Evaluating Social Marketing in Nutrition: A Resource Manual

Prepared by:

James Hersey, Ph.D.
Sarah Daugherty, M.P.H.
Research Triangle Institute
Washington, DC

and

Health Systems Research, Inc.
Washington, DC

Prepared for:

Office of Analysis, Nutrition and Evaluation
Food and Nutrition Services
U.S. Department of Agriculture
Alexandria, VA
Project Officer: Patricia McKinney
Contract No. FCS-53-3198-5-038

December 1999

ACKNOWLEDGMENTS

This report was prepared by James Hersey and Sarah Daugherty of Research Triangle Institute under the direction of Loren Bell of Health Systems Research, Inc. under contract to the Food and Nutrition Service (FNS) (Contract No. FCS-53-3198-5-038).

We wish to acknowledge the support and assistance we received from Patricia McKinney who served as our Project Officer for this study. Valuable guidance and thoughtful reviews were provided by Steven Carlson, Ted Macaluso, and Carol Olander of FNS.

Special thanks go to all the individuals in State food stamp nutrition education networks and nutrition education programs who contributed the time and suggestions to this effort. Particular thanks go to Jamie Benedict of the University of Nevada at Reno; Sue Butkis of Washington State University; Darrell Carrington and Debra Keenan of Rutgers University; Joyce Counihan and Jackie McClelland of the North Carolina State University; Sue Foerster, Jennifer Gregson, and Jan Lewis of the California Department of Health; Judythe Gatchell, Sunita Raynes, and Christine Sady of the Maine Nutrition Network; Pat Hammerschmidt of Michigan State University Extension; Michael Holaday of Holaday Research in Lansing, Michigan; Larry Jones of the University of Wisconsin; Suzanne Murphy of the University of Hawaii; Robin Orr of the University of Illinois; and Karen Zotz of North Dakota State University for their work which provided the examples used in this manual. It is hoped that the application of the evaluation principles in this manual can contribute to our understanding of how to improve social marketing initiatives directed at low-income Americans.

Table of Contents

Chapter 1. Background and Purpose 1

A. Social Marketing of Healthful Nutrition 1
B. The Need for Evaluation 2
C. Specification of a Program Rationale 4
D. Organization of this Manual 6

Chapter 2. Formative Evaluation 7

A. Uses of Formative Evaluation 7
B. Strategies for Formative Evaluation 8

Chapter 3. Process Evaluation 10

A. Uses of Process Evaluation 10
B. Strategies for Process Evaluation 11

1. Assessing the Development of Partnerships 11
2. Assessing the Leverage of Resources 14
3. Tracking Message Dissemination 15

a. Evaluation Televison and Radio Distribution 16
b. News and Consumer Affairs Coverage 16
c. Dissemination of Paid Ads and PSAs 16
d. Monitoring Print Distribution 18
e. Documenting Distribution via Kiosks and Interactive Information Systems 19
f. Documenting Information Requests 20
g. Documenting Community Activities 20
h. Assessing the Cost-Effectiveness of Information Channels 21

4. Documenting Policy, System, and Environmental Change 22

a. Policy Change 22
b. System Change 23
c. Environment Change 23

Chapter 4. Impact Evaluation 26

A. Challenges in Evaluating Behavioral Impact 26
B. Assessing the Implementation of Social Marketing 27
C. Measurement Selection 28

1. Clarifying Information and Decision Making Needs 28
2. Understanding Campaign Rationale 28
3. Uses of Theory to Inform Measurement Selection 30
4. Selection of Measurement Approaches 32
5. Selecting Appropriate Measurement Tools 33

D. Research Design 36

1. Length of Time between Data Collections 39
2. Statistical Power 39
3. Estimating Response Rate 40

E. Eliminating Plausible Alternative Explanations 42

1. Selection Bias 45
2. Problems Relating to the Use of Repeated Interviews 45

a. Differential Attrition 45
b. Effects of Repeated Testing 46

3. Extraneous Factors Related to the Passage of Time 46

F. Resources and Tradeoffs in Evaluation Design 48

1. Resource Tradeoffs 48
2. Potential Enhancements to Research Design 49

G. Conclusion 50

Bibliography

Glossary

Appendix A

Appendix B

CHAPTER 1

Background and Purpose

This manual is intended to serve as a resource to State and community programs. It will provide guidance on evaluating social marketing of nutrition education directed toward low-income audiences such as participants in the Food Stamp Program. The sections that follow provide an overview of social marketing, explain why evaluation of these types of programs is important, and describe strategies to meet the challenges posed by the evaluation of social marketing efforts.

A. Social Marketing of Healthful Nutrition

Social marketing is an audience-centered approach that features multiple, reinforcing channels of communication along with public policy and environmental changes to influence behavior. Andreasen (1995) offers the following definition:

> Social marketing is the application of commercial marketing technologies to the analysis, planning, execution, and evaluation of programs designed to influence the voluntary behaviors of target audiences in order to improve their personal welfare and that of their society (p. 7).

Although it is less intensive than ongoing instruction, social marketing can reach more people than is otherwise possible in traditional nutrition education classes. Social marketing also can reinforce the messages that have been taught through individually directed nutrition education. Social marketing efforts reach audiences through multiple communication channels, such as television, print, point-of-purchase ads, and interpersonal communications in culturally appropriate ways. Social marketing emphasizes the importance of keeping the target audience

involved in needs assessment, message development, and refinement of messages and delivery strategies. Social marketing techniques include analysis and segmentation of audience needs that influence message development and delivery, use of the media to reach large numbers of people, and an emphasis on ultimately changing behavior.

B. The Need for Evaluation

Program evaluation of social marketing can serve two important functions. First, it can assist in providing feedback to support continuing improvement of activities. Second, it can help to establish the effect and hence the accountability of the program. Both of these goals are important, and it is often possible to design evaluation approaches that address both objectives in ways that result in more useful information than if only one objective had been sought. For instance, results indicating the attitudes and beliefs, as well as the behavior, that have been affected can help to corroborate the strength of evidence of a social marketing activity as the source of the behavioral change and also provide feedback about next steps to be undertaken in an ongoing intervention effort.

Yet despite the importance of the endeavor, evaluation of social marketing efforts faces a number of special challenges.

- Unlike a direct teaching program where one knows who participates in a nutrition education class, many of the messages in social marketing efforts are delivered via TV, radio, print media, and point-of-purchase ads. This makes it necessary to assess the extent to which individuals were actually exposed to nutrition education messages.

- Social marketing often incorporates efforts to affect policy, system, or environmental changes (e.g., the availability of low-fat foods). This range of activities can be important to measure because these intermediate outcomes can enable and reinforce behavioral changes.

- Because the exposure to nutrition messages via social marketing methods is often less intense than in an ongoing nutrition education class, it is important to use evaluation designs and measurement tools that are sensitive to small changes in behavior.

- Social marketing entails an ongoing program of evolving messages that are tailored to meet the needs of a target audience. Because of this, it is important to design studies in a way that generates timely feedback to inform continual improvement and refinement of activities.

This manual suggests a set of evaluation strategies that can address these challenges. These strategies encompass three major types of evaluation:

- Formative Evaluation. Social marketing efforts include planning, needs assessment, and materials development. Formative evaluation techniques, such as focus groups, cognitive testing, and in-depth interviews, are often widely applied in the pretesting of evaluation instruments to ensure that survey questions are understandable and culturally appropriate.
- Process Evaluation. Social marketing of nutrition education to low-income populations is a continually evolving process, and many of the networks and programs engaged in these activities are relatively new. Process evaluation can help to answer questions regarding the implementation of a social marketing effort, describe barriers encountered, and offer guidance about strategies to address these barriers. For instance, process evaluation can describe how effectively social marketing activities actually reached members of a particular target audience. Process evaluation can also describe the dissemination of messages, and the effects of efforts in terms of system, environmental, or policy changes that can influence behavior.
- Impact Evaluation. This report discusses two types of evaluation under the category of impact evaluation: efficacy evaluations, which are often thought of as evaluation of pilot projects with rigorous implementation, and effectiveness evaluations, which are concerned with evaluation of widespread implementation of a program. Impact evaluation assesses immediate indicators of behavior, such as changes in food resource management, food safety, or dietary intake.

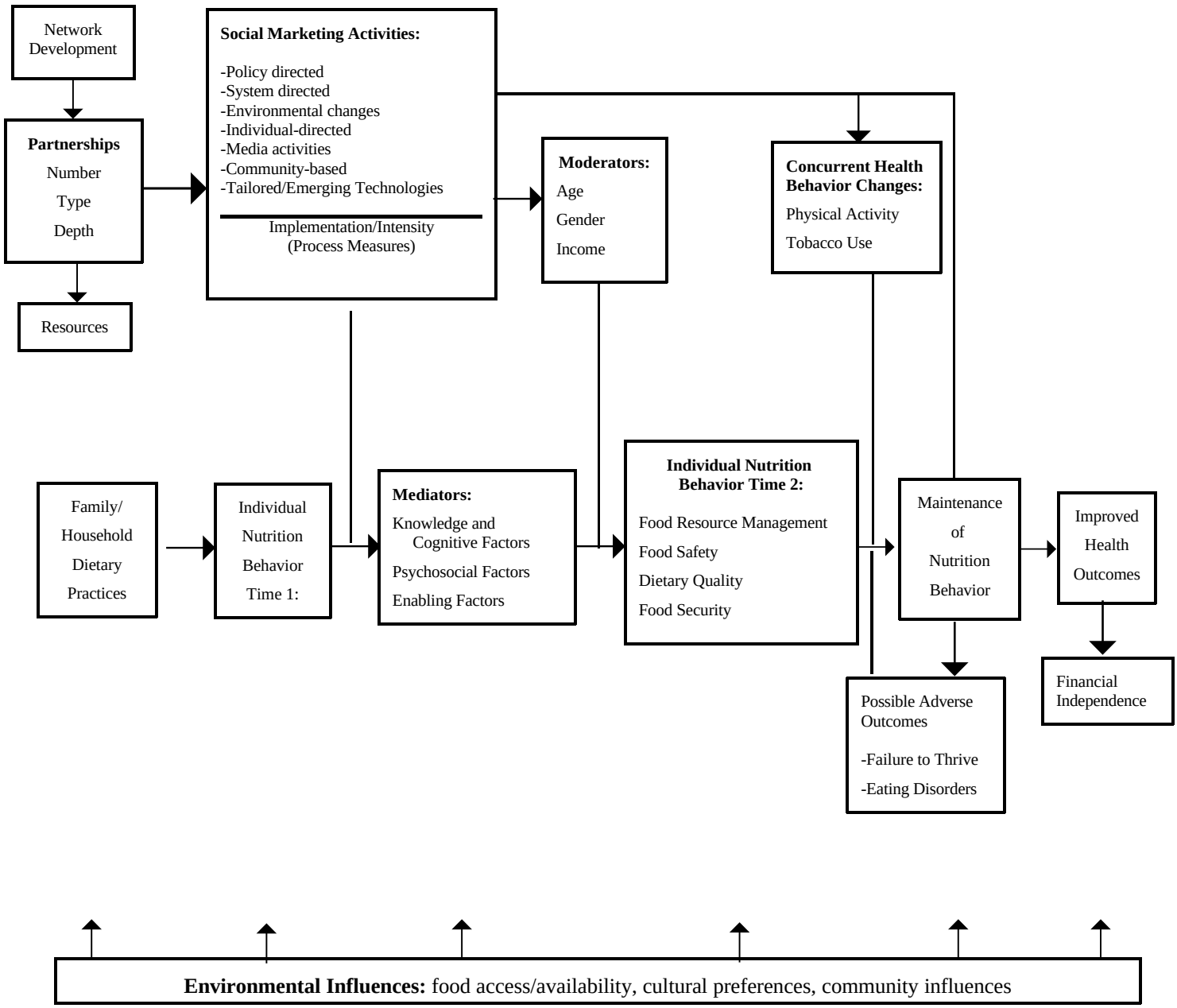
In contrast, this report does not focus on long-term health outcomes. The linkage between dietary behavior and nutritional status has been reasonably well established in other research, and the time needed to conduct these long-term studies often exceeds the schedule for providing feedback to a social marketing effort. Similarly, this report does not discuss biomedical indicators of nutritional status. While biomedical indicators can be important in validation studies and epidemiological research, collecting them can be expensive and can raise issues of respondent burden and cultural sensitivity that need to be carefully worked through.

C. Specification of a Program Rationale

Whichever of these types of evaluation is employed a useful first step in planning an evaluation effort is to describe the rationale or conceptual framework for a program. Discussion with program developers can help to detail a comprehensive list of the targeted behavioral changes, as well as the activities intended to produce these changes. This list can lead to a set of "hypothesized" causal linkages that can prioritize the issues discussed for the evaluation and provide guidance for selection of specific measures.

An example of a program rationale is shown in Exhibit 1-1. This example was drawn from USDA's Statewide Nutrition Education Network Demonstration Project that issued cooperative agreements to 22 States to develop collaborative networks of public/private partners that would incorporate social marketing in efforts to provide nutrition education to limited resource audiences (Hersey, Anderson, Bell et al., 1999). The rationale for the project was that networks would develop partnerships and identify resources that can be used to deliver a variety of social marketing activities. These activities include efforts to affect the policy, system, and environmental factors that influence subsequent nutrition education activities. Social marketing can include interventions directed at individuals, as well as media and community interventions. Activities developed by the Nutrition Education Networks under this program often were focused on changing client behaviors associated with (1) food resource management, (2) food safety, (3) dietary quality, and (4) food security. In addition, social marketing activities were also designed to reinforce the maintenance of existing healthful behavior.

Exhibit 1-1. Conceptual Framework: Evaluation of Social Marketing in Nutrition



For purposes of simplicity, this exhibit does not show the interactions among different types of interventions.

Individual programs are expected to have more detailed rationales, depending on the types of activities they engage in and the focus of their social marketing effort. Still, this general conceptual framework helps to guide the discussion of evaluation that follows in the rest of this document. Describing the program rationale can do much to guide priorities for evaluation of individual programs. For instance, a program rationale should identify a series of hypothesized linkages between a program's activities and its impacts. The evaluation approach can build a monitoring system that provides routine management information about expected process accomplishments (e.g., distribution of social marketing activities) and then focus on the linkages where the strength of the hypothesized effects has not been firmly established (e.g., attention to materials by the target audience, effects of those materials on attitudes, and immediate behavior).

D. Organization of this Manual

This document is not intended to be a complete text on evaluation; a number of excellent books have been written regarding this topic (as cited in the Bibliography). Similarly, excellent texts have been published on nutrition measurement (e.g., Willet, 1998) and a set of white papers, supported by USDA offers guidance about tools that can be used to evaluate nutrition education and social marketing programs among low-income audiences. Other sources provide excellent reviews of work in social marketing *(*e.g., Andreasen, 1995), and prior evaluations in nutrition education (Balch, 1994; Lytle, 1994; Contento, Balch, Bronner, et al., 1995). Rather, this manual is intended to offer guidance about how to apply these tools to the challenges of conducting evaluation of social marketing of nutrition education programs in network like settings.

Subsequent sections of this manual describe approaches to formative evaluation (Chapter 2), process evaluation (Chapter 3), and evaluation of impacts on behavior (Chapter 4). Attachments to this report provide information about other evaluation resources, information about sample size estimation, and examples of evaluations of social marketing activities conducted of state- and community-level efforts.

CHAPTER 2

Formative Evaluation

Formative evaluation provides feedback from the target audience about proposed activities and materials to help develop a social marketing effort. Formative evaluation helps answer the question, "What works better?" It is used to test and modify messages and modes of communication in the planning stages of projects. This chapter describes use of formative evaluation in social marketing initiatives to improve the nutrition of low-income audiences.

A. Uses of Formative Evaluation

A defining feature of effective social marketing programs is the use of formative evaluation to ensure that messages, strategies, and materials are carefully adapted to meet the needs of the target audience. Formative evaluation is used in needs assessment, concept development, concept testing, and testing of materials. During planning, formative evaluation provides important information for concept testing and message development about what is likely to work in actual operation of a campaign. Formative evaluation also helps to test social marketing materials before they go into full-scale production and release. Results from formative evaluation provide information about how to modify messages to better reach the target audience and achieve and maintain desired behavior change (e.g., improved food resource management, food safety, or dietary practices).

Some of the questions formative evaluation attempts to answer are:

- What meaning does the behavior or issue in question have for the target audience and for subgroups?

- What are the best ways (i.e., channels of communications) to reach the target audience?

- What barriers exist to communicating with the target audience? What strategies exist for overcoming these barriers?

- Does the target audience understand the message? Is its format appealing to the target audience?

- Which message statements are most powerful for influencing the knowledge, attitudes, and behaviors of the target audience?

In circumstances when total evaluation resources are limited, it is recommended that formative evaluation activities be given high priority.

B. Strategies for Formative Evaluation

Formative evaluation can take advantage of the comparative strengths of a variety of data collection methods depending on feedback requirements, as described in Exhibit 2-1. For instance, ethnographic and observational studies can help to inform the development of practical programs by learning about the ways that people utilize foods within the cultural and family context of their community. Focus groups provide in-depth qualitative information that can be helpful in understanding the target audience by hearing them discuss nutrition issues in their own words. On the other hand, intercept methods can provide a timely method to obtain reactions to materials. Survey information can help round out this assessment by obtaining information on a representative sample that can provide a context against which to interpret qualitative findings.

Exhibit 2-1. Assessment of Data Collection Methods for Formative Evaluation

Method	Application	Strength	Limitation
Focus groups	Obtain target audience reaction to proposed message in small group discussion setting	Detailed, elaborative responses in audiences' own language	Nonrepresentative sample
Individual in-depth interviews	Obtain target audience reaction to proposed messages	Opportunity for probing and follow-up questions	Limited sample size
Observation	Observe pilot delivery of messages	Insight into factors influencing message acceptance	Nonrepresentative data
Ethnographic analysis	Investigate in-depth cultural factors influencing behavior	Understanding of cultural influences on message acceptance	Length of time to complete studies
Central location intercepts	Brief review or test of messages and materials	Inexpensive method to allow respondents to see and react to materials	Nonrepresentative sample
Mail surveys	Obtain feedback from staff and volunteers delivering programs	Low cost method to obtain insight on implementation	Special effort to ensure adequate response rate
Telephone surveys	Validate ideas and concepts and obtain tracking data	Timely, representative data	May require special tracking strategies to reach low-income audiences
In-person surveys	Obtain feedback on materials	Coverage of non-telephone households	Expensive

A number of excellent source materials on these formative topics exist (as described in the attachment to this document), and training courses in these methods are widely available, and most marketing agencies can offer staff with experience in the application of formative evaluation methods. Hence, rather than attempt to replicate this material, this document simply underscores the importance of formative evaluation techniques to assure that materials and programs that are developed meet the needs of target audiences.

CHAPTER 3

Process Evaluation

Process evaluation is the use of empirical data to assess the delivery of social marketing efforts. Process evaluation is used to assess the fit of the social marketing effort with the program goals at one or more points during implementation. It helps answer questions like, "What was actually delivered?", and "Was the program delivered as intended?" This chapter discusses the use of process evaluation to assess key intermediate outcomes, such as the development of partnerships, the leverage of resources, and effects on system, policy, and environmental change that can influence the nutrition-related behavior of low-income populations.

A. Uses of Process Evaluation

Process evaluation assesses the implementation of a social marketing program and investigates how well and with what intensity messages are delivered to the target audience. Process evaluation can also look at initial outcomes, such as the development of effective partnerships and the leveraging of resources to help in the delivery of nutrition education or to improve access to nutritious foods. Information from process evaluation can identify ways of improving the implementation process. In such instances, the focus of an evaluation should be upon the dissemination process:

- How can the network best leverage the participation of State, local, and private-sector groups in nutrition education efforts?
- What channels best reach a particular audience?

Process evaluation also can help to document progress of social marketing efforts in terms of:

- Increasing the diversity and strengthening the coordination, support, and collaboration of partners working on issues of nutrition education;
- Leveraging additional resources for nutrition education as well as increasing the coordination (and avoiding duplication) of resources;
- Increasing the reach of nutrition education; and
- Promoting policy, system, and environmental changes that provide greater access to healthful foods and nutrition education.

These types of questions are by no means mutually exclusive, and evaluations that address both implementation and impact issues are often best able to identify the effects of an effort and provide the feedback that can support continued improvement.

B. Strategies for Process Evaluation

Specifically, as discussed in the sections that follows process evaluation of social marketing initiatives is practically useful in (1) assessing the development of partnership, (2) assessing the leverage of resources (3) tracking the dissemination of messages, and (4) documenting policy, system, and environmental change.

1. Assessing the Development of Partnerships

Process evaluation can help to document the number, type, and strength of partnerships involved in social marketing of nutrition education efforts. This can be an important indicator of progress because the greater the number and variety of community partnerships and the deeper the collaborations with these partners, the greater will be the exposure of target audiences to social marketing messages and affordable and nutritious food. Partnerships can help to create a more positive environment for behavior change and a shared commitment for improving the nutrition of members of the local community. Scales describing depth of partnership build on work in community development (Francisco, Paine, and Fawcett, 1993; Hogue, 1993), and there is some evidence that nutrition networks with deeper degrees of

partnership mounted more extensive social marketing programs than networks with less deep partnerships (Hersey, Anderson, Bell, et al., 1999; Orr, 1999; Jones, 1999). Progress over time can be monitored in terms of the number and type of partners, the depth of partnerships, and the types of contributions that partners make toward a social marketing effort in nutrition (Exhibit 3-1).

Exhibit 3-1. Describing Relationships and Contributions of Nutrition Partners

Depth of Organizational Relationship	
Network	Organization has signed on as a member of a formalized network. There is ongoing dialogue and information sharing.
Cooperator	Organization assists with information such as referrals, announces classes, provides space for brochures, and provides access to clients to increase community awareness. The goals with this relationship are to ensure that work is done.
Coordination or Partnership	Organization maintains autonomous leadership, but there is a common focus on issues and group decision making. The emphasis in this relationship is on sharing resources to create something new.
Coalition	Organization has longer term commitment to joint action in the area of nutrition education. Shared leadership exists, roles are defined, and generation of new resources occurs.
Collaboration	Organization contributes to joint nutrition activities and has identified personnel who help advise and make decisions about effective educational programming. An interdependent system is built that benchmarks shared impacts. Consensus decision making and formal links and role assignments are common.
Contributions of Partners	
Education	Organization provides space, service, or personnel to increase the frequency and number of persons who can be reached by nutrition education.
Access to Nutritious Food	Partner enables Food Stamp clientele to have a better selection of nutritious food. This may be documented by self-report from grocers or other retail establishments, self-report from food management staff such as nursing home or daycare or school food service personnel, pre/post observation of food offered in retail stores, pre/post observation of nutritious refreshments or snacks available in organizational meetings, community food recovery efforts, increased farmers markets' acceptance of food stamps or electronic benefit transfer cards, etc.
Coordination	Partner coordinates nutrition messages with other programs' messages, coordinates education with other services at the same site, seeks to reduce duplication of services, reduces application/eligibility processes for clients, etc.
Policy	Partner makes client-supportive changes in its own policies, encourages joint policy discussions with other organizations, or encourages public discussions of current policies. These can include both public policies, such as "Good Samaritan" laws, or private policies, such as increased insurance coverage for diabetes education.

Source: "Reporting Results" section of University of Wisconsin Cooperative Extension Service Web site (http://www.uwex.edu/ces/wnep/).

2. Assessing the Leverage of Resources

It is also useful in process evaluation to track the leveraging of resources devoted to nutrition education and social marketing. An easily accessible method to do this is to compile information on growth in budgets and expenditures through State Food Stamp Nutrition Education Programs (FSNEP). Although they may be a useful indicator of resources for nutrition education, FSNEP expenditure figures underestimate the total effects of programs in leveraging additional resources for nutrition education. The major reason for this underestimation is that FSNEP funding only provides data about "matchable" funds from governmental organizations and does not cover in-kind and staff contributions of private for-profit and nonprofit organizations.

An accurate picture of leveraged resources should include in-kind contributions of private organizations. Some prudence will need to be exercised in gathering such cost data. Of course, it is possible to apply standard accounting practices to estimate the value of in-kind office space, staff time and volunteer time contributed by nongovernmental organizations. The application of such detailed reporting practices, however, can be sufficiently burdensome that simpler methods, such as staff full-time equivalents (FTEs) devoted to nutrition education, may be a reasonable proxy for leveraged resources.

On the other hand, in areas such as media distribution, it may be reasonably easy to obtain dollar estimates of the value of donated broadcast time. For instance, commercial monitoring of television public service announcements (PSAs) can generate estimates of the dollar value of the air time. These estimates are based on the cost to purchase this time for commercial advertising. In many situations in which a social marketing campaign purchases advertising, it is possible to negotiate for additional donated broadcast time, and the stations can provide estimates of the dollar value of such time. It also may be possible to obtain estimates of the value of news coverage. For instance, when a biennial statewide survey showed a decrease in the consumption of fruits and vegetables, the California Nutrition Network used this opportunity to make a series of appearances on television news and talk shows in which presenters discussed convenient, economical, and delicious ways to prepare fruits and

vegetables. It was estimated that the value of the broadcast time associated with this publicity exceeded $500,000 (Hersey, Anderson, Bell, et al., 1999). Where such figures can be compiled, they help convey a clear picture of how a social marketing effort leverages resources.

3. Tracking Message Dissemination

An essential task in conducting process evaluation of a social marketing campaign is determining the dissemination of social marketing materials and messages. Such information is essential to answer an initial question of any evaluation effort: "Did social marketing messages reach the target audience?" Previous evaluation studies have indicated that a common cause for the failure of social marketing campaigns to achieve their objectives has been the relatively limited distribution of messages. While exposure of a target audience to campaign messages does not necessarily result in behavioral change, adequate levels of audience exposure are a prerequisite. In addition, evaluation of the dissemination of materials can provide useful feedback about the most cost-effective channels of communication to use in reaching target audiences.

Data about dissemination, when combined with information on the process of campaign implementation, yield lessons about effective approaches to reach various target audiences. Such information can enable a program to answer questions about what types of distribution strategies and organization efforts result in frequent airing of PSA spots; news publicity for nutrition education; and the mobilization, implementation, and institutionalization of community group activities aimed at improving the nutrition-related behavior of low-income populations. Most fundamentally, it is through documentation of dissemination of social marketing information that one can obtain quantifiable measures of the independent variables of the larger evaluation:

- What methods were most successful in distributing campaign materials to target audiences?
- What types of efforts were most successful for securing intermediary groups' involvement in the social marketing campaign?

- What types of support did community groups find most helpful and what other types of support would they find helpful in future efforts?

a. Evaluating Television and Radio Distribution

Television and radio are likely to play a key role in social marketing efforts. Information can be broadcast through (1) news coverage and consumer-oriented talk shows and (2) paid advertising and PSAs.

b. News and Consumer Affairs Coverage

Despite a tendency to think of social marketing as advertising, nutrition is often news itself and often can be featured on consumer-oriented talk shows. Given the interest of nutrition stories to the agricultural communities in many States and the compelling interest of nutrition in peoples lives, new information about nutrition often warrants news coverage. Since news shows have high viewership, the documentation of news coverage can be important in monitoring the distribution and dissemination of campaign messages. Consumer talk shows also can be useful in reaching particular target audiences.

There are several ways to monitor this coverage. The easiest method is to have local participants in a social marketing effort record instances in which they appeared on a news or consumer affairs show and instances of local news coverage of a social marketing activity. While this approach is fairly inexpensive, it requires considerable self-discipline on the part of community participants and can prove cumbersome if a program needs to monitor activities in a wide number of sites. Another approach is to purchase information about television news coverage from a commercial service.

c. Dissemination of Paid Ads and PSAs

A centerpiece of many social marketing campaigns is television and radio public service announcements. With broadcast media, it is important to monitor information on the timing and placement of broadcasts because television viewership and radio listening

varies dramatically during the course of the day. PSAs are most often used to fill unsold advertising slots that occur during the less popular viewing times. Knowledge of differences in the timing of PSAs and ads can help to assess the extent of campaign reach.

Data on differences in viewership over the course of day (Exhibit 3-2) derived from A.C. Nielsen indicate the importance of recording data on the timing of ads and PSAs. For instance, Nielsen data indicate that morning television viewership is approximately 10 percent of the total viewing audience. In the afternoon, viewership rises to about 15 percent of the total audience, although daytime viewing by women is somewhat higher. Viewership rises during the early evening, and by prime time (8:00 to 11:00 p.m.) television is viewed by nearly half of all adults. In the late evening (after 11:30 p.m.), viewership drops to around 18 percent. In radio, listening is highest during the morning "drive time," reaching about 20 percent of the target audience, and drops by the evening except for teenagers, who are evening listeners.

There are variations in viewership according to the placement of ads and PSAs as well. For instances, spots tend to be more heavily viewed if they are shown during a program than during a station break and are lower if a spot is shown within a series of advertisements rather than as the first ad in the series.

Information on the day, time, station, and placement of a spot can be used to estimate the total number of people watching or listening to a particular ad. Usually when broadcast time is purchased, the station (or the media firm that arranges for placements) can provide data estimates of the size and characteristics of an audience reached by an ad. Often the stations can provide information on the income distribution of their audience, enabling the development of estimates of the number of low-income viewers/listeners reached by the ads.

Process evaluation should focus on the proportion of the target audience reached over the course of a particular campaign (i.e., the cumulative reach of a campaign) and the frequency with which individuals view or hear a particular ad (i.e., the cumulative

frequency of exposure). Data on the frequency of exposure to campaign materials can provide an indication of whether an effort is achieving the level of activity needed to affect change in awareness, attitudes, and behavior.

Data on the viewership of television ads and the listenership for radio ads can often be obtained from the stations where the ads are broadcast. For television, two independent commercial monitoring services can also arrange to monitor ad placement in the 75 largest media markets in the country. It is also possible to obtain information through station logs and local monitoring, but the time and expense with this are such that it is often preferable to rely on commercial monitoring sources.

Monitoring will provide information on the "gross rating points" of the percent of a target audience viewing a particular ad at a particular airing. It is common to sum these various ratings over time to generate estimate of the "total impressions" (i.e., the number of times that members of a target audience were exposed to an ad). It also is possible to develop algorithms to estimate the cumulative reach and frequency of exposure. Such algorithms are sometimes available from media ad placement firms, and a set of algorithms for translating information on gross rating points into indicators of cumulative exposure has been developed for media campaigns (Maloney and Hersey, 1984).

It also can be useful to gather process information about factors that result in high levels of broadcast exposure in order to provide feedback to improve this aspect of social marketing. Prior process evaluations on this topic have found that involving local groups in contacting station managers and acknowledging the contributions of stations were associated with high levels of broadcast (Maloney and Hersey, 1984).

d. Monitoring Print Distribution

In many social marketing campaigns, newspaper coverage, feature articles, and print advertisements can provide a valuable medium for nutrition education. The advantage of newspaper stories is that they can provide more information than a 30-second television

ad or news story can. Moreover, since print material can be saved, it can be read more than one time or forwarded to other individuals. Newspapers can be a particularly useful medium for reaching older audiences. Newspapers typically can provide an estimate of their readership for an estimate of reach (recognizing that not all readers read all articles). Hence, documenting the distribution of print coverage can be quite useful in evaluating a social marketing effort.

In addition, a number of social marketing efforts feature print materials that are distributed to a target audience. Examples of such materials include pamphlets, brochures, recipe guides, and posters. In addition, several State nutrition networks mail newsletters to Food Stamp participants. It can be helpful, then, to track the distribution of materials in the various State and local channels in which they are used. Tracking the distribution of print material requires close collaboration with local volunteer and agency groups participating in a social marketing campaign. While the task can be formidable, a hallmark of a well-organized community effort is its ability to provide documentation of distribution as part of its own organizing effort. It often is possible to develop tracking systems that can serve to support both campaign and evaluation needs.

While the distribution does not mean that the materials were adequately or appropriately used by different groups, distribution is a first step in reaching a target audience. Knowledge of distribution efforts can provide useful information on the ways in which materials are distributed and the types of audiences that may have been reached.

One of the challenges in assessing distribution of print materials is determining if the target audience uses the materials. Print must be read, comprehended, and accepted by a target audience before it can be acted upon. One process evaluation by the State of Washington conducted a survey of a sample of food stamp families with children who had received newsletters on nutrition in order to learn whether recipients of materials liked and read the newsletter they received.

e. Documenting Distribution via Kiosks and Interactive Information Systems

Many emerging technologies, such a kiosks, can benefit from two levels of tracking information dissemination. First, it is helpful to document the number of sites, the types of sites, and the length of placement of kiosks in such settings. Second, analysis of information maintained by the kiosks or computer-based technology can provide information about the number of times the kiosks were used and the number of times each module was requested. Depending on the design of the kiosk, it is sometimes possible to embed a question that describes the demographic characteristics of the user and obtain feedback about immediate knowledge gains and intentions to use the information provided.

It also can be important to document the process by which such kiosks and interactive information systems are used. For instance, a process evaluation of an informatics tool in a health maintenance organization (HMO) found that the effects of the tool on selection of medical treatment alternatives, while significant in the initial year of the program, disappeared in the second year of the program. This reduction was related to a change in clinic procedure that required patients to make a second visit to the clinic location to use the information tool (Hersey, Matheson, and Lohr, 1998).

f. Documenting Information Requests

Similarly, it is often possible to build in information reporting systems as part of toll-free hot lines that allow analysis of information about the location (e.g., area code) and interests of a caller. It is also possible to route a random sample of callers to an interviewer to obtain more detailed information about the caller's characteristics and perceptions of the information provided. It also might be possible to compile a list of the types of information requested and the zip codes on mailing labels, which can be related to the economic population characteristics of the census tracts within those zip codes. Some States have conducted surveys of information requesters as a way to gain a better understanding of the factors that influenced people to respond to materials advising individuals of how to obtain nutrition information.

g. Documenting Community Activities

A major focus of social marketing activities is on community events and direct education of or services to low-income populations. Such information is important to gather a full picture of a social marketing effort. The source of this information is likely to be records maintained by community participants in the social marketing effort. In some programs, the information is routinely reported as part of management information systems such as the EFNEP report system (e.g., ERS_4), which provides information on the delivery of individual, small-group education to various audiences and the conduct of community events.

It is sometimes useful to supplement information on program activities with more complete information on the potential population who could be reached by an effort. For instance, the Maine Nutrition Network recorded that it trained teachers to integrate unutrition education modules in their regular classroom activities in more than 50 school settings, which composed 45 percent of the school districts in which more than half of students received free or reduced-price school lunches.

h. Assessing the Cost-Effectiveness of Information Channels

Process evaluation can help to assess the effectiveness of various distribution channels in reaching a target audience. For instance, one question sometimes raised about social marketing is whether the use of mass media can effectively target a food stamp population. Findings from the Kent County, Michigan, pilot test indicated that the campaign there reached a low-income audience (Holaday, 1999). The pilot test employed a television ad, which ran 394 times on cable channels (i.e., USA, TNN, BET, Nickelodeon, and Lifetime), along with billboards in 20 locations, 200 bus posters, newsletters, and take-home information on lunch menus in schools in low-income neighborhoods.

A random-digit-dialing telephone survey of 800 adults found that the campaign had achieved a high level of awareness, particularly among low-income adults. The campaign

message was recalled by 52 percent of adults with income below $20,000 (combining measures of unaided and aided recall). This was significantly higher than the rate of recall (42%) among adults with higher income levels. The campaign message "Eat healthy, your kids are watching" was recalled unaided (or "top-of-mind" awareness) by more than twice as many low-income adults (7%) and all adults (5%) as the national "Got Milk" campaign, which had run for more than a year and was recalled by about 2 percent of low-income and higher-income adults.[1] Further analysis indicated that school lunch menus were a particularly cost-effective way to reach low-income adults, although process evaluation observed that the campaign materials were printed on the back of the first menu of the school year, and it is not certain that later menus would achieve equal levels of attention. Newsletters, billboards, newspapers, and TV ads were about equal in cost-effectiveness. Posters and bus signs were the least cost-effective medium.

4. Documenting Policy, System, and Environmental Change

In addition to immediate process evaluation, an evaluation may benefit from assessing the impact on some of the intermediate outcomes that can influence ongoing behavioral change. These include efforts that result in policy, system, and environmental changes.

a. Policy Change

Policy change refers to changes in governmental or organizational policies or regulations that can result in increased access to nutrition education or nutritious food. Policy includes "local, State, Federal policies and laws that regulate or support healthy actions and practices for disease prevention, early detection, and management" (Glanz and Rimer, 1995). This group of activities can also include both mandatory and voluntary changes in regulations (McKenzie and Jurs, 1993). An application of policy work in health promotion programs can be through efforts concerned with ensuring that institutions work the way they should. For example, nutrition advocates were instrumental in convincing

[1] Since January 1995, the National Fluid Milk Processors Promotion Board has spent approximately $80 million per year on television and print ads featuring athletes and celebrities sporting a "milk mustache."

the U.S. Food and Drug Administration to change food label requirements to provide more detailed information on nutritional status of foods (Amidei, 1991; Wallack, Dorfman, Jernigan, et al., 1993). State-level examples of policy change include work by the Pennsylvania Department of Health to convince the Blue Cross/Blue Shield program to cover nutrition education for patients with diabetes (Hersey, Hare, Roussel, et al., 1994).

In New Jersey, the Department of Labor changed its policy to allow nutrition education to be included as part of the Employment Services' (ES) job skills/life skills workshops. A memorandum documenting this change in policy was sent to all ES regional managers and ES local office managers.

A local California program has been working with local teenagers to advocate education to allow delivery of nutritious foods in schools. This can include use of yogurt and low-fat foods in school cafeterias. In Georgia bottled water can now be sold, in addition to soda, in several schools districts. Students also are allowed to bring water to class. Network staff report that water has now become a "socially accepted" beverage to consume instead of nonnutritious soda beverages.

b. System Change

System change refers to changes in organizational or community practices that can result in increased access to nutrition education or nutritious food. One example of system change occurred in Brown County, Wisconsin, where a community hunger coalition looked at factors the affected access to a variety of reasonably priced foods. As part of this effort, the city transit authority redesigned bus routes to make stops near supermarkets. Another example of system change includes the standard monitoring and reporting of the delivery of nutrition education to patients with diabetes in managed care organizations.

c. Environment Change

Environment change refers to changes in the physical environment that can influence nutrition behavior (Perry, Baranowski, and Parcel, 1990; McKenzie and Jurs, 1993). In the context of nutrition, primary environmental factors include access to nutritious foods and information regarding nutrition. Community-based social marketing efforts often have worked with local stores to increase access to low-fat milk in stores.

In order to increase the access of low-income families to fresh fruit and vegetables the Maine Nutrition Network has worked to encourage farmers markets to accept Food Stamps. An assessment of Food Stamp redemption data indicates that in the year-following the implementation of the recruitment effort the dollar value of Food Stamps redeemed at farmers markets increased by 15 percent.

Reporting on policy, system, and environmental change is perhaps best communicated through a description of what change was effected. Given the variety of changes that can be affected in this area, it is difficult to anticipate the best way to categorize all of these changes. The best guidance for reporting may be to ask for a description of the change that was affected and a brief narrative description of what action lead to those changes. An example of a format to present this information as a case study, "Writing your success story" developed by the University of Wisconsin, is shown in ***Exhibit 3-3***. It is also possible to develop other innovative methods to disseminate this information. For instance, the Brown County example was conveyed through the use of a short video that described the project and the impact of the programs with an immediacy that would be difficult to replicate with a printed report.

Exhibit 3-3. Suggestions for Writing a Success Story!

What makes a great success story? As the saying goes, if anything's worth doing, it's worth doing well! This sheet will provide you with tips on how to make your success stories more interesting.

- **Who are the main characters of your story?** If you are writing about children, then pick out a few unique cases in which certain children indicated they had learned from the program. Tell the story of these specific children.
- **Who presented the program?** Don't forget to give credit where credit is due! This includes the person who gave the program as well as the unit with which that person is affiliated.
- **What was learned?** If the main character(s) indicated they learned something from the program, what was it? What did they *not* know before the program that they knew after?
- **How was the program conducted?** Tell how this information was presented. What was unique about the program that caught people's interest?
- **When did the program occur?** Was your program at night? Was your program in the Spring?
- **Where was the program held?** Was your program given at the local YMCA? Also, remember to include the name of the county in which the program was held.
- **How many people attended the program?** If 100 8th graders attended your program, say so.
- **What was said?** Stories are more interesting if you can use quotes. Not only quote the main character(s) of your story, but spice things up by including the comments of participants.

Some things to remember as you write your success story:

- **Use examples!** Don't just tell us the seniors were interested in your program; tell us how! Did they laugh? Did they ask a question which showed they were interested? What was the question?
- **Use simple language!** When you write, stick to language that everyone understands. A junior high student should be able to read and understand your story.
- **Use details!** Don't ever assume your reader knows what you are talking about! For example, don't just tell us that a new recipe tasted wonderful; tell us what made that recipe wonderful.

Good Luck! And may all your success stories be **great success stories!!!!!**

Source: "Reporting Results" section of University of Wisconsin Cooperative Extension Service Web site (http://www.uwex.edu/ces/wnep/).

CHAPTER 4

Impact Evaluation

A. Challenges in Evaluating Behavioral Impact

The heart of any evaluation of a social marketing initiative in nutrition is the measure of impact on the individuals toward whom the program is directed. The ultimate goal of any social marketing campaign is to bring about changes in behavior; these may include the FSNEP behaviors related to food resource management, food safety practices, dietary quality, and food security among food stamp participants and families in similar economic circumstances.

Impact evaluation faces two main design challenges. The first is the tendency to incorrectly ascribe impact to a social marketing effort when other factors may represent the true cause of change. For example, in the area of nutrition, an event other than a social marketing effort, such as publicity about a new nutrition finding, can affect nutrition behavior.

A second danger that an evaluation must avoid is failure to detect actual impacts. There are a number of possible causes for this failure. The most common error is use of inappropriate measures of change, such as focusing on long-term, general behavioral effects while failing to assess the intermediate changes in beliefs or specific behaviors that might be the critical impacts of a campaign. Such errors can be compounded by policy staff using the results of an evaluation as a final verdict rather than as an opportunity to learn. Successful social marketing is a continuing process in which incremental efforts build on the knowledge gained in earlier work. The ability of an evaluation to skirt these twin dangers and effectively guide ongoing health education efforts depends on three factors:

- Assessment of the social marketing implementation process,
- Selection of appropriate measurements, and
- Rigorous research design and analysis.

B. Assessing the Implementation of Social Marketing

To establish the impact of a social marketing effort, an evaluation must obtain high-quality data about the nature and extent of campaign implementation. As described in the prior chapter on process evaluation, data about the exposure of various audiences to campaign messages compose the independent variables in any analysis of the effects of social marketing. These data are the indicator that the intervention is delivered to the target audience.

No impacts may suggest that the social marketing effort failed to obtain the exposure needed to effect behavioral change. A number of social marketing campaigns dependent on donated media time, or required to purchase advertising on limited budgets, have experienced minimal audience exposure. An early example of this was the "Feeling Good" television series, which covered nutrition and other health topics. The series significantly affected nutrition knowledge and attitude of the few low-income people who watched the series (Mielke and Swinehart 1976); however, its overall success was quite low because it only reached one to two percent of the viewing audience.

In general, an assessment of the extent to which various audience segments were exposed to campaign materials is essential to understanding impact results. Specifically, it is important to know if the social marketing campaign reaches enough people to have a discernable impact? Was it properly implemented for the target audience? These questions typically are addressed through a process evaluation, which can indicate the most effective ways of disseminating nutrition materials as well as mobilizing community activities and support. Such information is critical in helping to plan future waves of a social marketing campaign.

C. Measurement Selection

A second element in the design of a useful evaluation is the selection of appropriate measurements. There are four essential steps to selecting appropriate measurements:

- Clarifying the information and decision making needs of policy and program staff,
- Specifying the rationale underlying a social marketing effort,
- Considering the theoretical framework which can guide the measurement of key elements within that rationale, and
- Selecting appropriate measurement tools.

1. Clarifying Information and Decision Making Needs

An initial step in the measurement development process is to clarify the goals and objectives of the evaluation. In some cases, the emphasis of an evaluation will be on the question, "Did it work?" In such an instance, it may prove useful to detail the areas in which a social marketing campaign worked:

- How many individuals in a target audience were exposed to messages?
- What were the effects of the social marketing effort on the knowledge and awareness of the target audience?
- What were the effects on attitude and intentions?
- What were the effects of the social marketing effort on behavior?

Because resources are typically limited, it is important to clarify the objectives and thus the information needs of program staff.

2. Understanding Campaign Rationale

Questions about which areas of impact to assess are best answered by careful analysis of the conceptual framework or rationale underlying a social marketing effort. The rationale is essentially a series of assumptions of causality that underlie an approach (as discussed earlier in

Chapter 1 Section C). Development of such a rationale is an essential first step in the planning of any social marketing effort. In many cases, such a rationale may not have been clearly specified. One of the benefits of evaluation, therefore, is that it helps program staff to clearly understand the rationale underlying an effort.

An important component of campaign rationale is the types of behaviors that are the focus of a social marketing effort. Food Stamp Nutrition Education Programs (FSNEPs) are encouraged to report on evaluation of social marketing impacts in several broad behavioral areas. The broad behavioral areas include:

- Food Resource Management. Food resource management incorporates issues such as spending habits and shopping behaviors, including decision making patterns, shopping skills, and other factors that may influence shopping behavior.
- Food Safety. Food safety attempts to eliminate hazardous and risky behaviors that lead to food-borne illness. Four behavioral constructs have been identified: time/temperature control, personal hygiene, adequate cooking, and cross-contamination.
- Dietary Quality and Physical Activity. Different social marketing activities will focus on such activities as fruits and vegetables; lower fats; cereals and grains; calcium-rich foods; poultry, eggs, meat, fish, and beans; reduced sugar; and balancing diet and physical activity.
- Food Security. Food security was defined by the American Institute of Nutrition (1990) as

  > Access by all people at all times to enough food for an active, healthy life. Food security includes at a minimum: the ready availability of nutritionally adequate and safe foods and an assured ability to acquire acceptable foods in socially acceptable ways (e.g., without resorting to emergency food supplies, scavenging, stealing, or other coping strategies).

  Food security measures reflect individual perception in large part. Hence, investigation of the effects of social marketing initiatives on food security can benefit from inclusion of measures of behaviors that are expected to result in food security (such as planning to manage food throughout the course of a month).

- Behavioral Antecedents. Within these areas, behavioral antecedents may serve as effective indicators marking an impact of an educational program. The stages of change theory is a relevant model to use in framing an evaluation targeting specific behavioral changes. The five stages are precontemplation, contemplation, action,

maintenance, and relapse. Although individuals targeted by nutrition education programs may be in different stages, identification of the incremental movement through the cycle may provide evidence of the program's effectiveness at motivating people to make desirable changes in behavior.

In addition, as described in the prior chapter, FSNEPs have been encouraged to report on accomplishments to affect system, policy, and environmental changes that support behavior change in these areas.

From a practical standpoint, it will often be necessary to make choices about which of these areas to include. It is unusual for single programs to place equal attention on all of these areas, hence, the selection of measures for an evaluation should be guided by the focus of the social marketing effort. If resources permit, however, it can sometimes be useful to include both one area (e.g., fruits and vegetables) that is central to a social marketing effort and another area (e.g., low-fat sources of calcium) that may be less central. This will allow one to test the hypothesis that greater change should be observed in the area that is central to the focus of the social marketing campaign rather than to more general influences (e.g., social desirability of response) that might be expected to affect both areas approximately equally.

Within each area, decisions about which tools to select should be based on the tools' reliability, validity, length, and relevance to the focus of the social marketing activity.

3. Use of Theory to Inform Measurement Selection

In selecting outcomes measures to include, a sound theoretical model for behavior change that is supported by evidence from prior research is recommended. Since a well-supported model portrays the behavioral process being assessed in an evaluation, the use of such a model increases the probability that the effects of a social marketing effort will be detected, and therefore the relationships in the chain of impact from awareness to behavior will be established.

The choice of impact measures should be informed by the theory of behavior change that underlies an intervention. A number of theories, such as the health belief model (Mullen, Hersey, and Iverson 1987; Rosenstock 1974), the theory of reasoned action (Ajzen and Fishbein 1980; Griffith, Neuwirth, and Dunwoody 1995), and social cognitive theory (Bandura 1986, Maibach and Parrott 1995) are helpful in describing the types of beliefs that can be influenced by a message. Work in stages of changes is helpful in describing the steps in movement from awareness to contemplation of change to behavior (Prochaska and DiClemente 1984). The PRECEDE model (Green and Kreuter 1991) is useful in describing the role of predisposing, enabling, and reinforcing factors needed to maintain behavioral changes in the nutrition area. For instance, a communication persuasion model (McGuire 1989, Parrott 1995) suggests the importance of the following elements:

- Source Factors: Characteristics of the individual who presents the message;
- Message: Characteristics of the message itself, such as the organization, repetition, and type of appeal;
- Channel Factors: Variables that have to do with the medium through which a message is transmitted;
- Receiver Factors: Variables that involve the characteristics of the target audience; and
- Destination Factors: Variables that have to do with the target behavior at which a message is aimed.

An evaluation that provides measures of these different factors can prove very useful in refinement of current messages and the design of future message strategies.

The choice of theoretical approach depends in large part on the particular aspect of the social marketing effort about which decisions will be made. A theoretical framework provides a far richer and more sensitive set of measures than might otherwise have been selected. A close link between theory and evaluation also contributes to practical and theoretical development of strategies to improve nutrition-related behavior among low-income populations.

4. Selection of Measurement Approaches

A final set of measurement issues involves the selection of appropriate measurement approaches. Four broad types: mail surveys, telephone surveys, in-person survey, and archival data exist, with clear tradeoffs between them (see Exhibit 4-1), and within each of these methods there are a variety of choices in terms of the use of closed-ended vs. open-ended responses. Closed-ended items (such as Food Frequency Questionnaires or aided recall of campaign exposure) offer more standardized responses but may miss the depth provided in a open-ended questionnaire. On the other hand, open-ended responses (e.g., 24-hour dietary recall) provide more inclusive data but more variable response data. Recent advances in-person technology such as audio-computer-assisted personal interviewing (ACASI) tend to yield data that provides more accurate data about sensitive health topics (e.g., substance use) than lower cost methods (Turner, Lesseler and Devore 1992; O'Reilly, Hubbard, Lessler, et al. 1994; Turner, Ku, Sonenstein, and Pleck 1996; Turner, Forsyth, O'Reilly, et al. 1999). Visual displays (e.g., video clips of TV ads or pictures of serving sizes) that can be incorporated in ACASI interviews can improve accuracy of reporting. Although the advantages of ACASI in terms of the accuracy of nutrition related behavior has not been investigated, these methods have applicability to low-income linguistic minorities (Hendershot, Rogers, Thornberry, et al. 1996) and to telephone survey methods (Turner, Miller, Smith, et al. 1996).

Mail surveys, despite their appeal of low cost, often have serious limitations in terms of response rate. Nonetheless, mail surveys can have some use if questionnaires can be kept very brief and if an incentive can be offered to motivate response rate. E-mail surveys also can be used in surveys of professionals or network members (but are generally not applicable to low-income audiences). Telephone surveys often offer a reasonable compromise for a general population. In some areas, however, it may be necessary to employ a combination of telephone and in-person interviews because low-income populations can have lower rates of telephone ownership. Mixed mode surveys also can offer advantages in terms of cost and flexibility.

Exhibit 4-1. Comparison of Data Collection Approaches

Criterion	Approach			
	Mail Surveys	**Telephone Surveys**	**In-Person Surveys**	**Archival Data**
Reliability	low/med (response bias; order, etc.)	mod/high	mod/high	mod/high
Representativeness (Response Rate)	low	mod/high	mod/high	mod/high
Breadth of Topics	low (10 mins)	moderate (20 mins)	very high (40 mins)	very low
Response Time	moderate	quick	moderate	slow
Cost	low	moderate	high	low

5. Selecting Appropriate Measurement Tools

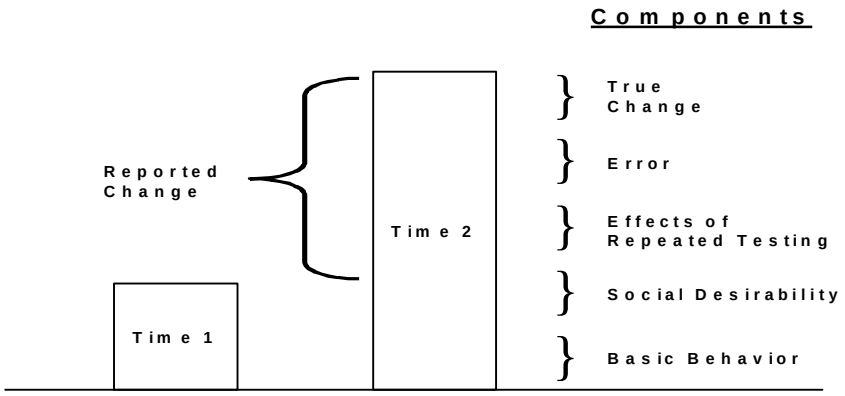
The goal in measurement selection is to choose an instrument with established reliability and validity of information collected. Reliability refers to the ability of an instrument to yield similar results if taken a second time. Reliability is often measured in terms of internal consistency (coefficient alpha: Cronbach 1951), from 0.0 for nonreliable measure to 1.0 for a completely reliable scale. In general, an internal consistency of 0.7 or higher is considered to be a fairly reliable instrument (Nunnally and Bernstein 1994).

Validity, on the other hand, refers to the accuracy with which a measure indicates the actual construct. For instance, validity is often expressed in terms of the correlation between a report of food intake with biochemical indicators of nutritional status.

Exhibit 4-2 shows the sources that might influence an increase in scales of behavior between Time 1 and Time 2. The Time 2 score is influenced by the initial behavior of an individual. Indeed, because longitudinal samples control for this by conducting analysis over time in the same individual, they require a smaller number of cases to detect effects as statistically significant than do designs which use different samples of respondents at baseline and again at follow-up.

Another source of potential change is random error in answering a question. Typically, scales comprising more items are more reliable than shorter scales because random errors in answering any individual item tend to cancel themselves out. On the other hand, part of a change can be associated with social desirability, particularly if an intervention helped make people more aware of the correct answer to a question about dietary behavior. Finally, there is a true change component. The goal in constructing scales is to increase the accuracy of a scale by minimizing the influence of social desirability. As a rule, longer scales are more reliable and more sensitive to change, but the added length of the data collection tool can adversely affect response rate. Often an evaluation is well-served by careful selection to gather information on a focused set of topics. Information about the reliability, validity, sensitivity to change, and cultural relevance of measurement tools can inform decisions about measurement selection.

Exhibit 4-2. Components of Measurement Change



D. Research Design

Effective research design is a function of the issues to be investigated, the resources (time and money) available for a study, and the opportunities that avail themselves for investigation. Since those decisions need to be made within the context of a particular nutrition education program, it is not possible to recommend single design for the evaluation of all social marketing efforts. Rather, this section discusses the more basic issues that should be considered in approaching questions of design.

There are two prerequisites for establishing a social marketing effort as the cause of observed effects. The first is demonstrating covariation between independent and dependent variables. This is improved by the following:

- Careful Measurement of the Extent of Exposure to Nutrition Messages in Various Groups. Such measurement provides the independent variable in the analysis. The more accurately exposure can be measured, the greater the possibility of detecting campaign effects.

- Sensitive Measurement of Effects that Are Likely to Occur. Issues regarding the selection of appropriate measurement have been discussed in the prior section.

- Adequate Sample Size and Response Rate. Larger sample sizes increase the ability of a study to detect the effects of a social marketing effort because the random error of measurement and individual background begins to be balanced out across individuals and the effects of the common effects of exposure to a message can be seen. Similarly, high response rates offer greater assurance that findings from a sample generalize to a broader audience. Since larger samples cost more, decisions typically involve questions about the minimum sample sizes requirements to detect a change as statistically significant. The sections that follow offer guidance on this topic.

The second prerequisite for causal attribution is the ability to eliminate plausible alternative explanations for any observed results. This is accomplished by judicious formulation of a research design.

1. Length of Time between Data Collections

One design issue that may best be determined by specific social marketing endeavors is the length of time between baseline and follow-up. Sufficient time should be allotted for the social marketing effort to have been fully implemented. On the other hand, too long a time interval could result in a level of turnover that would increase the possibility for a seasonal change or historical event (e.g., a change in food stamp outreach activities) to influence study results. In light of these considerations, the appropriate time interval between baseline and post-intervention data collection should be between three and six months.

2. Statistical Power

Questions about desired statistical power ultimately depend on informed judgment about the amount of change that could be expected. To some extent, the magnitude of change that is considered important will vary with the nature of the variables measured. For instance, a small change in behavior such as consumption of servings of fruits and vegetables, may have greater practical significance than a larger change in the awareness about eating a variety of foods. Once decisions have been made about the magnitude of change that one desires to detect, it is then possible to select a sample size with the statistical power to detect such changes.

If no prior information exists, it is useful for sample size of an evaluation to be large enough that a study is able to detect changes on the magnitude of between one-tenth and three-tenths of a standard deviation. The authors' preference is to select two-tenths (0.2) of a standard deviation as a desirable level for the magnitude of changes that can be detected in a field evaluation study of a social marketing campaign. This would require a sample size of approximately 200 in each group. Alternative solutions to this equation are shown in Exhibit 4-3.

In cases involving a longitudinal samples, a method for estimating sample sizes is provided in Appendix A.

Exhibit 4-3. Number of Cases per Cell Needed for Various Levels of Statistical Power

Size of Differences in Means that Can Be Detected	Number of Cases per Cell
0.10 standard deviations	800
0.20 standard deviations	200
0.25 standard deviations	128
0.28 standard deviations	100
0.30 standard deviations	89

3. Estimating Response Rates

Accurate results rest on the premise that the sample drawn from a study is representative of the targeted population. When collecting data, regardless of the method, a high response rate is desired. The response rate is defined by the number of persons providing information divided by the number of eligible reporting units in the sample as shown in Exhibit 4-4. The greater the percentage of surveys completed, the more likely the returned information in the sample is accurately representing the larger targeted population.

Exhibit 4-4. Method for Calculating Response Rate

The response rate (RR) is the number of complete interviews with reporting units divided by the number of eligible reporting units in the sample. The calculation is as follows:

RR1 = CI/ CI+ PI+R+NC+O

Where:

CI = *completed interview*

PI = *partial interview*

R = *refusal and break off*

NC = *noncontact*

O = *other, unknown*

The response rate above (RR1) is a preferred response rate for household surveys, although it is also a minimum response rate.

In telephone surveys, it is common to estimate what proportion of cases of unknown eligibility (i.e., "Ring-No-Answers") are actually eligible. One approach to this estimation is to assume that the proportion of eligible and ineligible cases among the cases whose eligibility status is known would also apply to the cases of indeterminate eligibility (Lessler and Kalsbeek, 1992). In this case, the response rate would be:

RR3 = CI/ (CI+ PI) + (R+NC+O) + e(UH+UO)

Where in addition to the terms defined above:

In planning research studies it is important to emphasize the need for high response rates. In many cases it can be more important to devote the resources associated with follow-up to obtain an acceptable response rate (and to encourage continued participation of individuals in a longitudinal study) than it is to further increase sample size. Without a meaningful response rate, it is very difficult to generalize study results. Typically, one seeks to obtain a response rate in the 80 percent rate for in-person data collection efforts, and in the 70 percent range for telephone and mail surveys. If response rates drop below 50 percent to 60 percent, it is difficult to draw conclusions about the generalizability of the effort, even if one has conducted analysis to assess differences between the characteristics of respondents and nonrespondents. Response rates and attrition rates should be reported separately for the intervention and the control group in a study.

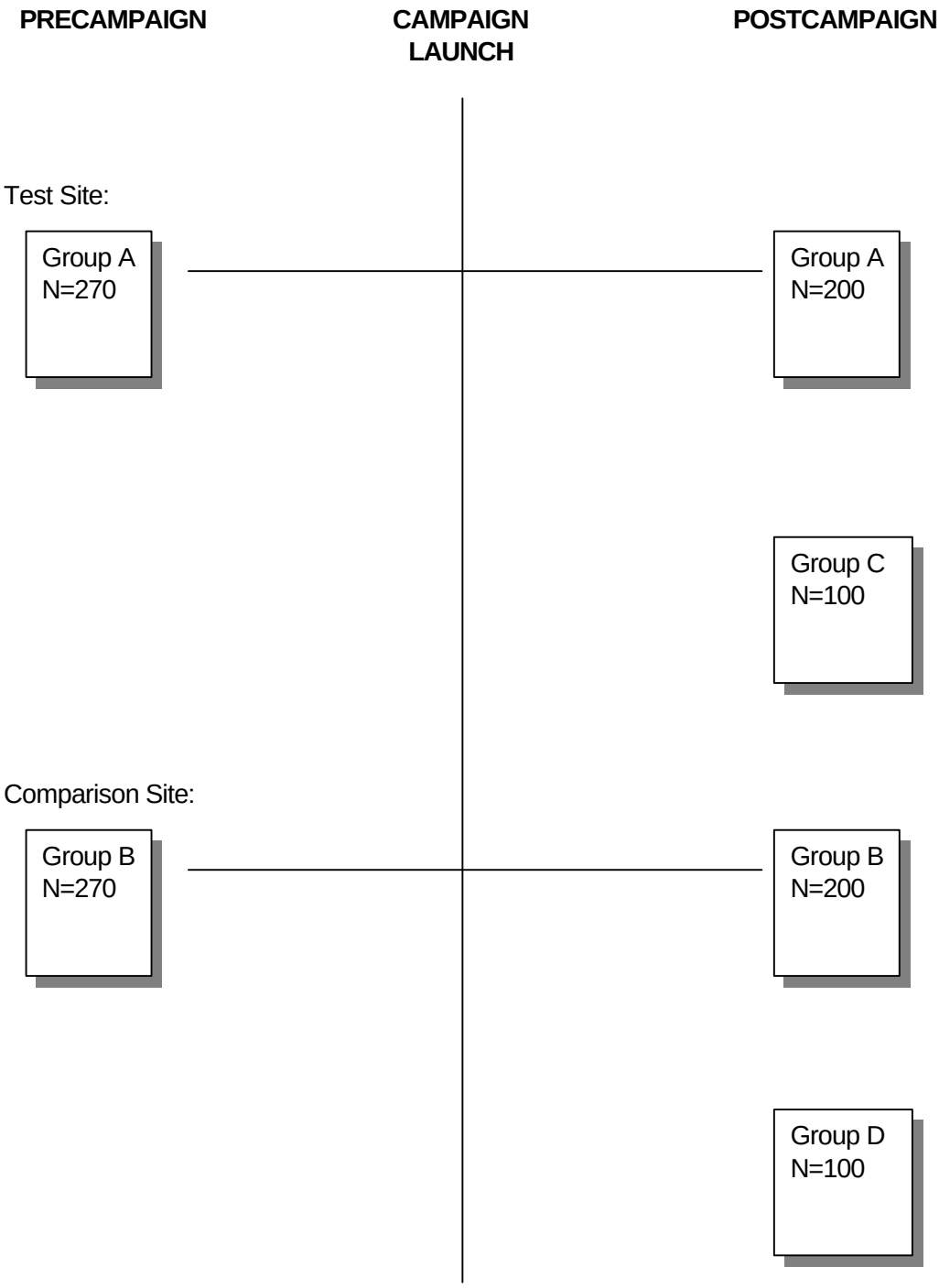
Response rate can be enhanced by keeping the survey abbreviated, increasing the attempts to reach participants, providing multilingual surveys, offering incentives for participation and, in the case of mail surveys, providing prepaid postage (Wholey, Hatrey, and Newcomer 1994).

E. Eliminating Plausible Alternative Explanations

To establish valid campaign impact, the evaluation design must have the ability to rule out alternative explanations which might otherwise account for observed results. As characterized by Campbell and Stanley (1963), this involves removing threats to the internal validity of a study.

As a context for this discussion of validity threats, one might consider a hypothetical example from a commonly employed evaluation design. This design (shown in Exhibit 4-5) involves precampaign interviews six months apart with a randomly selected panel of low-income adults in two communities: Site A, which experienced a campaign, and Site B, a "matched" comparison site without a campaign. This "nonequivalent control group" design is a relatively powerful design that illustrates design that the validity threats most likely to be encountered in any

Exhibit 3-6. Sample Design Illustration



evaluation of a public education campaign. The design collects information from three groups of respondents:

- Group A. A Pre-/Postcampaign Panel. This group consists of a random sample of 270 adults in a campaign site. There is 25 percent attrition in the sample so that postcampaign interviews are conducted with only 200 respondents.

- Group B. A Pre-/Postpanel Comparison Group. This group consists of a random sample of 270 adults in Site B, a closely matched comparison site where there is no campaign. Again there is 25 percent attrition so that post-test interviews are conducted with only 200 respondents.

- Groups C & D. A Post-Test-Only Comparison Group. These groups consist of a random sample of 100 adults in the intervention site and 100 adults in the comparison site who are interviewed only at Time 2. The addition of this group helps to provide a control for the effects of repeated interviews.

The illustration assumes that a study has collected data on the percent of low-income adults who consumed five or more servings of fruit and vegetables a day. The percent of panel respondents (Group A) who reported five or more servings of fruit and vegetables a day increased from 35 percent to 50 percent in the intervention site; the percent of respondents in the panel group from the comparison site (Group B) agreeing with this statement increased from 25 percent to 30 percent, and 45 percent of respondents in the post-test only group in the intervention site (Group C) and 25 percent of respondents in the post-test-only group in the comparison site (Group D). With a sample size of 200 per cell at Time 2, a difference in the changes between Group A (15 percent increase) and Group B (5 percent increase) would be statistically significant. If implementation analysis indicated a fairly intense campaign in Site A, it would be tempting to ascribe this change to the impact of the social marketing campaign.

Before cheering the campaign success, however, it would be wise to rule out other possible explanations of these findings. Three main types of validity threats would bear particular close attention:

- Differences in the selection of respondents,

- Problems related to repeated interviews, and

- Extraneous factors occurring with the passage of time.

1. Selection Bias

Perhaps the most serious threat to this design is bias resulting in differential selection of respondents for the comparison groups. Selection bias in this design can result either from differences in the populations of the two sites, or from differences in the way respondents were selected in the two sites. For instance, differences in the educational background of groups could result in greater exposure to messages about consumption of fruits and vegetables. Comparisons of the precampaign differences in the background of Group A and Group B respondents provides a test for the significance of selection bias.[2]

2. Problems Relating to the Use of Repeated Interviews

This design is subject to two problems that can result from using repeated measures: differential attrition and the effects of repeated testing.

a. Differential Attrition

The differential loss of respondents from comparison groups could also account for differences in the changes between groups. In the example described above, post-test differences in consumption could result if more males than females failed to participate in the second interview at one of the sites, while in another site, there was no difference between men and women in attrition rate. If there were differences in the eating patterns of men and women for the behavior of interest, then this differential attrition rate could account for the observed differences at Time Two between respondents in the two sites.

One way to test for the influence of differential attrition is to compare the pretest scores of dropouts and nondropouts and investigate the differences in the respondents who dropped out in the two sites. In the example, there was little difference between the

[2] Problems introduced by the possibility of selection bias becomes particularly severe in a panel study where groups are selected for treatment or control groups on the basis of pretest scores, since subsequent changes can be influenced by regression artifacts. For instance, if individuals at nutritional risk were selected for participation in a community involvement program, while individuals who were not at nutritional risk who were placed on a waiting list were selected as the control; changes could occur simply because individuals at greater nutritional risk had a greater opportunity to improve.

Time 1 scores of the subsequent dropouts from the two sites (25 percent of study dropouts in Site A and 20 percent of study dropouts in Site B consumed five or more servings of fruit and vegetables a day) and in both cases the dropouts were about 5 percent below the nondropouts in their respective sites.

b. Effects of Repeated Testing

In many instances, having taken an interview once can influence responses in a subsequent interview. For instance, asking people about nutrition ads can cause them to pay closer attention to the ads they do see. Questions about the frequency of consuming fruits and vegetables can also cause people to become more aware of their intake, thus yielding more accurate information in a subsequent interview. In the example above, the fact that changes are bigger in the campaign site than in the comparison site suggests that the effects of repeated testing are minimal. The influence of repeated interviews can be tested by comparisons of post-test only groups (Group C and D) with the post-test scores of the panel group (Group A_2 and B_2) in the same site.

3. Extraneous Factors Related to the Passage of Time

Other validity threats are concerned with what happens during the passage of time. Such effects could be associated with:

- Maturation. With some groups (e.g., children or older adults), change in age can itself influence changes in attitudes and nutrition behavior. While six months, the length of time between interviews in this example, may be too short to expect maturational change, there are periods, such as the transition to or from employment, when six months can have an important impact on a person's nutrition practices. The best control for maturation effects is the presence of a comparison group of the same age.
- Seasonal Trends. Dietary practices are influenced by the time of year (e.g., increased availability or lower cost of fruits and vegetables in different seasons, or increased consumption of comfort foods during holidays). Evaluation must therefore be sensitive to changes which might be related to the season of the year. As a first precaution, known seasonal variations should be avoided when scheduling data collection. In addition, a comparison of Group C with the pretest scores of Group B, provides one way to test for the possible effects of seasonal trends; a comparison of findings with trends in other studies is another.

- Historical Changes. Patterns of nutrition behavior have changed dramatically during the past decade. In general, the threat of historical trends increases with the length of an evaluation. However, the changes can arise, in just a few months if, for example, there was a new discovery concerning the health benefits of a particular food.
- History-Selection Interaction. Particularly difficult problems of interpretation can arise from events, such as promotion of a particular type of food item by supermarkets in a market, or the publicity associated with another health promotion campaign, which occur in a particular site.

When a social marketing effort is evaluated in just one site, it is virtually impossible to disengage the effects of a campaign from the effects of other events which might have occurred in a community. What can be done, however, is to have maintained a close enough track of the implementation process to be able to document possible community events which might have influenced changes.

In the illustration, as in most cases where alternatives are carefully investigated, the findings of campaign effects would appear to withstand the scrutiny of a close investigation of alternative explanation of effects. The pretest scores of the two groups are similar enough to suggest that effects of selection bias are small; there appears to be no artifacts that result from repeated testing or differential attrition; and the fact that the campaign group showed a significantly larger change in belief over time than did the comparison group suggests that the influence of the campaign is operating over and above the effects of any general historical or seasonal trends. The one possibility this design does not rule out is that some other series of events in Site A may have caused the observed effects. The only good way to control for the confounding influence of local events would be to conduct a multisite study to see if findings of campaign effects could be replicated in several locations. Nonetheless, the findings from this illustrative study suggest the merit of such replication. Indeed in most instances a well-executed design combined with a rigorous test of alternatives explanations tends to increase rather than obscure the chances of detecting effects of a social marketing effort.

F. Resources and Tradeoffs in Evaluation Design

1. Resource Tradeoffs

The earlier sections of this manual have described a number of approaches for designing an evaluation; this section addresses the issue of choosing among them. The first criterion of choice depends on the purpose of the evaluation. What emphasis should be placed on understanding the distribution process? What emphasis should be placed on refining message strategy? Answers to such questions provide the starting point for decisions about evaluation design. A second issue involves the tradeoff between the certainty one can place in the findings (internal validity) and their generalizability (external validity). While there is no inherent incompatibility between these concepts, limitations on the resources available for an evaluation normally impose a priority. The authors recommend giving greater emphasis to internal validity, that is doing a rigorous evaluation in a small number of sites, rather than a less rigorous evaluations across multiple locations.

A common issue concerns the perceived tradeoffs between feedback and accountability needs. Most researchers involved in social marketing recognize the priority that needs to be assigned to providing feedback to continually informing and improving subsequent activities. It is useful to note, however, that there are often a number of design decisions that allow satisfaction of multiple objectives. For example, a major premise of this manual is that it is essential to document the exposure of social marketing messages; not only does that help to quantify the independent variable for an evaluation, but it provides feedback to improve the distribution process. Similarly, inclusion of attitude items as well as information on behavior can help to establish the effects of a social marketing effort. Since the major cost of data collection is the initial contact with respondents, the inclusion of both attitudinal and behavioral items on a survey is a relatively modest cost.

Some steps can be done to undertake this integration in a more formal basis. For instance, the Maine Nutrition Network is conducting a survey of food stamp participants as part of their

needs assessment and planning activities, and these same individuals will be reinterviewed in six months so that the planning information provides a baseline for an evaluation effort.

2. Potential Enhancements to Research Design

There are a number of design enhancements across sites that also can be helpful. For instance, a design alternative would be a "rolling site" design, depicted in Exhibit 3-6, in which the control site for an initial study would serve as the intervention site at a later point in time (benefitting from the baseline information obtained in its experience as a control site). The design may be particularly appropriate for organizations with a large number of sites that, for logistical reasons, may need to phase in implementation over a period of time. In this situation, the rolling site design permits random assignment of different sites to treatment and control conditions, with the incentive to control sites that they will be able to implement the program later on. The rolling site design also contributes to the ability of a study to identify the effects of a program implementation, because it offers the opportunity to demonstrate that effects observed between Time 1 and Time 2 can be replicated between Time 2 and Time 3 when the program is implemented in the original control sites.

Exhibit 4-6. Rolling Site Design

Condition	Time 1		Time 2		Time 3
Group A Sites	A-1	X	A-2		A-3
Group B Sites	B-1		B-2	X	B-3
Group C Sites			C-2		C-3

Note: When multiple sites are available, this design could be adapted to include multiple sites in each condition. For instance, if there were nine sites available, then three sites could be included in the group of sites in which the social marketing activities were implemented after Time 1 (Group A); three sites could be included in the sites in which the social marketing effort took place after Time 2 (Group B); and three sites could be included in the sites in which the social marketing initiative took place after Time 3 (Group C).

Similarly, it is also possible to utilize the results from one set of sites to inform the development of enhanced evaluation efforts. The results from each of these initial evaluations can be used to

inform improved implementation of social marketing activities, thus enhancing the social marketing program.

G. Conclusion

In closing, one will note that this manual has not recommended a particular research design. Such a recommendation would be premature in the absence of information about the primary purposes of a study, the rationale underlying a social marketing effort, and the resources and opportunities available for conduct of an evaluation. Rather, this manual has attempted to outline the types of research issues and evaluation options that should be considered in planning an evaluation. Perhaps, more importantly, this manual has helped to describe a planning process that can lead to making informed decisions about the best evaluation design in a particular circumstance.

The planning for an evaluation begins with careful consideration of the information needs of decision makers and program staff. Such information can help to determine the points in a nutrition education process which should be the primary focus of attention. Next planning should consider rationale underlying the social marketing effort and draw on communications theory to see what is needed to test and refine this rationale. This work will help in deciding what should be measured and inform decisions about the most appropriate evaluation approach to use in measurement. After these issues have been decided upon, it will be possible to develop a research design which can maximize the ability to reach conclusions about the effects of a social marketing effort.

This manual has made a number of specific suggestions about evaluation approaches. For example, it suggests:

- Adequately documenting the dissemination of materials and delivery of social marketing messages,
- Measuring a range of dependent variables along the chain of exposure through behavioral change,
- Selecting a design that maximizes the internal validity of results, and

- Devoting sufficient resources to analysis of the implementation process of the social marketing campaign to understand how success was achieved.

The important point, however, is the emphasis on the process of reaching decisions about evaluation research. The effort in clarifying a research agenda and specifying the rationale for the social marketing effort is not an easy task. Yet it is that type of hard work and effort that can enable both an evaluation and a social marketing initiative to further our knowledge about the ways in which social marketing can have a positive effect on the nutrition-related behavior of Americans.

Bibliography

Ajzen, I., and M. Fishbein. 1980. *Understanding Attitudes and Predicting Social Behavior.* Englewood Cliffs, NJ: Prentice-Hall.

American Association for Public Opinion Research. 1998. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for RDD Telephone Surveys and In-Person Household Surveys.*

Andreasen, A.R. 1995. *Marketing Social Change.* San Francisco, CA: Jossey-Bass.

Balch, G.I. 1994. *Nutrition Education for Adults: A Review of Research.* Food and Consumer Service, U.S. Department of Agriculture.

Bandura, A. 1986. *Social Foundation of Thought and Action.* Englewood Cliffs, NJ: Prentice-Hall.

Campbell, D., and J. Stanley. 1963. *Experimental and Quasi-experimental Designs for Research.* Chicago: Rand McNally.

Caudle, S.L. 1994. "Using Qualitative Approaches." In *The Handbook of Practical Program Evaluation,* ed. J.S. Wholey, H.P. Hatry, and K.E. Newcomer, 69-95. San Francisco, CA: Jossey-Bass.

Cochran, W.G. 1977. *Sampling Techniques.* New York, NY: Wiley.

Contento, I., G.I. Balch, Y.L. Bronner, L.A. Lytle, S.K. Maloney, S.L. White, C.M. Olson, S.S. Swadener, and J.S. Randell. 1995. "The Effectiveness of Nutrition Education and Implications for Nutrition Education Policy, Programs, and Research: A Review of Research." *Journal of Nutrition Education* 27:279- 422.

Cook, T.D., and D.T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis for Field Settings*. Boston, MA: Houghton Mifflin Company.

Cronbach, L.J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16:297-334.

Fitzpatrick, R., and M. Boulton. 1996. "Qualitative Research in Health Care: I. The Scope and Validity of Methods." *Journal of Evaluation and Clinical Practice* 2(2):123-30.

Foerster, S.B., J. Gregson, L.D. Beal, J. Hudes, H. Manuson, S. Livingston, M.A. Davis, A.B. Joy, and T. Garbolino. 1998. "The California Children's 5 a Day—Power Play! Campaign: Evaluation of a Large-Scale Social Marketing Initiative." *Family and Community Health* 21:46-64.

Francisco, V.T., A.L. Paine, and S.B. Fawcett. 1993. "A Methodology for Monitoring and Evaluating Community Health Coalitions." *Health Education Research* 8:403-416.

Frey, J.H. 1989. *Survey Research by Telephone.* Newbury Park, CA: Sage.

Glanz, K. 1997. "Behavioral Research Contributions and Needs in Cancer Prevention and Control: Dietary Change." *Preventive Medicine* 26: S43-S55.

Glanz, K., A.M. Hewitt, and J. Rudd. 1992. "Consumer Behavior and Nutrition Education: An Integrative Review." *Journal of Nutrition Education* 24:267-277.

Goetz J., and M. Lecompte. 1984. *Ethnography and Qualitative Design in Educational Research.* New York, NY: Academic Press.

Goodman, R.M., K.R. McLeroy, A.B. Steckler, and R.H. Hoyle. 1993. "Development of Level of Institutionalization Scales for Health Promotion Programs." *Health Education Quarterly* 20:161-178.

Green, L.W., and M.W. Kreuter. 1991. *Health Promotion Planning: An Educational and Environmental Approach.* Mountain View, CA: Mayfield.

Griffith, R.J., K. Neuwirth, and S. Dunwoody. 1995. "Using the Theory of Reasoned Action to Examine the Impact of Health Risk Messages." In *Communication Yearbook,* ed. B. Burleson, 201-228. Newbury Park, CA: Sage.

Hendershot, T.P., J.P. Rogers, J.P. Thornberry, H.G. Miller, and C.F Turner. 1996. "Multilingual Audio-CASI: Using English-Speaking Field Interviewers to Survey Elderly Korean Households." In *Health Survey Research Methods,* ed. R.B. Warnecke. Hyattsville, MD: National Center for Health Statistics.

Hersey, J.C., S. Anderson, L. Bell, N. Eisen, M. Harkins, L. Harris-Kojetin, J. Matheson, L. Shiveley, C. Woodsong, and B. Zimmerman. 1999. *Evaluation of the Statewide Nutrition Education Network Cooperative Agreement Demonstration.* Washington, D.C.: Research Triangle Institute and Health Systems Research, Final Report to Food and Nutrition Service, U.S. Department of Agriculture.

Hersey, J.C., L.S. Klibanoff, D.J. Lamm, and R.L. Taylor. 1984. "Promoting Social Support: An Evaluation of California's Friends Can Be Good Medicine Campaign." *Health Education Quarterly* 11:293-312.

Hersey, J.C., M.L. Hare, A.E. Roussel, and M. Odell Butler. 1994. "Lessons from Implementing State-Based Diabetes Control Programs: An Evaluation." Prepared for the Centers for Disease Control and Prevention, Division of Diabetes Translation, and the Office of Program Planning and Evaluation. Battelle.

Holaday, R.M. 1999. "Evaluation of the Michigan Nutrition Support Network: Network Building and 'Eat Healthy' Campaign." *Final Report to Michigan State University Extension.* Lansing, MI: Holaday Research and Consulting.

Jones, R. 1995. "Why Do Qualitative Research? It Should Begin to Close the Gap Between the Science of Discovery and Implementation." *British Medical Journal* 311(6996):2 [editorial].

Kitzinger, J. 1995. "Introducing Focus Groups." *British Medical Journal* 311:299-302.

Kraemer, H.C., and S. Thiemann. 1987. *How Many Subjects: Statistical Power Analysis in Research.* Newbury Park, CA: Sage.

Lavrakas, P.J. 1993. *Telephone Survey Methods: Sampling, Selection, and Supervision.* Newbury Park, CA: Sage.

Lefebvre, R.C., and J.A. Flora. 1988. "Social Marketing and Public Health Intervention." *Health Education Quarterly* 15:200-315.

Lessler, J., and W.D. Kalsbeek. 1992. *Nonsampling Error in Surveys.* New York, NY: Wiley.

Lytle, L.A. 1994. *Nutrition Education for School Age Children: A Review of Research*. Food and Consumer Service, U.S. Department of Agriculture.

Maibach, E.W., and R.L. Parrott. 1995. *Designing Health Messages.* Thousand Oaks, CA: Sage.

Maloney, S.K., and J.C. Hersey. 1984. "Getting Messages on the Air: Lessons from the National Alcohol Abuse Prevention Campaign." *Health Education Quarterly* 11:273-292.

McGuire, W.J. 1989. "Theoretical Foundations of Campaigns." In *Public Communication Campaigns*, ed. R.E. Rice and C.K. Atkin, 43-66. Newbury Park, CA: Sage.

McLeroy, K.R., D. Bibeau, A. Steckler, and K. Glanz. 1988. "An Ecological Perspective on Health Promotion Programs." *Health Education Quarterly* 15:351-377.

Mielke, K.W., and J.W. Swinehart. 1976. *Evaluation of the "Feeling Good" Television Series.* New York: Children's Television Workshop.

Miles, M.B., and A.M. Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook.* 2d ed. Newbury Park, CA: Sage.

Mosteller, F., and G.A. Colditz. 1996. "Understanding Research Synthesis (Meta Analysis)." *Annual Review of Public Health* 17:1-23.

Mullen, P.D., J.C. Hersey, and D.C. Iverson. 1987. "Health Behavior Models Compared." *Social Science and Medicine* 24:973-981.

Nunnally, J.C., and I.H. Bernstein. 1994. *Psychometric Theory.* New York, NY: McGraw-Hill.

O'Reilly, J.M., M.L. Hubbard, J.T. Lessler, P.P. Biemer, and C.F. Turner. 1994. "Audio and Video Computer-Assisted Self-Interviewing: Preliminary Testing of New Technologies for Data Collection." *Journal of Official Statistics* 10:197-214.

Parrott, L.R. 1995. "Motivation to Attend to Health Messages." In *Designing Health Messages,* ed. E. Maibach and R.L. Parrott, 7-23. Thousand Oaks, CA: Sage.

Patton, M.Q. 1990. *Qualitative Methods in Research and Evaluation*. Newbury Park, CA: Sage.

Patton, M.Q. 1997. *Utilization-Focused Evaluation: The New Century Text*. 3d ed. Thousand Oaks, CA: Sage Publications.

Prochaska, J.O., and C.C. DiClemente. 1984. *The Transtheoretical Approach: Crossing the Traditional Boundaries of Therapy.* Homewood, IL: Dow Jones-Irwin.

Reichardt, C.S., and C.A. Bormann.1994. "Using Regression Models to Estimate Program Effects." In *Handbook of Practical Program Evaluation,* ed. J.S. Wholey, H.P. Hatry, and K.E. Newcomer, 417-455. San Francisco, CA: Jossey-Bass.

Reichardt, C.S., and T.D. Cook. 1978. "Beyond Qualitative versus Quantitative Methods." In *Qualitative and Quantitative Methods in Evaluation Research,* ed. T.D. Cook and C.S. Reichardt, 7-32. Thousand Oaks, CA: Sage.

Reichardt, C.S., and S.F. Rallis, eds. 1994. *The Qualitative-Quantitative Debate: New Perspective,* New Directions in Program Evaluation, no. 61. San Francisco, CA: Jossey-Bass.

Rosenstock, I.M. 1974. "Historical Origins of the Health Belief Model." *Health Education Monographs* 2:328-335.

Thompson, F.E., T. Byers, and L. Kohlmeier. 1994. "Dietary Assessment Research Manual." *The Journal of Nutrition* 124:2245S-2317S.

Thorne, S. 1997. "The Art (and Science) of Critiquing Qualitative Research." In *Completing a Qualitative Research Project*, ed. J.M. Morse. Thousand Oaks, CA: Sage.

Turner, C.F., B.H. Forsyth, J. O'Reilly, P.C. Cooley, T.K. Smith, S.M. Rogers, and H.G. Miller. 1999. *Computer-Assisted Survey Information Collection.* New York: Wiley.

Turner, C.F., F.L. Ku, F.L. Sonenstein, and J.H. Pleck. 1996. "Impact of ACASI on Reporting of Male-Male Sexual Contacts: Preliminary Results from the 1995 National Survey of Adolescent Males." In *Health Survey Research Methods*, ed. R.B. Warnecke. Hyattsville, MD: National Center for Health Statistics.

Turner, C.F., J.T. Lessler, and J. Devore. 1992. "Effects of Mode of Administration and Wording on Reporting Drug Use." In *Survey Measurement of Drug Use: Methodological Issues*, ed. C.F. Turner, J.T. Lessler, and J.C. Gfroerer (DHHS Publication No. ADM-92-1919). Washington, DC: Government Printing Office.

Turner, C.F., H.G. Miller, T.K. Smith, P.C. Cooley, and S.M. Rogers. 1996. "Telephone Audio Computer-Assisted Self-Interviewing (T-ACASI) and Survey Measurements of Sensitive Behaviors." In *Survey and Statistical Computing 1996*, ed. R. Banks, J. Fairgrieve, L. Gerrard, et al. Chesham, Bucks, UK: Association for Survey Computing.

U.S. Department of Agriculture, Food and Nutrition Service. 1995 and 1996. "Cooperative Agreements for State Nutrition Networks." Request for Applications.

Wholey, J.S., H.P. Hatry, and K.E. Newcomer. 1994. *Handbook of Practical Program Evaluation.* San Francisco, CA: Jossey-Bass.

Willet, W.C., and L. Sampson. 1997. "Dietary Assessment Methods." *The American Journal of Clinical Nutrition* 65:1097S-1368S.

Willet, W. 1998. *Nutritional Epidemiology*. 2d ed. New York: Oxford University Press.

Woodsong, C., and L. Harris-Kojetin. 1997. "Expanding, Enhancing, and Improving Focus Group Research Methodologies." Invited workshop presentation at the annual meeting of the Association for Health Services Research, Chicago, IL.

Glossary

Glossary of Terms		
Abbreviation and/or Name		**Definition**
Cooperative Agreements		These were cooperative agreements awarded to States in 1995 and 1996 to support the development of Statewide nutrition education networks
Cooperators		This refers to the organizations supported by the award of a cooperative agreement to the State Food Stamp Program to support development of Statewide Nutrition Education Networks
Cooperative Extension		The county and State university extension organization affiliated with the Cooperative State Research Education and Extension Service (CSREES)
Cumulative Frequency		The number of times that an individual or target audience was exposed to an nutrition message over a specified period of time.
Cumulative Reach		The total percentage of a target audience exposed to nutrition messages over a specified period of time
Dietary Guidelines for Americans		These are the guidelines illustrated in the USDA Food Pyramid
Effectiveness evaluation		Evaluation of the impact of a social marketing program in widespread dissemination.
Efficacy Evaluation		Evaluation, typically in a pilot-test, designed to determine the impact of a social marketing effort when it is rigorously implemented.
Environmental Change		Changes in the physical environment that can influence behavior.
External Validity		The ability of study findings to generalize to the wider population.
FNP	Family Nutrition Program	The term used by some States to refer to the Cooperative Extension component of the State Food Stamp Nutrition Education Program.
FNS	Food and Nutrition Service	The organization that administers the Food Stamp Program
Formative Evaluation		Evaluation conducted while developing a social marketing effort to obtain feedback about proposed activities and materials.

Glossary of Terms		
Abbreviation and/or Name		**Definition**
Frequency		The number of times that an individual or target audience was exposed to a nutrition message.
FSNEP	Food Stamp Nutrition Education Plan	This term refers to nutrition education programs supported by the USDA School Lunch Program.
FSP	Food Stamp Program	The USDA organization that administers the Food Stamp Program.
GRPs	Gross Rating Points	An estimate of the percentage of the target audience exposed to a broadcast at any one time.
Impact Evaluation		Evaluation of the impacts of a social marketing campaign, typically changes in behavior.
Internal Validity		The ability of a study to rule out alternative explanations for study findings.
NEP	Nutrition Education Plan	The nutrition education plan submitted to and approved by the Food Stamp Program to support approved nutrition education activities under the State's Food Stamp Program Plan of Operations.
NET	Nutrition Education and Training Program	This program provides training for nutrition education related to USDA school lunch programs.
Policy Change		Changes in governmental or organizational policies or regulations that can facilitate changes in behavior (such as by increasing the availability to nutrition education or access to nutritious food.)
Primary Grantee		This refers to the lead organization awarded funding by the State Food Stamp Program under the cooperative agreements.
Process Evaluation		Evaluation to assess the delivery and implementation of a social marketing effort.
PSA	Public Service Announcement	Unpaid advertising donated by a television or radio station, newspaper or magazine, to promote a message of benefit to the general public.
Reach		The percentage of a target audience exposed to nutrition messages.
Reliability		The accuracy of measurement to yield the same result when it is repeated.
Social marketing		A comprehensive, audience-centered approach involving multiple, reinforcing channels of communication and environmental change to promote healthy behavior.
State's Food Stamp Program Plan of Operations		This is the plan submitted by the State Food Stamp Program that once approved by FNS becomes the basis for expenditure of administrative funds by the State Food Stamp Program.
System Change		Changes in organizational or community practices that can facilitate behavior change.

Glossary of Terms	
Abbreviation and/or Name	**Definition**
Validity	The accuracy with which a measure assesses what it is intended to indicate.
WIC	Supplemental Food Program for Women, Infants, and Children

Appendix A: Sample Size Requirements in Longitudinal Studies

When information is available that allows for an estimate of expected baseline levels and the expected effects of a social marketing campaign, it is possible to develop more refined estimates of sample size requirements. This section presents information regarding estimates of sample size requirements in longitudinal studies, in which the same individuals are interviewed before and after exposure to a social marketing effort. The sample sizes needed to detect various increases in improved nutrition behavior are given in ***Exhibit A-1*** and ***Exhibit A-2***. This exhibit assumes that there is a 5 percentage point increase in the control group (e.g., secular trends not associated with the intervention led to a five percentage point increase in the frequency of desired behavior between the baseline and the post-intervention period.) These estimates were developed for two scenarios:

- **Estimates of Changes in Individuals.** First, the tables indicate the sample size requirements to detect meaningful changes in individual dietary practices as measured in longitudinal surveys of individuals. The changes in nutrition-related behaviors in prior studies were mostly in the range of 10 to 20 percentage points. Hence, it can be useful to select a sample size of individuals large enough to detect a difference in rates of change of 15 percentage points. The most conservative estimates are based on the assumption that at baseline (as shown in Column A), 50 percent of respondents engage in a desired behavior. Assuming a survey with a design effect of 1.2, the sample size to detect a change of 15 percentage points (Column B) would be 95 (Column C). That is, it would take 95 individuals in each group to detect a significant difference, an increase from 0.5 to 0.65 in the intervention group versus an increase from 0.5 to 0.55 in the comparison group. It also would be prudent to set the initial sample size large enough to accommodate study dropout. Assuming that the completion rate is 80 percent, one would need to start with a sample of 119 individuals in both the intervention and the comparison groups.

Exhibit A-1. Sample Size for Detecting an Increase of δ in Behavior (Assuming 5% Baseline Increase in Control Group)

Initial Proportion (A)	Increase δ (B)	Longitudinal Sample Size (n) of Adults* (C)	Cross-Sectional Sample Size (n) of Adults (D)	Cross-Sectional Sample Size (n) of School Children (E)
0.8	0.05	1,180	2,360	2,178
0.5	0.05	796	1,592	3,576
	0.10	207	414	891
	0.11	172	344	735
	0.12	146	282	617
	0.13	125	250	525
	0.14	108	216	453
	0.15	95	190	394
	0.16	84	168	346
	0.17	75	150	306
	0.18	67	134	272
	0.19	61	122	244
	0.20	55	110	220

Notes: In Column C, this table shows the number of providers in the intervention group at each time point in longitudinal survey that would be required to detect a change as statistically significant (one-tailed), assuming a longitudinal survey of equal numbers of individuals in a comparison group. For instance, to measure a change in the intervention group between 0.5 at the baseline and 0.65 following the intervention (assuming an increase from 0.5 to 0.55 in the comparison group), it would require 95 individuals in the intervention group and 95 individuals in a comparison group. The sample size for a cross-sectional sample (n = 190) is twice this size (Column D).

In Column E, this table shows the number of students required to measure changes between the intervention group and a comparison group, assuming independent samples at both the baseline period and following the intervention. For instance, to measure a change in the intervention group between 0.5 at the baseline and 0.65 following the intervention (assuming an increase from 0.5 to 0.55 in the comparison group), it would require 394 students in the intervention group and 394 students in a comparison group at each time point. The column assumes a design effect of 1.4 for school children associated with their clustering within school classrooms.

* n per group per time period

Exhibit A-2. Sample Size for Detecting an Increase of δ in Prevention (Assuming 5% Baseline Increase in Control Group)

Initial Proportion (A)	Increase δ (B)	Longitudinal Sample Size (n) of Adults* (C)	Cross-Sectional Sample Size (n) of Adults (D)	Cross-Sectional Sample Size (n) of School Children (E)
0.2	0.05	471	942	2,492
	0.10	125	250	628
	0.11	105	210	520
	0.12	89	188	438
	0.13	76	152	373
	0.14	67	134	322
	0.15	59	118	281
	0.16	52	104	247
	0.17	46	92	219
	0.18	42	84	196
	0.19	38	72	176
	0.20	34	68	159

Notes: In Column C, this table shows the number of providers in the intervention group at each time point in longitudinal survey that would be required to detect a change as statistically significant (one-tailed), assuming a longitudinal survey of equal numbers of individuals in a comparison group. For instance, to measure a change in the intervention group between 0.2 at the baseline and 0.35 following the intervention (assuming an increase from 0.2 to 0.25 in the comparison group), it would require 59 individuals in the intervention group and 59 individuals in a comparison group. The sample size for a cross-sectional sample (n = 118) is twice this size (Column D).

In Column E, this table shows the number of students required to measure changes between the intervention group and a comparison group, assuming independent samples at both the baseline period and following the intervention. For instance, to measure a change in the intervention group between 0.2 at the baseline and 0.35 following the intervention (assuming an increase from 0.2 to 0.25 in the comparison group), it would require 281 students in the intervention group and 281 students in a comparison group at each time point. The column assumes a design effect of 1.4 for school children associated with their clustering within school classrooms.

* n per group per time period

A study that was based on independent cross-sectional samples (e.g., different individuals at each point in data collection) would require approximately twice the sample size to detect the same difference between treatment and comparison groups. In the example above, this would require a sample size of 190 individuals in the treatment group (Column D) and 190 individuals in the comparison group. Repeated cross-sectional samples are sometimes employed in telephone surveys of respondents before and after implementation of a social marketing campaign. In general, the cost efficiencies are such that longitudinal surveys are preferable **School Children.** The manual also developed estimates of the sample sizes needed to detect meaningful changes in practice when they can be implemented.

- **Estimates of Changes in** s using clustered independent samples at two points in time. This might be the situation in a school study in which it was not permissible, for confidentiality reasons, to link the responses of individual students between a baseline and a post-test survey. Assuming a design effect of 1.4, analyses suggest that it would take a minimum of 394 students in the intervention group and (with 394 students in the comparison group) to detect a change of 15 percentage points.[3] An analogous situation can arise if one were to develop independent sample of patients served by a health clinic.

It is useful to note that these are conservative estimates of sample size requirements. The sample sizes shown in the two exhibits are based on point estimates for individual items. Many of the items in a survey can be formed into item scales that would significantly decrease the variance and measurement error. This would reduce the attenuation that is associated with a one-individual item that has a greater amount of random variation associated with a response. Given the general reliability of available scales (coefficient alpha > 0.80; Cronbach 1951), it is likely that the use of scales rather than individual items would increase the effective power of comparisons by two to four times (Nunnally and Bernstein 1994).

[3] Estimates of design effects have been taken from prior studies. In general, the design effects associated with clustering of students within a school are expected to decrease as the number of schools increases. Design effects also can be reduced by stratification, pairing, or blocking, whereby similar schools are placed in the same stratum and then, within each stratum, schools are randomly assigned to treatment or control conditions.

Appendix B: Source Documents

This attachment contains two examples of evaluations of social marketing campaigns:

- A process evaluation of a Pilot Test of the "Eat Healthy, Youth Kids are Watching" campaign in Kent County, Michigan. This evaluation, conducted by Holaday Research under contract to Michigan State University, is notable for its ability to assess the cost effectiveness to different types of channels of communication in reaching low-income populations.
- An impact of the "Friends Can Be Good Medicine" Campaign. Although this campaign focused on social support, rather than on nutrition, it provides a useful example of an efficacy evaluation of an integrated community-based and media activity campaign in a pilot-test setting.

These evaluations were included in this document because they were conducted for State programs within the budget limitations that are typical of state nutrition education programs.