number between 1 and the sampling interval. That is, if the sampling interval is 5, choose a random number between 1 and 5.
- Select participants, starting with the random number, say 3, according to the sampling interval: 3, 3 + 5, 3 + 10, 3 + 15, 3+ 20, etc. until the sample size has been obtained.

# Annex 6: Instrumentation Checklist

1) _____ The instrument is concisely written (i.e., each piece of information requested is necessary).
2) _____ The language used in the instrument is appropriate for the people who will be providing the information.
3) _____ If the instrument has been translated, it has been retranslated into the original language as a quality control measure.
4) _____ The language used is specific in nature.
5) _____ There are no double barreled questions.
6) _____ Questions are worded neutrally in order to minimize response bias.
7) _____ There are no double negatives.
8) _____ The options given for closed-ended questions cover each of the major possible responses .
9) _____ Field-level forms ask for facts, not judgments.
10) _____ Calculations are not required on field-level forms.
11) _____ The instrument begins with “easy” questions before progressing into more sensitive or difficult topics.
12) _____ The layout is clear, easy to follow, and uncrowded.
13) _____ Questions are numbered.
14) _____ There is ample space given for responses.
15) _____ The instrument is clearly labeled.
16) _____ The instrument has been pre-coded for data entry and compilation (if possible).
17) _____ There are spaces for identification codes and the date.
18) _____ There is space at the end of the instrument for additional comments and observations.
19) _____ There is consistency between all forms being used.
20) _____ The instruments have been pre-tested and revised.

# Annex 7: Illustrative Use of the Conceptual Framework Model for Backwards Mapping

The first Tamil Nadu Integrated Nutrition Project (TINP I) which operated in South India from 1980 to 1990 provides a useful illustration of the use of the conceptual framework that was introduced in Step 1. Although the project did not explicitly use such mapping, it did employ an extensive M&E system, examples of which have been used in this guidebook.

TINP I, the largest nutrition project operating in a developing country at that time, was designed to provide nutrition services to young children and mothers using locally identified community nutrition workers chosen on the basis of being "successful" mothers. Monthly growth monitoring was employed and coupled with nutrition counseling and micronutrient provision. Children found to be severely malnourished or growth faltering were placed in a "special attention" category which included daily on-site feeding (justified primarily as an educational tool) plus counseling for caretakers. Children who failed to return to proper growth curves after ninety days were referred to health clinics. A similar system was to have been put in place for pregnant mothers.

Had a conceptual framework been developed for the project, it might have been used not only for M&E system design, but also as a guide for systematic problem identification. Once data were available on outcomes and impacts, "backward mapping", as illustrated in the paragraphs and framework below, could have helped pinpoint the weak links in need of attention.

Beginning at the right side of the framework, impact data collected by the project would indicate that while severe malnutrition among young children was being reduced beyond expectations, improvements in moderate malnutrition had been more elusive. This would be of particular concern given the fact that, universally, moderate malnutrition, which affects larger numbers of children and is more menacing than earlier suspected, contributes to a larger number of total deaths than severe malnutrition.

# Tamil Nadu Integrated Nutrition Project (TINP I)

Inputs → Outputs → Outcomes → Impacts → Benefits

| Inputs | Assumptions | Outputs | Assumptions | Outcomes | Impacts | Benefits |
|---|---|---|---|---|---|---|
| • Facilities and equipment for growth monitoring and screening | | • All children are weighed; all pregnant women are measured | • Mothers are able and motivated to bring all children to growth monitoring sessions; pregnant women are able and motivated to come to screenings | | | |
| • Village mapping | • All households are included on map | | | | | |
| • Development of messages and protocol for nutrition counseling | • Materials are produced and available on site | • Targeted households receive counseling | • Knowledge and attitudes regarding nutrition, health and hygiene change | • Nutrition, health and hygiene practices improve | • Reduced severe and moderate malnutrition | |
| • Food supplements and necessary equipment for daily feeding for at-risk children and at-risk mothers | • Supplements are acceptable, meet quality standards, and are available at feeding sites in adequate quantity | • Severely malnourished and growth faltering children consume food on a daily basis; targeted at-risk pregnant women consume food on a daily basis | • Mothers are able and motivated to bring all children to daily feeding; pregnant women are able and motivated to come to daily feeding | • Total caloric intake for “at risk” pregnant women increases | • Reduced incidence of low birth weight infants | • Reduced morbidity and mortality |
| • Micronutrients | • Micronutrient supplements are available at distribution points | • Mothers and children receive micronutrients | | | • Reduced micronutrient deficiencies | • Improved learning and productivity |
| • Development of referral system | | • Difficult cases are referred to the health center | • Health centers are equipped to accept referrals | | | |
| • Staff identification and training | • Community Nutrition Workers (CNWs) are motivated and knowledgeable | | | | | |

Moving to the left on the conceptual framework in search of an explanation, one would note that while nutrition counseling (an output) was systematically provided, no information was collected on changes in knowledge and attitudes (output assumptions) or on the extent to which practices had actually changed (an outcome). Since behavioral change was originally envisaged as an important means of addressing moderate malnutrition, this now recognized information gap might have been addressed by a special study.

If such a special study found that critical practices had not changed, this might also explain the high relapse rate (an outcome) found by the M&E system. The absence of behavioral change would require, moving again to the left on the framework, an assessment of knowledge and attitudes (output assumptions), a re-examination of counseling (an output), and perhaps, finally, an assessment of message development (an input).

The “mystery of the moderates” might also have been addressed by considering the criteria for selection of children receiving daily feeding and special attention. This review, in fact, did lead to a decision to expand the criteria to include the moderately malnourished in TINP II.

Moving back again to the right side of the conceptual framework, one would discover that little data was collected on the prevalence of low birth weight infants (an impact), or on the total caloric intake of mothers (an outcome). In addition, evaluations revealed that few mothers actually received food supplements (an output) and that actual screening of pregnant women was rare (an output). Further investigation would reveal a serious administrative problem. While the village-based Community Nutrition Workers responsible for child growth monitoring and feeding were administered by the state’s Social Welfare Department, responsibility for maternal screening resided with the health center-based Multi Purpose Health Workers under the Department of Health. Coordination between the two departments was inadequate. These problems were addressed in the design of TINP II which included systematic coordination between the two departments, joint directives, and some common training.

Finally, looking at the micronutrient component of the project at the bottom of the map, one would discover that while careful records were maintained on the delivery of iron and Vitamin A (outputs), there was, important in the case of iron, no information collected on the extent of compliance in taking the pills daily or, in turn, on the effect of the supplementation on anemia.

# References



Anderson M.A. 1977. *Care Preschool Nutrition Project, Phase II Report* New York: Care.

Beaton G. and Ghassemi H. 1982. *Supplementary Feeding Programs for Young Children in Developing Countries* AJCN Vol. 35 (4) (suppl) 864–916.

Converse J.M. and Presser S. 1986. *Survey Questions: Handcrafting the Standardized Questionnaire*. Newbury Park: Sage Publications.

Dean A.G., Dean J.A., Coulombier D., Brendel K.A., Smith D.C., Burton A.H., Dicker R.C., Sullivan K., Fagan R.F., Arner T.G. 1995. *EpiInfo, Version 6: A Word-Processing, Database, and Statistics program for Public Health on IBM-compatible computers.* Centers for Disease Control and Prevention, Atlanta, GA, USA.

Florio D.H., Behrmann M.M., and Goltz D.L. 1979. "What Do Policy Makers Think of Evaluational Research and Evaluation? Or Do They?" *Educational Evaluation and Policy Analysis*. 1(6): 61–87.

Franklin J.L. and Thrasher J.H. 1976. *An Introduction to Program Evaluation*. New York: John Wiley & Sons.

Hamilton D. and Gaertner U. 1992. *Goal Oriented Project Planning (GOPP): An Introduction to the Methodology*, GTZ-RPMAS, TG-PMC, UNDP/DTCP.

Heaver R. 1991. *Participatory Rural Appraisal: Potential Applications to Family Planning, Health and Nutrition Programs.* Washington DC: The World Bank.

Hennekens C.H., Buring J.F., and Mayrent S.L. (Eds.). 1987. *Epidemiology in Medicine*. Boston: Little, Brown and Company.

Ho T.J. 1985. *Economic Issues in Assessing Nutrition Projects: Costs, Affordability and Cost Effectiveness.* Washington DC: The World Bank (PHN).

Levinson F.J., Dladla C., McLachlan M., and Olivar M. 1995. *Nutrition Program Constraints Assessment: Gauteng Province, South Africa.* Tufts University: Medford, MA.

Levy S. and Lemeshow S. 1991. *Sampling of Populations: Methods and Applications*. New York: John Wiley & Sons, Inc..

MacGregor A.J. 1979. *Graphics Simplified: How to Plan and Prepare Effective Charts, Graphs, Illustrations, and Other Visual Aids.* Toronto: University of Toronto Press.

Musgrove P. 1991. *Feeding Latin America's Children: An Analytical Survey of Food Programs.* Washington, DC: The World Bank. (Latin American and the Caribbean Regional Studies Report #11).

Newman J., Rawlings L., and Gertler P. 1994. Using Randomized Control Designs in Evaluating Social Sector Programs in Developing Countries. The World Bank Research Observer, Vol 9, no. 2, pp. 181–201. Washington, DC: The World Bank.

Phillips M. and Sanghvi T. 1996. The Economic Analysis of Nutrition Programs. Nutrition Toolkit #3. Washington, DC: The World Bank, Human Development Department.

Phillips M., Sanghvi T., Suarez R., et al. 1996. *The Cost-effectiveness of three vitamin A interventions in Guatemala.* Soc Sci Med: Vol. 42 No. 12, 1661–68.

Rosen D.S., Haselow, N.J., and Sloan N.L. 1993. *How to Use the HKI Food Frequency Method to Assess Community Risk of Vitamin A Deficiency*. New York: Helen Keller International, Incorporated.

Rossi P.H. and Freeman H.E. 1993. *Evaluation: A Systematic Approach 5*. Newbury Park: Sage Publications.

Rothman J. 1980. *Using Research In Organizations: A Guide to Successful Application*. Sage Publications: London.

Santo Pietro D. (Ed.). 1986. *Evaluation Sourcebook for Private and Voluntary Organizations*. New York: American Council of Voluntary Agencies for Foreign Service.

Valadez J. and Bamberger M. (Eds.). 1994. *Monitoring and Evaluating Social Programs in Developing Countries.* Washington, DC: The World Bank Economic Development Institute.

World Bank. 1980. *Tamil Nadu Nutrition Project Implementation Volume*. Washington, DC: The World Bank, Population, Health and Nutrition Department.

Zeitlin, M.F. , et. al. 1984. *Vol. IV, Household Evaluation, Nutrition Communication, and Behavioral Change Component, Indonesian Nutrition Development Program.* Washington, DC: Report to the World Bank.

Zeitlyn, S. No date. *Feeding Practices in Bangladesh with special reference to pregnant, postpartum and lactating women and infants and children: A review of the literature.* Dhaka, Bangladesh: UNICEF.