human systems management

Volume 4, No. 2, 1983

The study of community-level nutrition interventions: an argument for reflection-in-action

William D. DRAKE
University of Michigan, Ann Arbor, MI 48109, U.S.A.

Roy I. MILLER
Community Systems Foundation, Ann Arbor, MI, U.S.A.

Donald A. SCHON
Massachusetts Institute of Technology, Cambridge, MA, U.S.A.

pp. 82–97
The study of community-level nutrition interventions: an argument for reflection-in-action

William D. DRAKE
University of Michigan, Ann Arbor, MI 48109, U.S.A.

Roy I. MILLER
Community Systems Foundation, Ann Arbor, MI, U.S.A.

Donald A. SCHON
Massachusetts Institute of Technology, Cambridge, MA, U.S.A.

After analyzing longitudinal anthropometric data from eight community-level nutrition programs to determine their impact, it was concluded that conventional approaches to analysis do not eliminate indeterminacy because: (1) the data was inaccurate or inconsistent, (2) the measures or measurement methods produced misleading results, and, most importantly, (3) a lack of information about the local context of the interventions precluded the elimination of competing explanations of observed outcomes. In that analysis, as with most similar analyses, the traditional experimental approach (applying a predesigned experiment using controls in a presumably constant environment) failed because the experimental context was unstable, unpredictable, and unique in each case. Furthermore, the instability, unpredictability, and uniqueness of each case called for a flexible intervention strategy to cope with the changing context.

As an alternative approach to both analysis and intervention, reflection-in-action is proposed. Six features of this model are: explicit specification of the framework underlying the intervention strategy; continuous monitoring of both data gathering procedures and intervention strategies; periodic redesign of those procedures and strategies; collaboration between researchers, practitioners, and subjects throughout; use of on-the-spot experimentation to test particular hypotheses; and explicit enumeration and accounting for potential factors confounding both the analysis and the intervention itself. By actively using the data for continuous monitoring, field practitioners, working with analytic specialists, are more likely to reduce or eliminate indeterminacy due to inaccurate data and/or contextual changes than would traditional researchers or evaluators who maintain distance between themselves and the intervention.

Reflection-in-action, in part, is illustrated in the context of a recent evaluation conducted in Sri Lanka where a revisit to the field with preliminary quantitative results caused modification in the interpretation of those results. Problems remain, however, in achieving full implementation of this approach. Practitioners and scientists will have to change their attitudes and behavior to accommodate R-I-A, the role of quantitative analysis in program management and evaluation will have to be placed in proper perspective, and institutions supporting intervention activities will have to modify their approach to both funding and evaluation.

Keywords: Intervention, nutrition, program evaluation, reflection-in-action, program management, experimental design, planning.

William D. Drake has worked as a practitioner and researcher on issues surrounding planning and implementation. He chaired the Ann Arbor Metropolitan Transportation Authority during its formative years, and was instrumental in initiating an interdisciplinary Ph.D. Program in Urban and Regional Planning at the University of Michigan. Currently he is President of the Board of Directors of a Savings and Loan Association and President of Community Systems Foundation. He is co-author of a seven-volume work Analysis of Community-Level Nutrition Programs.

Roy I. Miller is a Research Associate of Community Systems Foundation in Ann Arbor, Michigan. For the past five years, Dr. Miller has worked on developing techniques for monitoring and evaluating nutrition interventions. As an evaluation specialist, he has worked with individuals administering food aid programs in Honduras, India, and Sri Lanka. He is co-author of the seven-volume work Final Report: Analysis of Community-Level Nutrition Programs.

Donald A. Schon. As an industrial consultant, government administrator, and a president of OSU (a non-profit social research and consulting organization), Dr. Schon has worked as researcher and practitioner on the problem of organizational learning and professional effectiveness. He was invited in 1970 to deliver the Keith Lecture on the BBC. His books include: Invention and Evolution of Ideas (formerly, The Displacement of Concepts), 1963: Technology and Change.
1. Findings from an analysis of community-level nutrition programs

In the analysis of social intervention, three questions are of paramount importance:

1. Is a particular intervention having the desired effect.
2. How can any given intervention be improved in design and/or implementation, and
3. Can the lessons of a single attempt at intervention be translated into better program designs in the future?

In short, can we learn from ongoing social interventions?

Attempting to answer these questions, we analyzed a large sample of community-level nutrition interventions. We considered the interventions as experiments aimed at improving the nutritional status of their target populations, often preschool children between the ages of zero and six. Employing conventional measures of malnourishment, we attempted first to determine whether a change in nutritional status had occurred during the project by comparing measurements of nutritional status in the target population before, during, and after the period of intervention. Second, where a change in nutritional status could be detected, we attempted to attribute it to the intervention. This we did by formulating several plausible explanations of the change, seeking then to discriminate among these explanations on the basis of available data. In short, we attempted to apply a widely accepted model of interventions as social experiments to the particular domain of community-level nutrition intervention.

Of the roughly 25 projects considered initially, only eight approximated the minimal criteria necessary for treatment of them as experimental interventions. One criterion which was especially deci-

1 Nutrition interventions are of many types. Some focus on specific nutritional deficiencies with known treatments; for example, mass pill distribution to eliminate widespread vitamin A deficiency. Others focus on poor nutrition in general, a problem requiring a more diversified multi-disciplinary attack. Our work dealt with this latter type.

sive was the availability of longitudinal data on nutritional status.

In only one of the eight cases did those responsible for the intervention view their effort as an experiment in the traditional sense. (By traditional sense, we mean as a test of a specific hypothesis through a careful application of treatment according to a pre-defined experimental design using appropriate controls.) In all other cases, the activities were first and foremost efforts to provide help to those in need. Learning was always important to the change agents but, except for the one case, the primary motivation behind each intervention design was to maximize program impact, rather than to facilitate learning. Thus the interventions reviewed feature diverse philosophies and varied commitment to the collection and analysis of quantitative data to promote learning.

Because we refer repeatedly to the eight interventions by name in the following text, we provide a brief description of each here.

1. Candelaria - A program featuring home visitations by volunteer ‘barefoot doctors’ in the town of Candelaria, Colombia. Services include: education on nutrition, hygiene, and the utilization of health services; weighing of children; and referrals to the medical establishment.

2. Candelaria Revisited - A resurvey of the village of Candelaria two years after termination of the program.

3. Primops - An extension of the Candelaria concept to an urban setting in Cali, Colombia.

4. Esperanca - A program based on the establishment of health posts in rural villages in the Central Amazon region of Brazil. Services include: the provision of basic health care by a visiting doctor and/or a locally trained ‘barefoot doctor’; nutrition; and health education; and child weighings.

5. SCF-Honduras - A program following the Save The Children community development scheme, CBIRD, carried out in the Peshire region of Honduras. Food supplements were administered in parallel with more general development efforts.

6. SCF-Indonesia - Another Save The Children program, this time in the special province of Aceh in northern Sumatra.

7. Thailand - An experiment to test the effect of rice fortification on nutrition and health in the Chiang Mai region of Thailand.
(8) Kottar – A community development oriented program run by the Kottar Social Services Society in Tamil Nadu, India. Services include food supplementation; nutrition, and health education; immunizations and more general community development assistance.

The results of our analysis can be summarized as follows:

1. In only two of the projects was it possible to establish, with reliability, that a positive change in nutritional status had taken place in the target population (that is, a change clearly free of noise in the data, and independent of changes in the make-up or aging of the target group).

2. In none of these cases was it possible to attribute change unambiguously to the intervention; on the basis of available data, we could not discriminate among some plausible competing explanations.

3. Although the projects analyzed varied greatly concerning the extent to which they were designed and carried out as rigorous field experiments (according to a model which we call ‘the model of rigorous experiment’), the indeterminacy of the data was as great in the more rigorously designed as in the less rigorously designed projects.

We were able to identify three main sources of indeterminacy in the data:

1. Practices of data-collection, storage, and processing which produced inaccurate or inconsistent data. (We shall call this ‘dirty data’.)

2. Imperfect measures and measurement methods which lead to misinterpretation of analytic results (We shall call this ‘inadequate measurement’.)

3. A lack of timely, situation-specific data which would have permitted discrimination among competing explanations of change in nutritional status, given settings that were in many respects complex, unique and rapidly changing. (We shall call this ‘inadequate data’.)

We found that it was possible, by adopting certain techniques of data cleaning, processing, and analysis, to reduce some of these sources of indeterminacy. By performing tests with regard not only to the range assumed by relevant variables but also their internal consistency within and between observations in a longitudinal sequence, we were able to eliminate many, if not most, mechanical errors without discarding valuable data. By exploring a variety of analytic techniques to test alternative theoretical propositions, we could quantify the effect of some confounding variables, eliminating some of the sources of indeterminacy. In this way, we were able to identify reliable patterns of change in two projects, and were able to eliminate some contending explanations of change in some others. But in other cases, and for some plausible competing explanations in all cases, it was not possible to arrive at an unambiguous explanation of the data.

For example, in the analysis of the data from Kottar, we found obvious keypunching errors in 40% of the anthropometric data. Once we cleaned the data, we were able to convince ourselves that measurable, positive, nutritional change had occurred only after experimenting with a locally generated growth standard for Indian children. The recovery was not as rapid as that called for by international standards but was apparent when viewed with a less ambitious expectation as set by a local standard. We were able to eliminate the possibility that the observed change was due to a bias in the selection of new participants over time but we were unable to refute the possibility that the change was induced, in part, by the reduction of rice prices as a highly localized drought ended.

Reflection on this experience strongly suggests lessons that are equally important for two objectives – the conduct of informative community-level nutrition experiments, and the conduct of effective community-level nutrition interventions. And further, our experience leads us to question what we take to be the prevailing model of the proper relation of experimental research to the practice of community-level intervention.

2. The model of rigorous experiment and its shortcomings

The most common approach to learning from community-level nutrition interventions is based on techniques which have proven successful in several of the natural sciences – an approach we call the model of rigorous experiment. An essential element of the model of rigorous experiment is good experimental design. As a discipline within the field of statistics, experimental design grew out of the need to sort out the variations in certain variables due to some specific action or event from the variations which normally occur in those varia-
bles. As Cox explains in his introduction to the subject,

...as soon as the effects under investigation become comparable with the uncontrolled variations, the problems we shall be concerned with become important. [3, p.3]

As originally conceived, statistically efficient experimental design becomes possible when the researcher can control the assignment of experimental units (the recipients of some sort of treatment; for example, preschool children, mice, or plots of land) to different treatment groups. By appropriate randomization in such assignments, it becomes possible to eliminate most, if not all, of the potential explanations for changes unique to the treated group other than the desired explanation that the treatment itself caused the changes.

As social scientists began to grapple with situations calling for experimental verification of hypotheses that were not conducive to the neat random assignment procedures available in the laboratory, the concepts of experimental design were extended to include quasi-experimental designs. These are more complex than the true experimental designs and less able to eliminate all competing explanations. Campbell and Stanley, whose work has contributed most to the development of quasi-experimental research in the social sciences write:

Insofar as the designs discussed in the present chapter become complex, it is because of the intransigency of the environment: because, that is, of the experimenter's lack of complete control. [2, p.1]

In their landmark publication, Campbell and Stanley itemize eight threats to internal validity of quasi-experimental designs and another four threats to external validity. Internal validity pertains to the question of whether the treatments considered in a particular experiment truly make a difference in that experiment. External validity pertains to the question of generalizability; that is, to the question of the degree to which the observed results carry over to other settings. Their discussion shows that even the most complex quasi-experimental designs are susceptible to challenge from, at least, some of these threats to validity.

In the nutrition field, the response thus far to difficulties encountered in applying quasi-experimental design concepts in research and/or evaluation research nutrition interventions has been to try for better controls or more complicated experimental designs. Habicht and Butz write:

Therefore, any experimental design which does not randomly distribute the intervention and its controls within a village or a region must have sufficient villages or regions covered by each treatment (replicates) so that one can estimate the probable contribution of non-specific influences at the village or regional level. Adjacent villages and regions must have different treatments, and the villages and regions should be so stratified that any other random non-specific influences are controlled for. Designs which show differences between regions or villages but do not have these required replicates must remain suspect. [6, p.150]

We believe that it is not possible to remove several sources of indeterminacy (eliminate the threats to validity), of the sorts which we have encountered, by the more exacting application of the model of rigorous experiment and analysis. To help explain our position, let us consider four characteristics of controlled experiments.

(1) Control and experimental groups must be matched: A control group must be established which is like the experimental group in all respects except for the presence of the intervention. Otherwise, outcomes (observed results) could be attributed with equal plausibility to the inherent differences in the groups.

(2) Control and experimental groups must experience identical change: During the experiment, both the control and treatment groups must be insulated from change unique to either group. Otherwise, outcomes could be attributed with equal plausibility to events occurring in one group but not the other.

(3) Pre-design of experimental conditions: In practice, where it is difficult to define and maintain such rigorous controls, it is possible to apply assorted multivariate statistical techniques to account for known differences between the groups. The use of statistical techniques to account for factors which confound experimental modes is often called the statistical control model [7, p.393]. To implement such a model, it is necessary to identify the potential confounding factors before the experiment is actually initiated. Otherwise, the data required for implementation of the statistical control model would be lacking.

(4) Separation of practice from research: Finally, those who deliver the intervention and collect the data should not be a party to the research design
and analysis. Otherwise, they are likely to bias their observations and/or become an integral, but unreplicable, part of the intervention. (We call this the principle of ‘research distance’.)

In practice, however, experiments with these characteristics cannot be attained reliably in the actual environment of community-level interventions. With respect to interventions in the nutrition arena, this is especially so because of the long time-frame assumed by most experiments and the high degree of interaction between nutrition in a community and the broader economic, social, and political context. Each experimental situation is likely to be **unique**; therefore, it is difficult to identify relevant factors for matching groups. Each situation is also likely to be **unstable**; therefore, it is difficult to avoid differential change in the groups over time. And, because such changes are **unpredictable**, it is virtually impossible to predefine the experimental conditions to account for those changes.

Paradoxically, because of the separation of practice from research (to ensure objectivity and guarantee replicability), experimentalists often fail to recognize the degree to which the uniqueness, instability and unpredictability of context distort analytic results. And, when they are aware of such distortion, they are unable to eliminate the resulting indeterminacy from their conclusions. This suggests the need for reliance on **local information systems** to sort out the ambiguities arising from the breakdown of controls. After all, those most likely to be able to detect confounding phenomena, and to offer plausible interpretations of ambiguous data, are precisely those closest to and most knowledgeable about the experimental situation—the practitioners and the subjects of intervention.

3. The conditions of effective community-level intervention

The factors which prevent effective implementation of the model of rigorous experiment also constrain the design and delivery of effective community-level interventions. That is, the factors which inhibit informative experimental research also inhibit effective practice.

(a) Because each community situation is unique, it is impossible to predict ahead of time what type of intervention is likely to be effective in achieving a positive change in nutritional status. The features of effective interventions must be discovered in and for the particular situation. That is, the design of effective intervention must be arrived at through situation-specific inquiry; it cannot be taken with reliability, ‘off the shelf’ (although it may have features of intervention-types familiar through earlier practice).

(b) Because the community situation is likely to be unstable, no intervention—however effective it may be initially—is likely to be effective for very long. If an intervention is to be effective over the long term, it must be continually revised in response to detection of changes in the situation including changes triggered by the intervention itself.

(c) In the design and redesign of effective interventions (as in the design and redesign of informative experiments), much will depend on the site-specific knowledge of practitioners and subjects who are closest to the data and most familiar with the context. It is they who will be in the best position to suggest explanations of surprising outcomes, detect unanticipated changes and identify (and correct) sources of error in data-collection.

In short, effective community-level nutrition interventions are likely to be ones in which intervention is treated as experimental inquiry—but experimental inquiry of a particular sort. One might describe this sort of inquiry in several ways, depending on the features chosen for attention. It might be called ‘contextual inquiry’, if we focus on the fact that inquiry must address itself to the unique characteristics of the particular site. We might call it ‘on-the-spot experiment’, if we focus on conduct of derived experiments in response to sources of indeterminacy discovered in the course of inquiry. We might call it ‘rapid information feedback’, if we focus on the on-line use of data to inform the redesign of measures, data-gathering procedures, and interventions. The term, ‘reflection-in-action (R-I-A)’, calls attention to the fact that inquirers must continually re-examine, in the midst of action, their framing of the problem of malnutrition, their models of the phenomena associated with malnutrition, and their strategies of intervention. R-I-A names the comprehensive process by which inquirers respond to the detection of surprising outcomes by surfacing, criticizing, restructuring and testing the context-specific frames, theories and strategies which they
have brought to the situation.

But R-I-A refers also to the model of field research which answers best to the four constraints on rigorous experiment.

4. Community-level nutrition: intervention as reflection-in-action

Given the weakness of the model of rigorous experiment in social settings and given also the conditions of effective community-level intervention, we conclude:

(1) in community situations in developing countries, instability, uniqueness and unpredictability prevent application of the model of rigorous experiment and call for reflection-in-action;

(2) in community situations of this kind, reflection-in-action is equally appropriate as a model of experimental research and as a model of effective intervention; and

(3) given a model of intervention and experiment as reflection-in-action, it is necessary to revise the prevailing view of the proper relation of research to practice from one advocating separation of the two to one urging their mutual support.

This third conclusion most strongly distinguishes R-I-A from more conventional approaches to research. In most conventional approaches, the context of experimental research is taken to be distinct from the context of intervention. The function of experimental research is taken to be the production of reliable general propositions as guides to future intervention. Finally, in order to insure objectivity, researchers must remain distant from practitioners and subjects of intervention. We propose, on the contrary, that intervention-oriented research can be carried out effectively only through actual intervention in particular communities, that such research cannot produce reliable generalizations strictly applicable to types of community situations in which it is undertaken, and that practitioners and subjects of intervention ought to be involved in the practice of intervention. Practice should be carried out by practitioner-researchers and research-practitioners.

Following R-I-A, a community-level nutrition intervention would take on the following distinctive features - each being a divergence from more conventional approaches to intervention research and practice.

(1) Explicit specification of the framework underlying the intervention strategy: Initially, description of the situation (context), framing of the problem(s) of malnourishment in that situation, and design of an initial strategy of intervention would be undertaken. This would be viewed as only a first step in the generation of experimental inquiry. One might call the description, frame and design a 'model of the local nutrition system'. This model serves not only as a basis for initiating intervention but also as the yardstick with which progress can be measured and the best guess at the factors most likely to interact with and confound the intervention strategy.

(2) Continuous monitoring of both data gathering procedures and intervention strategies: There would be continuous monitoring of both the data gathering procedures and the trends revealed in the data through analysis. This would facilitate the redesign of field procedures, measures and measurement methods and facilitate clarification of the meaning of trends in nutritional status (and other indicators of success) through discrimination among competing explanations of those trends.

Data gathering procedures are notoriously inconsistent in the field. Sloppiness in those procedures can cause far larger changes in impact indicators than would be expected from any intervention. Monitoring data gathering procedures would make it possible to minimize the disturbance due to faulty data gathering. Similarly, early identification of trends in the data makes it possible to recognize alternative causes for those trends soon enough to account for them or, at least, begin to measure them.

(3) Periodic redesign of data gathering procedures and intervention strategies: As a result of the periodic review of the data analysis, data gathering procedures (including what data is to be gathered) would be redesigned, the intervention itself might be altered, and the experimental conditions (e.g. controls) might be redefined. In other words, R-I-A attempts to cope with change by reacting to it rather than trying to eliminate it.
(4) Collaboration between researchers, practitioners and subjects throughout. Once again, we emphasize the need to involve all parties to an intervention in the continuous redesign process. This is likely to serve as a feasibility check for all proposed changes, as an incentive for rapid adoption of those changes and, most importantly, a review of the accuracy of the interpretation of the analysis leading to change.

(5) Application of disaggregation techniques and other forms of derived experimentation. Whereas conventional analysis in the area of nutrition intervention has stressed the testing of hypotheses regarding change in large groups, R-I-A would encourage researchers to discriminate among competing explanations of change by using strategies of disaggregation. In describing his own work with behavioral testing in the field of psychology, Skinner makes an eloquent plea for targeting on relevant variables through thorough examination of the response of fewer subjects to treatment rather than more.

Suppose that measurements have been made on two groups of subjects differing in some detail of experimental treatment. Means and standard deviations for the two groups are determined, and any difference due to treatment is evaluated. If the difference is in the expected direction but is not statistically significant the almost universal recommendation would be to study larger groups. But our experience with practical control suggests that we may reduce the troublesome variability by changing the conditions of the experiment. By discovering, elaborating and fully exploiting every relevant variable, we may eliminate in advance of measurement the individual differences which obscure the difference under analysis. This will achieve the same result as increasing the size of groups, and it will almost certainly yield a bonus in the discovery of new variables which would not have been identified in the statistical treatment. [9, p. 91]

By disaggregating to account for confounding factors, the experimenter, in effect, is designing a derived experiment which differentiates among parts of the target population and parts of the experimental context. For example, if conventional statistical analysis shows little or no change in nutritional status in response to food supple-

5. Experimental norms

Table 1 compares the models of rigorous experiment and R-I-A with respect to their objectives and methods for each of the experimental norms of validity, effectiveness, objectivity, and generalizability. The two models aim at very different sets of objectives. Most fundamentally, the model of rigorous experiment aims at valid general propositions applicable to the design of nutrition interventions in types of community settings. (For example, ‘supplementary feeding programs should be implemented in communities of type “x”, via nurse practitioners.’) On the other hand, R-I-A aims at valid propositions applicable only to the experimental community. (For example, ‘in Candelaria, given the residual effects of earlier sewer and water programs, interventions undertaken in the next three to five years will have to be associated with a reduction of more than 20% in the malnourishment rate in order to be judged significantly effective.’)

In the model of rigorous experiment, the effort is to achieve results that are objective in the sense that effects of intervention are kept independent of persons carrying out the intervention, or of reactions of the subjects to the mere act of experimentation. In R-I-A, where results are situation-
The models of experiment compared

<table>
<thead>
<tr>
<th>Norm</th>
<th>Model of rigorous experiment</th>
<th>Reflection-in-action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity</td>
<td>Uses pre- and post-intervention observations; seeks to apply Mill’s methods rigorously by the</td>
<td>Close monitoring of processes as well as of conditions pre- and post-intervention. Rapid feedback of information to redefine experimental conditions.</td>
</tr>
<tr>
<td></td>
<td>use of control groups, insulation of experimental and control groups, constancy of experimental</td>
<td>Draws on those close to the scene to detect confounding factors, offer plausible explanations. Uses derived on-the-spot experiments to discriminate among competing explanations.</td>
</tr>
<tr>
<td></td>
<td>conditions</td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Criteria and strategy of intervention based on pre-intervention diagnosis of community situation</td>
<td>Criteria and strategy revised iteratively, in response to rapid feedback of information re outcomes, conditions, processes.</td>
</tr>
<tr>
<td>Objectivity</td>
<td>Sought through distancing of researchers from practitioners and subjects</td>
<td>‘Hawthorne effect’ expected, exploited. Effort not to eliminate experimenter bias but to detect what it is, and take it into account in interpretation of results.</td>
</tr>
<tr>
<td>Generalizability</td>
<td>Aims at general intervention-oriented results by using logic of experimental design to interrelate general features of context, intervention, controls, outcomes</td>
<td>Experimental results always context-specific, generalizable only as elements of repertoire, themes that can inform invention of problem, theory or strategy in next case.</td>
</tr>
</tbody>
</table>

Specific, there is no need to separate intervention from the personalities of the change agent or from the act of experimentation. But for the sake of future learning from experiment, it is important to monitor and interpret the ways in which features of person and experimentation influence behavior in the community.

In the model of rigorous experiment, the criteria of effective intervention must be designed, prior to experiment, in ways generally applicable to a type of community situation. In R-I-A, both the criteria and strategies of intervention are designed and redesigned, iteratively, in ways that take account of the uniqueness of the particular experimental community.

In the model of rigorous experiment, experimental validity is sought through pre- and post-intervention measurements, coupled with the use of control groups and insulating mechanisms to account for eliminate confounding factors. In R-I-A, validity of experimental interpretations is sought through close observation of the processes which connect interventions, in the context of other factors, to outcomes; and it is sought through the performance of on-the-spot experiments, designed to discriminate among competing explanations of outcomes as these explanations are generated in the course of experiment.

The more specific meanings of the norms and methods of R-I-A will become clearer as we examine proposed responses to the sources of indeterminacy we have found in analyzing the data gathered from our sample of community-level projects.

5.1. ‘Dirty data’

We have named errors in the data which arise due to faulty practices of data-collection, storage and processing, ‘dirty data’. In our work, the most humorous example of ‘dirty data’ comes from the Primops intervention where a tropical bug invaded the computer cards used for data storage and ate machine readable holes into those cards. Our experience suggests that most instances of ‘dirty data’ are connected to the distance, in time and space, between the users of the data and those who collect it; the greater the distance, the dirtier the data.

R-I-A by definition reduces the distance between the source of the data and its users. The rapid feedback component of R-I-A generates the opportunity to locate errors early and to inform data-collectors and/or processors of those errors. Specific errors can be corrected and, if procedures are found to be consistently inadequate, they, too,
can be modified. Moreover, project staff are more likely to be motivated to correct errors if they can see the usefulness of the data that they collect.

5.2. Factors which confound estimates of change in nutritional status

In the nutrition field, there is a genuine concern that the methods for measuring the nutritional status of both individuals and communities are inadequate. These inadequacies tend to confound estimates of change in nutritional status of a population over time. Let us consider three aspects of inadequate measurement.

1) Misclassifications can occur where measures of size and health are not coincident; for example, it has been shown that a child who has suffered chronic malnourishment, particularly during the age periods of the human growth spurt, is incapable of exhibiting the catch-up growth necessary to regain normalcy as defined by a growth standard derived from healthy children. Also, it has been argued that improved nutritional intake may be applied to functions other than growth such as to metabolic functions or increased levels of activity.

2) The classification of a child with regard to nutritional status may vary as a function of the anthropometric standard used to define normalcy. In particular, international standards derived from healthy populations in the developed world often prescribe more rapid growth than do locally generated standards. Thus, a population may appear to improve relative to the less ambitious local standard while remaining stagnant or even worsening relative to an international standard. In our own analysis, we showed that only a percentage of children suffering from second or third degree malnourishment over a two and one-half year period from 50% to 42.5% using a local standard but, using an international, sex-differentiated standard, we showed a drop from 50.4% to 45.5%.

3) Related to the choice of standard, the choice of "cut-points" for defining malnutrition can also alter the picture of change in nutritional status in a community. The selection of a level of deprivation relative to a standard for defining malnourishment is rather arbitrary. Because of the tendency for the "scores" of a population in a community to cluster about the traditionally accepted cut-off points, small shifts in these points lead to rather large numbers of children crossing the border line between categories of malnutrition.

In addition to indeterminacy from inadequate measurement, we note two other factors which confound estimates of the change in nutritional status.

4) The analysis of nutritional change can produce different results depending on the statistical methods used and on the variables included in the analysis. The assertion that results depend on the choice of statistical methods is surely nothing new; however, in the nutrition field, that choice takes on special significance. The concept of improved nutrition in a community cannot be defined precisely. Some argue that mean percent of standard is an adequate measure of community nutritional well-being while others emphasize the need to show the greatest improvement among those initially worst off. Often, the selection of a statistical methodology is equivalent to the selection of a definition of improvement. In these cases, the statistics govern the description of change in nutritional status.

Similarly, the selection of variables to monitor change is a proxy for defining improvement. The most common anthropometric "ratios" - weight for age, height for age, and weight for height - measure different aspects of malnutrition. The latter measures acute malnutrition; the second, chronic undernutrition; and the first, a composite of the two. Often, as one would expect, as height for age scores increase, weight for height scores decrease.

5) Finally, estimates of change in nutritional status are confounded by the method of accounting for the relationship between malnourishment and life cycle. The typical child in the developing world experiences a gradual deterioration of nutritional status from birth to some age of "maximum" risk (usually between 18 and 24 months) and then improves. Because participants in an intervention program improve, this pattern of growth must be acknowledged during the analysis. We have recommended that
nutrition status changes be mapped by using a device called a 'characteristic curve' – a graph of malnourishment against age at a single point in time [4, p.97]. However, even with a characteristic curve, the picture of change can be altered by redefining age categories or selection of different cut-points or standards.

None of these factors are easily remedied by using either the model of rigorous experiment or R-I-A. The use of anthropometrics to determine the severity of malnutrition in individuals will remain inadequate; however, research into the nature of the response of each anthropometric score to 'proven' intervention might establish precedents for assessing responses observed in field setting. Also, research into the use of other tests for malnourishment, such as blood serum tests, may enable such tests to supplement the use of anthropometrics. (Most social interventions cannot afford the use of such tests nor are the skills to administer them readily available in the developing world.)

Similarly, the methodological issues concerning the choice of statistical tests and the separation of intervention effects from life cycle changes remain relevant for both models. Here, however, the model of rigorous experiment, in its pure form, calls for the application of a set of pre-defined tests for change. For example, an initial statement might be made that the experiment will be considered a success if a two-way analysis of variance on mean percent of standard, by six-month age categories, demonstrates a positive statistically significant change in nutritional status over a two-year period. R-I-A is far more flexible in its use of the array of available statistical methods. Results can be sought using any of the available methods and, then, corroborated or disproved using others. Under the model of rigorous experiment, the statistical analysis leads to firm conclusions. With R-I-A, the statistical analysis is a tool used not only to help form conclusions but also to raise questions and to redefine the experiment.

5.3. Sources of indeterminacy in the attribution of change to an intervention

In all of the cases we have examined, there are many plausible accounts of changes in nutrition status. The nutrition intervention itself is only one possible cause of change among others. Conversely, changes in other features of the community context may mask the effect of the nutrition intervention. The task of interpretation, then, is always one of identifying competing plausible explanations of observed phenomena and discriminating among them.

These are the main sources of confounding effects; that is, effects which may mask the effects of intervention or provide a basis for alternative accounts of change in nutritional status.

(1) Changes in the make-up of the target population. The samples of children in experimental or control groups may change during the course of experiment because of in- or out-migration, death, or, simply, movement of children in or out of the sample group (even though they remain alive and resident in the community). In every data set reviewed during the analysis of community-level interventions, changes in the composition of the target population played a role in the interpretation of the quantitative results. In the Primops data, many children 'disappeared' from one round of data collection to the next. (This was probably due to a combination of factors including in- and out-migration, death, intra-barrio movement and errors in the recording of data.) A search for systematic patterns in the data revealed that a higher proportion of the malnourished disappeared in each round than of the nutritionally sound – 50% of the severely malnourished as compared to about 20% of the less malnourished. The improvement in nutritional status of the population from round to round, in the aggregate, may well have been an artifact of that selective disappearance.

R-I-A cannot prevent shifts in the make-up of either experimental or control groups; however, it provides researchers the opportunity to recognize the existence of such shifts while it is still possible to ascertain the reasons for and the magnitude of those shifts. In action oriented interventions, change agents are usually concerned more with delivering service to willing participants than in finding former participants to learn the causes underlying their disaffection. The inherent emphasis in R-I-A on monitoring impact to direct the intervention is likely to induce change agents to (a) track their initial sample group to facilitate the charting of progress in response to intervention, and (b) understand the causes of disaffection of the sample so as to better design the package of services to eliminate those causes.
(2) Confounding changes in project environment. Changes in project environment may be of many kinds — climatic, ecological, economic, political, social, or administrative. Three examples are drawn from our prior work.

(2.1) Climatic. In Esperanca, portions of the baseline survey were administered in different seasons (six months apart) but the resurvey was done in a single season. Therefore, it is possible that the different magnitude of change in the matched pairs of experimental and control villages may be a reflection of the seasonal shift in nutritional status taking place ordinarily. In Kottar, we could not separate out the impact of the easing of a highly localized drought from the effects of program participation.

(2.2) Administrative. Changes in the availability of other services or infrastructure often accompany intervention and confound the interpretation of analytic results. In Candelaria, the introduction of an improved water and sewage disposal system prior to the Promotoras phase of the intervention may have paved the way for the work of the Promotoras; that is, it is possible that the teachings of the Promotoras were able to induce behavioral change because of the existence of the improved infrastructure.

(2.3) Social. Various changes in social structure, norms or patterns of behavior, may act to alter nutritional status. In Kottar, leaders of the Social Service Society emphasized the need for social and behavioral change and, in fact, viewed the nutrition component of their program as the incentive to bring mothers into their classes so that the Society could work to organize the mothers and modify their social behavior. It is difficult to attribute change in nutritional status to supplementary feeding or nutrition education when, according to the Kottar change agents, their greatest successes have been in organizing communities to take advantage of all government offered services as well as their own labor and resources.

It is not possible to insulate experimental or control groups from such shifts in environment. More importantly, it is often impossible to predict, at the outset of an experiment, which changes in environment are likely to have an effect on nutritional status. After the period of experimentation, as we have discovered, the data in hand is usually insufficient to allow analysts to discriminate between an environmental change which may have influenced nutritional status and an intervention which was designed to do so.

The initial advantages of the model of R-I-A are twofold. First, the practitioners and subjects are closer to the project environment than are researchers — closer, often, both in spatial proximity and sustained contact in understanding. They are more likely to identify the changes in environment and the related changes in behavior, which may be affecting nutritional status. Second, when such environmental shifts are noted early in the experimental process, and when analysts discover early on that available data does not permit discrimination of the effects of such shifts from the effects of intervention, it may be possible to design derived experiments which will permit such discrimination.

Such experiments may be of two kinds, depending on the confounding changes at issue. The first, the strategy of differentiated environment, separates out groups or regions on the basis of high and low exposure to the environmental change in question. For example, in Kottar it might have been possible to differentiate high and low drought subregions; both groups would still have been subject to the experimental intervention. The second strategy of derived experiment, the strategy of differentiated treatment, is really a version of the familiar 'control group' experiment. Here, two groups of children would be identified, both subject to the shift in environment, but only one subject to experimental intervention. In Primops, for example, two groups of children might have been treated, but one subject to the benefits of improved sewage disposal but only one subject to the interventions of the Promotoras.

Both of these strategies are subject to possible difficulties. In differentiating groups by 'high' and 'low' exposure to environmental change, the experimenters may introduce new sources of variation. In a 'low drought' region in Kottar, for example, children might have been surrounded also by a more prosperous economy or by a richer supply of foods. A similar difficulty would arise in the strategy of differentiated treatment: the non-treatment group might be subject to differences from the treatment group in more ways than one. The first difficulty might be circumvented by a further differentiation of the sample. In the Kottar example, experimenters might try to disaggregate poor and less poor families in the low drought
area. And the second difficulty might be circumvented by the use of randomizing techniques in the selection of the non-treatment group. But both of these methods may prove difficult or impossible to apply in the field. Constraints on resources limit the number of new treatment groups that can be set up, and also limit the number of non-treatment groups that can be observed. Continued differentiation of the sample may produce ‘cell’ sizes too small for significant analysis. Finally, the establishment of non-treatment groups may be politically impossible or unacceptable on ethical grounds.

(3) Artifacts of the experiment. A third category of confounding effects consists of changes introduced by the process of experiment itself. These may be of several kinds. For example, in Honduras, the change agents introduced new criteria for gaining entry into the program which affected interpretations of change in nutritional status. Specifically, nine months into the feeding component of their intervention, they eliminated the worst cases from consideration producing, in the aggregate, a spurious improvement. In some projects, other change may occur in the methods of data-collection, the reliability of data, or the quality of the service delivery system. In still other cases, there may be ‘Hawthorne effects’; subjects responding not to the particular contents of an intervention but to the mere fact of being subjects of an experiment. Or, in a related effect, they may be responding to a feature of the experimental process (their relation to the experimenter, for instance) which experimenters themselves recognize as a substantive feature of intervention.

In the model of rigorous experiment, it is difficult or impossible to recognize and correct for such effects. Researchers at a distance from the actual context of intervention and observation are unlikely to detect clues of the existence of processes of this sort. If they should detect them it would be difficult or impossible, in the course of analysis conducted after termination of the experiment, to build their awareness into their interpretation of the data. Distant from the actual delivery of services, they are unlikely to spot important features of that process which have not been named as significant in the experimental design. And the ‘double-blind’ or ‘placebo’ methods sometimes used to respond to such difficulties are difficult or impossible to administer in the context of communities in developing regions.

With the model of R-I-A, however, researchers have direct, on-line exposure to intervention process and context. They are in a position to observe what intervenors actually do and how subjects of the experiment respond. Thus, they are more likely to be able to detect the presence of artifacts of experiment.

In some cases, their observations might lead to changes in experimental procedure (for example, by rectifying a distortion of the experimental sample). Their observations may lead them to discount spurious changes in outcome measures. When researchers suspect that subjects may be discounting the mere fact of experiment, the experimental period may be prolonged to allow for a decay in this effect. When researchers suspect that some unnamed and unanticipated feature of the experimental process is producing significant results, they need not respond by seeking to eliminate or control for that feature. On the contrary, they may seek to build it into the experimental design in a conscious and systematic way. (For example, they may seek to cultivate a particular kind of relation between experimenter and subjects of the experiment.) For the experimenters are not bound by a need to keep experimental conditions constant, or to work toward the formulation of propositions generalizable to a type of community context; their objective is specific to the project in which they are engaged. ‘Hawthorne effects’, in the most general sense, need not be eliminated. However, they must be detected and described and they may be exploited in the redesign or redescription of experiments. These are functions researchers are better equipped to do when they have a continuing, on-line relation to the process of experiment.

5.4. Experimental norms and sources of indeterminacy in R-I-A

We are now in a position to summarize the ways in which researchers might respond, under the model of R-I-A, to the sources of indeterminacy we have encountered in our analysis of nutrition intervention data.

Some sources of indeterminacy, as we have encountered them, are not peculiarly related to the differences between the models of rigorous experiment and R-I-A. These include the choice of measures, cut-points, and standards for determining
nutritional status, and the pattern of variation of nutritional status with age. Although these factors may be taken into account in the analysis of experimental data – for example, by using characteristic curves, or by deriving local standards – it is not clear that such analytic responses require a shift from one model to the other.

But other sources of indeterminacy would be significantly affected by a shift to the model of R-I-A. Errors in the collection and processing of data ('dirty data') are subject to a feedback cycle of their own. Their detection and correction would be far more likely, both from the point of view of both opportunity and motivation, if data were processed and results of processing fed rapidly back to practitioners in the course of experimentation.

Confounding shifts in context of experiment, both in the target population and in the environment, are not predictable before initiating intervention, at least not predictable with enough specificity to inform the design of experimental conditions. After intervention, when such shifts are discovered, it is unlikely that their confounding effects can be discriminated on the basis of data then available. It is possible, however, to detect such shifts in the process of experiment; and most likely to do so are the practitioners and subjects of intervention. Such detection opens up the possibility of gathering additional data and the subsequent application of statistical techniques to account for the changes in experimental context or the environment. It is also possible, some of the time, to design derived experiments, through the strategies of differentiated environment or differentiated treatment, to sort out the differential effects of intervention and environmental change.

It is also possible through close observation of the processes of intervention and responses to intervention to discover phenomena which permit discrimination among the contending explanations of change in outcome measures. Such process observations may also uncover features of intervention, or its context, which inform new interpretations of the contents of intervention and suggest redesign of experiment. But, again, such observations are most likely to be made by observers who are educated by the attempt to make sense of the data; that is, to do on-the-spot research.

It is in these ways that R-I-A offers a route to experimental validity under conditions which cause the model of rigorous experiment to fail. However, two caveats must be kept in mind.

The first is that the objectives of R-I-A are different, and in a sense more modest, than those of rigorous experiment. R-I-A aims at producing experimentally valid propositions only for the community context in which intervention is undertaken and only within the bounds of a particular state of that context. It does not aim at producing experimentally valid propositions generalizable to other community contexts, although processes undertaken and results achieved in one community may serve in another community to stimulate and guide R-I-A there.

The goals of intervention in R-I-A are discovered, in their specific forms, in the course of iterative, on-the-spot experimentation, just as strategies of intervention appropriate to that particular context are discovered in that way. Thus, R-I-A aims at context-specific effectiveness and at context-specific discovery. The process of iterative data-collection, analysis, redesign and re-observation has two faces. From the point of view of research, it is an iterative process of experiment, analysis and further experiment whose aim is to remove sources of indeterminacy in the data. From the point of view of the intervention, it is an iterative process of learning to produce, in the context and according to criteria of effectiveness peculiar to the context, desired changes in outcome measures.

A similar consideration informs the definition of 'objectivity'. Because interventions are seen as context-specific, there is no attempt to describe relationships between features of intervention, context, and outcome which are generally valid across contexts. If the experimenter influences processes and outcomes, perhaps as a consequence of his hypothesis, the task is not to remove that influence but to observe it and understand it. It is not necessary that other experimenters, in other contexts, be able to perform the same actions and achieve the same results. The task is, rather, to describe such influences and (when they are seen as positive) to incorporate them into the intervention.

The second caveat is that there is no guarantee that, in any given context, R-I-A will converge on unequivocal interpretation. The proliferation of derived experiments may multiply sources of indeterminacy. In such cases, inquirers might, in any
case, persevere in the conduct of new experiments until a positive change in nutritional status has been achieved. The final 'moves' in the sequence of interventions would then have been affirmed (that is, positively appreciated) even though the theory of intervention underlying those moves would be neither confirmed nor disconfirmed. Inquiry would then have achieved a significant though limited goal - the production of a move affirmed through its observed results. From the point of view of the practice of intervention, theory testing is always in the service of more-testing, so that such a limited goal is by no means negligible. To put the matter differently, in ordinary intelligent practice, the intervenor carries out on-the-spot experiments to test his strategies and assumptions only to the point of producing a move whose results he likes. There is always the possibility that explanations of those results might be other than the ones that informed the practitioner's intention. That possibility usually does not, and need not, bother the practitioner. It is when a present case is seen as a preparation for future projects that such possibilities are truly worrisome.

6. A partial application of R-I-A

Upon completion of the research project on the analysis of community-level nutrition programs described earlier, we believed that it would be possible to put into practice most of the elements contained in the R-I-A model in the context of a nutrition evaluation effort. Such a project was undertaken during 1982 with the objective of testing as much of the R-I-A model as possible. The project was an evaluation of the Food For Peace grant program in the country of Sri Lanka [5]. One central question of the study was whether the food donation program had a favorable impact on the nutritional status of children under six years of age and, if so, by how much?

In order to maximize the involvement of field-level practitioners in the evaluation, a condition essential in implementation of R-I-A, the project was partitioned into three phases. Phase I followed the ordinary approach of designing a study which attempted, as best as possible, to resolve the question of nutritional impact from the program. For the children under six years of age who participated in the Thripasha program, a field data collection effort was implemented which ultimately generated some twelve thousand weight and age observations for eighteen hundred children. Since Thripasha is distributed in Sri Lanka as one component of a package of health services, caution was taken to partition the sample in a variety of ways to control for possible differences in impact reflecting variations in the delivery system or recipient population. The sample was stratified by geographic region, by the type of health care provided, by the governmental unit responsible for staffing and funding the distribution center, by the socio-demographic characteristics of the family of the recipient child, and by the age of the child. Social and demographic family variables included age of mother and father, parental occupation as a surrogate for income, number of siblings and birth order. The selection of this particular stratification was based on the cost and feasibility of data retrieval balanced against what our prior experience showed to be most critical. Our judgments were reviewed by local practitioners and modified to conform to their experience as well. Phase I did not depart from the model of rigorous experiment except in one important aspect - there was a-priori recognition that there would be, at best, some indeterminacy resulting from the analysis even if measurable change in nutritional status was confirmed.

Phase II consisted of an exhaustive analysis of the data gathered in Phase I. A report was prepared consisting primarily of tables, charts and graphs showing the relationships between variables deemed to be of potential importance. Characteristic curves portraying age cohort comparisons of the nutritional status of children who had participated in the program for a substantial period of time to the status of children who had just recently entered in the program provided concrete evidence that participants of all ages were nutri-

5 Thripasha is the Sri Lankan name for the food supplement used in the Sri Lankan Maternal Child Health Program. It means three nutrients in the local languages and consists of extruder-cooked, locally-grown corn and soy mix with vitamins, minerals, a small amount of sugar and donated instant corn soy meal (ICSM). The children are brought to a health clinic where they receive not only Thripasha but also treatment for a variety of health problems. The health clinics in Sri Lanka are staffed by fully trained doctors and nurses.
tionally better off than new entrants. Tables were included presenting evidence that the magnitude of the benefit derived from participation appeared to vary by the type of organization rendering service and by the physical setting of the clinic, rural or urban. Nutritional status was shown to be related to several of the other stratification variables as well.

In Phase III, armed with specific results of the analysis, the evaluation team engaged in a dialogue with local practitioners concerning the causes of the observed results. In several instances, the interpretations of the analysis were modified substantially from what seemed obvious from the analysis. For instance, differences in outcome which seemed to be related to the type of unit providing service and type of health care offered in conjunction with the food supplement were found to be highly correlated with the clinic staffing levels per beneficiary. This competing explanation was first proposed by one of the local staff and then tested with on-the-spot observations of clinics in several different parts of the country. Thus, what appeared to be strong evidence in support of the hypothesis that the type of governmental unit administering the program played a role in causing positive response to it was reinterpreted to be evidence supporting the role that staffing levels play in insuring effective clinic operation.

It is important to note here that data on staffing levels had been gathered at the onset and, initially, found to be of no importance. The reason for this lack of detection of a relationship during the early rounds of analysis was that the variables selected for testing the hypothesis were too aggregated. In short, the study design team, even though it included knowledgeable local staff, had selected and measured the wrong variable. It was only after successive analyses that enough local perspective was gained to narrow and focus the definition of staff to a level which revealed the relationship between staff size and nutritional impact. It could be argued that had the correct variable been measured at the onset, this problem would not have arisen. However, it is our belief that the often subtle process of deciding precisely which variable is most appropriate to gather before analysis is extremely difficult. Regardless of the expertise of the analyst, mistakes and omissions are usually made and, as a consequence, the iterative process embodied in the reflection-in-action model becomes essential.

Since Phase III was concerned more with the determination of attribution rather than outcome, the issue of competing explanations was central to this part of the study. The Sri Lanka evaluation, like most such studies, contained several possible competing explanations for the observed outcome besides the intuitively appealing one that the program worked. The R-I-A model called for a comprehensive enumeration of competing explanations, even though some were beyond the scope of inquiry. During the four week in-country study period, the evaluation team was able to muster evidence for rejecting all significant competitors, thereby determining, with some degree of confidence, that the observed improvement in nutritional status was attributable to the bundle of services associated with Thripoja distribution. Had we not been able to articulate those competing explanations and perform on-the-spot research, our evaluation would have been inconclusive. It should be noted, however, that it was not possible to separate the effect of Thripoja from the package of other services rendered at the same time as the food supplement nor was it feasible to generalize the findings regarding food supplementation to other settings.

Adoption of R-I-A to an evaluation of the sort done in Sri Lanka provides some evidence concerning the usefulness of the approach. However, as mentioned earlier, only some elements of the model were applicable. Missing were the elements building in the evaluation and on-the-spot research as part of the ongoing program. An application including these elements remains to be done.

7. Problems in implementing R-I-A

In conclusion, we consider some of the problems likely to be encountered upon implementation of R-I-A. Problems exist at both the individual and institutional level.

1) New skills are needed: Usually, at the level of the individual, the researcher and practitioner come from different educational backgrounds and are motivated by different interests – the practitioner being action oriented and the researcher more interested in the search for knowledge. R-I-A calls for the attributes of both to be present in one individual or, at least, represented equally on a team. The model also calls for a blend of behaviors as well. At the field level, there is often a short-term conflict between expending resources
(time or money) on gathering and analyzing data and rendering services to beneficiaries. We have seen several instances in the field where this conflict is quite real. Too frequently, this conflict manifests itself when political pressures for expanded coverage are brought to bear on the change agents; that is, where short term political interests are placed above long term effectiveness.

Also, as scientists become more fully integrated into action-oriented intervention, they will have to modify their attitudes and behavior. Emphasis on the numeric analysis of clean data as presented in current academic journals must be replaced with emphasis on the quality of that hard data, the integration of ‘soft’ data sources into analytic frameworks, and a willingness to accept and articulate the indeterminacy of quantitative analysis in social settings. More importantly, scientists must sacrifice apparent academic productivity (measured by the number of scholarly publications produced) in order to become (and to remain) close to local conditions.

(2) Quantitative data must be kept in perspective: There is a tendency among individuals with too little experience in analytic endeavors to overreact to assertions supported by quantitative evidence. Because R-I-A calls for the continuous generation of quantitative assessments of program progress, there is considerable risk of such overreaction. Without a minimum critical level of expertise in experimental design and in the interpretation of quantitative analysis, practitioners can react prematurely or improperly to mathematical results and, in doing so, do more harm than good. Decisions may be based on improperly or incompletely developed ‘hard’ data at the expense of field wisdom. Over reliance on ‘hard’ numbers is not limited to practitioners in the field. A similar pattern exists within funding bureaucracies.

(3) Institutions must change the mechanisms for supporting intervention experiments: If a funding agency accepts R-I-A, then it can no longer require complete specification of the details of an intervention in advance. Only the starting position or ‘best guess’ can be stated along with the process by which successive iterations will converge on the most useful intervention design. But most funding agencies are geared to receiving requests for support in terms of services provided, staff time, supplies, materials and so forth. Generally, even if only one year at a time is funded, sponsors request fairly detailed projections of future years’ needs. Furthermore, evaluation of program efficacy is considered to be an activity useful primarily for decisions about future funding levels and, therefore, in need of objectivity brought by independent reviewers working at a distance from project administration.

In order for funding sources to support interventions embracing the R-I-A model, requests must be deemed acceptable which do not include detailed long-term program implementation plans. In turn, the absence of such plans forces the sponsor to change and conduct its own evaluations of program progress, evaluations to determine future funding level. No longer is it possible to evaluate based on a comparison of the delivery of services actually rendered by the program against the detailed plan provided in advance. Historical emphasis upon process indicators as surrogates for difficult-to-obtain outcome indicators is not acceptable. Rather, evaluations must focus on outcome indicators and the often complex task of attributing observed outcome to the intervention.

References