Charting the Course for Evaluation:

How Do We Measure the Success of Nutrition Education and Promotion in Food Assistance Programs?

Summary of Proceedings

February 28, 1997

CHARTING THE COURSE FOR EVALUATION:

HOW DO WE MEASURE THE SUCCESS OF NUTRITION EDUCATION AND PROMOTION IN FOOD ASSISTANCE PROGRAMS?

Summary of Proceedings

February 28, 1997

Editor:
Lynne Doner

Submitted to:

U.S. Department of Agriculture
Office of Analysis and Evaluation
Food and Consumer Service
3101 Park Center Dr.
Alexandria, VA 22302

Project Officer: Patricia McKinney

Submitted by:

KRA Corporation
1010 Wayne Avenue, Suite 850
Silver Spring, MD 20910

Project Director: Ed Rugenstein

This document was prepared under Contract Number 53-31984-035 between the Department of Agriculture and KRA Corporation of Silver Spring, Maryland. The contents of this document do not necessarily reflect the views or policies of the Department, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

CONTENTS

BACKGROUND

"Charting the Course for Evaluation: How Do We Measure the Success of Nutrition Education and Promotion in Food Assistance Programs?" brought together nutrition educators, traditional evaluators, market researchers, and experts at evaluation of health promotion efforts to establish a dialogue to identify and push forward the state of the art in evaluating nutrition education and promotion efforts. The conference took place on July 13 and 14, 1995 in Arlington, Virginia.

As the Food and Consumer Service (FCS) began to focus on integrating nutrition education into all its food assistance programs, it became clear that the FCS needed to be able to measure the effectiveness of such programs to ensure that limited resources were spent wisely. The agency also was looking for mechanisms to identify what program components worked best, under what circumstances, and at what cost. The goal was to assist everyone at the program delivery level to provide the best, most cost-effective nutrition programs possible.

To address these issues, the Food and Consumer Service assembled people with experience to provide their perspective evaluating a broad range of nutrition education, health promotion, and social marketing programs. This report shares what was learned at the conference by summarizing the major conference themes and presenting a synopsis of each session. The conference was divided into three sections:

1. **A retrospective look at what traditionally has been measured and how it has been measured in nutrition education programs**
2. **A look at state-of-the-art theories and methods for selecting evaluation techniques**
3. **Some lessons learned from ongoing and past programs**

Full transcripts of the conference can be obtained by contacting:

USDA Food and Consumer Service
Office of Analysis and Evaluation
Room 208
3101 Park Center Drive
Alexandria, VA 22302

OVERVIEW OF MAJOR THEMES

Designing and implementing nutrition education in nonclinical settings differs radically from conducting clinical research. Yet the speakers made convincing arguments that the experimental designs used for clinical science often guide expectations for what nutrition education programs can accomplish and how they should be evaluated. Their ideas about developing and evaluating behavior-focused nutrition education programs using models appropriate for population settings are summarized below.

1. Set appropriate objectives and manage expectations.

Nutrition education usually involves trying to change complex behaviors. Nutrition educators need to think through the type of intervention they will be doing before they set objectives. As one speaker noted, "the effect size for clinical interventions is large and hopefully fast. In a public health intervention, it's small, and at best, it's gradual."

Yet nutrition educators (and other health promotion practitioners) often set themselves up to fail by setting objectives for public health interventions that require large changes in behavior very quickly. In contrast, private sector marketers—who operate in the same environment as pubic health interventions—declare success with much smaller changes than health educators expect to make. A private-sector objective might be increasing sales by 2 or 3 percent, compared to a public-health objective of cutting the smoking rate in half. Furthermore, private sector marketers target consumers predisposed to their product, but nutrition educators are often working with those segments of the population that are least interested in making the change.

Public sector nutrition and health educators must guard against supervisors or funding agencies that expect to see change on a fiscal year basis; most of the successful health promotion efforts, such as decreasing stroke mortality and smoking, have measured change in decades, not years. Educators must make sure that expectations and objectives are appropriate for community-based programs directed at thousands of people, not clinical research looking at 60 subjects. Educators must also ensure that change is measured using evaluation models that are appropriate for their programs.

2. Define meaningful, measurable outcomes.

Often the outcomes selected for nutrition education programs are too global to meaningfully measure a program's effect, or to be measured with any accuracy. The science of measuring dietary change is relatively in its infancy. Most of the research on measuring diet has been epidemiological (particularly the relationship between diet and health outcomes), nutritional science (i.e., the relationship between diet and underlying biological mechanisms), and public health (broad-scale trends in large populations). Currently, a wide range of measures is used to evaluate dietary change, illustrating both the complexity of such change and the difficulty of measuring it.

Outcomes need to be realistic given the state of the target audience. For example, it would be unrealistic to expect a target audience with no prior knowledge of or interest in a particular behavior to embrace it immediately; a more realistic outcome would be to increase their knowledge of the behavior and its benefits to them.

Identifying and measuring intermediate variables in addition to outcomes is often critical to measuring progress. Behavior change can take a long time, frequently longer than the evaluation period, and intermediate variables help determine if progress is being made. More important, measuring intermediate variables helps identify those factors most important to behavior change, so efforts can be concentrated there.

Intermediate variables can take a number of forms depending upon the structure and objectives of the intervention. For example, if the intervention is a community initiative, intermediate variables can measure the new programs, policies and practices that are consistent with the initiative. Intermediate variables include anything that could influence the behavior change, such as interpersonal, environmental, accessibility, and availability factors.

3. Design interventions using appropriate theoretical models and design evaluations using the same models.

A recent FCS-sponsored review of the nutrition education literature found that nutrition education "works" when it is based on theory and has behavior change as a goal. Individual, social and environmental factors all play roles in behavior change, and interventions that influence all these factors are most likely to be successful.

The knowledge-attitude-behavior paradigm, used frequently with nutrition education programs, has serious deficiencies. It does not take into account any variable outside the individual, such as environmental factors. Research has not shown that it is necessary to change attitudes to change behavior (in fact, sometimes behavior changes first). Furthermore, the paradigm does not differentiate between types of knowledge, specifically "how-to" knowledge, versus motivational knowledge. Without motivation, people are unlikely to change behavior.

Theoretical models appropriate for nutrition education include: 1) stages of change, 2) social learning theory, 3) the health belief model, and 4) diffusion of innovations. Social marketing is not a theory; it is a process that can be used with any of these theoretical models to develop health promotion and disease prevention programs.

Often, some combination of these models should be used to develop successful programs. Also, at least one speaker thought program planners should pay more attention to developing policies and new services and less to messages about behavior. Finally, planners must ensure that the evaluation measures what the intervention was designed to do. Sometimes an intervention is designed to increase knowledge but measures behavior change instead. Then program managers wonder why they failed. Program planners and evaluators must work together to ensure the evaluation model fits the intervention model and measures the appropriate variables.

In addition, planners must determine the appropriate research approach. The gold standard for public health evaluations—randomized, controlled trials—is sometimes not appropriate for nutrition education programs delivered in community settings. Controlled trials assume that control or comparison communities receive no intervention. If the intervention is a community one, it may be impossible to hold out control or comparison communities.

Planners have a wide range of research approaches from which to choose. If a randomized, controlled trial is not the right approach for a particular intervention, alternatives include: time-series analyses, comparing indicator levels between groups with differential likelihood of exposure to a program, determining any other plausible explanations when change occurs rapidly, and assessing whether the observed outcome is in fact credibly explained by the process one thought would lead to change.

4. Include both formative and process evaluation activities.

All too often, outcome evaluation is the only type of evaluation used for nutrition education and communication efforts. Many speakers believed that formative and process evaluation are critical to the development and implementation of successful programs. If resources are limited, they recommend putting formative and process evaluation first. Skipping these activities may result in a flawed program. Outcome evaluators may conclude the intervention did not work, when in fact the materials needed refinement or delivery strategies required adjustment.

As one speaker put it, "one of the roles for research is to identify what matters for a particular target population in a particular place and then to guide the development of successful interventions." Formative research and evaluation assure the best possible program by identifying appropriate target audiences and ensuring program messages and activities are relevant and meaningful to them. Formative evaluation encompasses many aspects of program development: target audience selection, concept and message testing, pretesting, and market testing, to name a few.

Several speakers noted that identifying the determinants of behavior during the formative stage is important. One way to isolate determinants is to compare those who are engaging in the desired behavior with those who are not, and examine the variables upon which they differ. The importance of message testing before rolling out a program was repeatedly emphasized.

Process evaluation was deemed critical because it allows ongoing monitoring of programs and enables timely refinements—helping programs achieve success. Process evaluation activities include tracking participation in program events, tracking media coverage, and tracking progress made by change agents, such as coalition members. Carefully constructed process evaluation activities have many uses. They provide invaluable mechanisms for measuring intermediate outcomes. They allow mid-course adjustments to improve the program. They provide progress measures to funders and coalition members. And community leaders can use them to attract and maintain support and resources.

SUMMARY OF INDIVIDUAL SESSIONS

Contemporary Budget and Policy Realities:

The State of Nutrition Education in USDA and the Importance of Evaluation

Eileen Kennedy, Sc.D.

Executive Director

USDA Center for Nutrition Policy and Promotion

One out of six Americans is reached by a direct service delivery program operating out of USDA's Food and Consumer Service (FCS). Evaluation of FCS programs has been an ongoing effort. But there is often the perception that there is less we can say about the effectiveness of nutrition education programs than about the effectiveness of service delivery programs. Three questions are relevant to the discussion of evaluating nutrition education programs.

What works? Even when there is an agreed-upon outcome, there has been quite a difference of opinion on the paradigm that should be used for nutrition education and nutrition communications: nutrition education or social marketing. An example of the recent dialogue is shown in the following paragraphs:

Social marketing seeks to change consumer behavior by satisfying consumer desires and wants, but not by fostering consumer understanding of food and nutrition. The marketeers seek to promote specific behaviors without consumer understanding and in contrast, educators seek to build a framework into which learners can fit new information. . .Social marketing is cultural impoverishment because it promotes behaviors instead of teaching people to think.

Vanden Heede and Pelican, 1995

Nutrition educators do not understand social marketing. Social marketing provides a problem-solving process from which behavior change strategies are formulated and translated into discrete and integrated tactics aimed at specific behavior change. The emphasis is on consumer research to determine the most relevant and effective tactic to change behavior in a target audience, not what self-designated experts believe to be important for the target population to know or practice.

Lefebvre et al., 1995

This divergence of opinion also has implications for evaluation criteria and methodology in nutrition education. We have begun to look at trying to identify what the appropriate paradigm is, and, given that, what our ultimate measure of success should be. Should it be behavioral change, knowledge gained, or some combination? Over what time period? And clearly, given the policy reality, at what cost? Finally, what should the role of government be?

In what context does nutrition education work? Understanding the process leading to the outcome of a program is critical. For example, with the Supplemental Food Program for Women, Infants, and Children (WIC), we measured the effect on health outcomes. But we were also able to identify some key factors as to why WIC worked and was cost efficient in the United States when similar supplementary feeding programs in developing countries had high costs relative to outcomes. These key factors included higher levels of caloric supplementation and an integrated program, combining supplemental feeding with health care and nutrition education. Part of understanding the process of a program includes looking at when nutrition education works by itself and when it needs to be integrated with other interventions.

At what cost? More and more often, we are being asked which intervention achieves a given objective in the most cost-effective manner. Often, this question is "larger" than nutrition education. We are not asked to identify the nutrition education intervention that is most cost effective, but rather, given a nutrition education program, some jobs program, and a feeding program, for example, which is the most cost effective?

To answer these types of questions, we need to look more and more at the mix of programs needed to accomplish a particular objective. To be able to do this type of analysis for nutrition education, we need to do a lot of thinking about methodological needs and how we cost both the inputs into nutrition education and the outputs. For example, what is a gain in knowledge "worth" in a cost effectiveness analysis? This is an area that is ripe for very applied research. The science is there, but the application is lacking at the moment.

Implications for the Government

Contemporary budget and policy realities drive us, the Federal government, to reassess the role we have in nutrition education by answering two questions: 1) What is our comparative advantage? and 2) What is the most cost effective way of implementing the Federal government's role in nutrition education?

To summarize, in nutrition education we are headed toward a multifaceted approach to intervention strategies. Our evaluation strategies also have to be multifaceted and must include formative, process, and outcome evaluation research.

I. WHERE WE'VE BEEN

Overview: A Review of the Role of Evaluation in Recent Nutrition Education Research and Interventions

Isobel Contento, Ph.D.

Coordinator, Program in Nutrition and Education, Teacher's College,

Columbia University

Background

Recently, USDA contracted for a review of nutrition education research and intervention.[1] The focus of this presentation is on the evaluation measures used in the 217 studies reviewed. One criterion for inclusion in the review was that the studies had to be based on an experimental design involving random assignment, or on a quasi-experimental design where there was a comparison group. Another criterion was that there should be some validity and reliability of evaluation instruments, and these should be at acceptable levels. Only about a quarter of the studies screened met these criteria.

Summary of the Review Findings

The review examined two major questions:

1) *Does nutrition education work? If so, what are the success elements across interventions?*

2) *What are the implications for nutrition education program implementation, policy, research and demonstrations?*

The review found that nutrition education works. It is a significant factor in improving dietary practices, when behavior change is the goal, and the educational strategies are designed with that as a purpose.

The "behavior change as a goal" approach differs from interventions that disseminate information with the expectation that such information will result in changes in attitudes and behavior. In many cases the "knowledge-attitudes-behavior" model was misapplied. For this model to work, the "knowledge" must be motivational.

Knowledge-Attitudes-Behavior Model

The word "knowledge" has many meanings. There is "awareness knowledge" and "how-to knowledge." Social psychologists talk about "anticipated consequences knowledge" that is likely to enhance motivation to take action, while "instrumental knowledge" is the kind needed by people to act on their motivations.

Motivational knowledge is about the potential positive or negative consequences of behaviors. Examples of motivational knowledge include the following: 1) understanding about anticipated consequences, such as that eating lots of fatty foods may increase heart disease rates; 2) perceived susceptibility to disease, which is a kind of personal risk appraisal; 3) perceived severity, which is people's own fear about conditions such as breast cancer.

[1]This review is summarized in Contento, I. and others, "The Effectiveness of Nutrition Education and Implications for Nutrition Education Policy, Programs, and Research: A Review of Research," ***Journal of Nutrition Education***, 1995; 27(6), 277-418.

4) perceived benefits, such as that exercise will make you feel fit; and 5) a sense of mastery, or self efficacy, can also be motivating.

"How-to" information tends not to be motivational, although very important. It is the kind of information that seems to be used most often in the reviewed studies. The following examples are helpful if you are already motivated.

- A teaspoon weighs 5 grams and is 45 calories.
- Here is the Food Guide Pyramid. You can see that you should eat 6 to 11 servings of grains, cereals, pasta, and rice a day; 5 servings of fruits and vegetables; and 2 servings each of meat and dairy. Eat fats and sugars sparingly.
- This label means that each serving of this cereal has 8 grams of fiber. This is 32 percent of the Daily Value for fiber.
- Here is how you modify a popular recipe to make it lower in fat.

A school health education evaluation of 30,000 students and about 1,000 classrooms also sheds some light on the knowledge, attitudes, and behavior relationship. It found that 5 to 15 hours of instruction produced an enormous effect on program-specific nutrition knowledge. With a few more hours, an effect on more general knowledge could be produced. It took 25 to 50 hours to bring about a change in practices. Interestingly, attitude changes took 25 to 50 hours to produce, and even then the effects were small. The idea that knowledge leads to attitude change to behavior change may not be correct.

In general, the review found the more successful programs were those that: 1) set behavior change as a goal, 2) incorporated communications that were motivating, 3) taught strategies for behavior change, 4) included active involvement of both the individual and the community, and 5) attempted to build health-enhancing environments.

Evaluation Instruments Used in the Studies

Preschool children: Of 21 studies, 7 measured only knowledge; 6 measured knowledge and behavior, and a few measured attitudes. "Behavior" was measured differently in different studies, and included choosing a picture of a nutritious snack versus a low-nutrient snack; actually observing and recording behavior; and food preferences (because for preschool children, preferences are highly correlated with consumption). Some studies examined the role of peers modeling the behavior; some examined the use of rewards, positive attention by adults, and adult modeling in changing preferences.

School-aged children: Out of 17 general nutrition education studies, 15 measured knowledge, 12 measured attitudes, and most measured a variety of behaviors. Behavioral measures included observation of school lunch choices and school plate waste; observed snack preferences; 24-hour recalls; 3-day food records; food choice inventory; food frequencies; a lifestyle questionnaire; and frequencies of food-related behaviors. Skills were measured in several studies, and several looked at a variety of social-psychological variables, including self-esteem, self-efficacy, and health locus of control (the degree to which a person feels in control of his or her health or feels it is controlled by external factors). Self-efficacy was measured by asking children if they were "not sure," "sure," or "very sure" that they could choose the healthful alternative most of the time given two foods. Some

studies used physiologic measures, including serum cholesterol, urinary sodium, skin folds, blood pressure, body mass index, and weight.

Adults: Many of the 62 studies measured nutrition knowledge, but the proportion using knowledge as an outcome measure was much lower than for school-age children, probably reflecting the difference in goals for educational interventions with adults versus children. Knowledge measured was often program specific, rather than general nutrition knowledge. Attitudes were measured in relatively few studies. Behavioral measures included 24-hour dietary recalls, variously analyzed for food groups, fat intake, or nutrients; food frequencies; "dietary intakes," self-report of trying recommended behaviors; checking and knowing own cholesterol level; purchase intention; simulated purchase; sales of specific items in cafeteria/vending machines/restaurants; and sales of specific items in grocery stores. Physiologic measures were used in many studies and included serum cholesterol, weight, blood pressure, and overall risk.

Pregnant women/caregivers of infants: Fourteen studies were with pregnant women and 15 related to promoting breastfeeding. Knowledge and attitudes were measured in only a few studies; behaviors were measured more frequently and included mean number of prenatal visits, intake from each of the food groups, coffee and alcohol consumption, vitamin-taking behavior, and well-baby visits. Most of the interventions evaluated impacts on infants as the outcome measure. Breastfeeding behaviors included breastfeeding duration or percent of mothers breastfeeding at a specified time postpartum; qualitative behaviors (e.g., position used, latch-on); and delay of introduction of solid food. Physiologic measures were common and included mothers' weight gain during pregnancy, pregnancy-related complications, delivery complications, infants' birth weight, percent of low-birth-weight infants, number of weeks of gestation at delivery, APGAR scores, and postnatal complications score.

Older adults: Of 14 studies, 4 measured knowledge (3 measured general knowledge) and 2 measured attitudes. Behavioral measures included 24-hour recalls, food frequencies, observation of consumption in a residential dining room, shelf inventory of food items in the home, self-reported change of any kind, use of health services, health status, health behavior, and exercise. Health risk appraisals were used in three studies, and some studies used physiologic measures, including body weight, blood pressure, serum lipids, urinary sodium, and health insurance claims for medical visits.

Inservice training of nutrition intermediaries, such as teachers, paraprofessionals, professionals: Knowledge was commonly measured in the 21 studies; attitudes were measured in several. Behavioral measures used to evaluate the outcome of training included fidelity to the curriculum, whether teachers increased the amount of teaching they did because of their preparation in nutrition education, and quantity of teaching. In three of the six studies with food service workers, the evaluation measure was whether they actually reduced the fat and sodium in the school lunches as a result of training.

Summary

The most important observation that arises from an overview of these 217 studies is the wide variety of outcome measures used to evaluate nutrition education effectiveness. Most studies, regardless of population category, measured knowledge, attitudes, or other related social-psychological variables and behaviors. Many interventions also used physiological parameters as outcome measures. Within these categories, a wide variety of measures were used.

In particular, the types of measures used to evaluate "behavioral change" varied widely from study to study.

Often, the behaviors being measured were intakes of foods or nutrients as measured by food recalls or records, food frequency questionnaires, or observation. Other studies used behavioral intentions or self-reported likely choices among foods as proxy measures. Actual purchases were used to evaluate point-of-choice interventions, as were many other measures. This wide range of evaluation measures illustrates the complexity of dietary change and the difficulties of measuring such change.

Confounding Issues in Evaluations of Nutrition Interventions

William Smith, Ed.D.

Executive Vice President

Academy for Educational Development

Human Behavior in Relationship to Health: What We've Learned

Over the past 20 years, there has been some consensus in America that there is a relationship between human behavior and health, and some relationship between human behavior and disease. The second part of that consensus is that government has some responsibility to combat disease and protect health. Therefore, government has the obligation to help people adopt healthier behavior. Looking at causes of mortality in the United States in 1990, about 19 percent of it was due to tobacco. Diet and activity were a close second, and alcohol was number three. Further down the list were AIDS and firearms. Clearly, human behavior is important in reducing mortality.

We have made huge progress in this century in protecting human health. For example, there were 21,000 cases of polio in 1952 in the United States. In 1994, there were four cases. In 1941, there were almost 900,000 cases of measles; by 1993, there were 281. But there were 1,500 cases in 1983 and 25,000 in 1990. Why? Because we allowed our immunization program to go "through the floor," particularly in high-risk populations. We didn't eradicate measles, we put a strong program together which brought it down. When we discarded that program, measles went back up.

There has been tremendous success in the heart disease area. Stroke mortality is probably the biggest single success in the past two decades, having declined by 50 percent. Key behavioral measures increased significantly during that time period: people assessing their own risk increased from 50 percent to 75 percent, physician visits increased by 70 percent, and blood pressure readings increased by 95 percent.

HIV/AIDS and fatal drunk driving are two more examples. In San Francisco, using a very sophisticated model, they have concluded there were 8,000 new cases of HIV in 1987. In 1995, that number is anticipated to be around 650. Why? It's a combination of very complex factors, some of them regulatory, some educational and perceptional, some technological. With drunk driving, the number of alcohol-related fatalities in the United States is down by 31 percent. Why? Again, a combination of factors—air bags, driving age, better enforcement, community-based advocacy and very, very powerful messages.

Models to impact behavior and health can be put into three big categories: regulatory choices, new services (i.e., immunizations or counseling and testing services), and messages. We spend far too much time talking about messages, too little time talking about services, and a little bit less time talking about enforcement opportunities that help influence

behavior. All of these things can be used at different levels of society, government, the private sector, and the community.

The Vaccine Model: Inappropriate for Measuring Behavior Change

Four factors confound our understanding of different human behaviors and how they relate to health, evaluation and research: 1) the complexity of human behavior; 2) the instability of human behavior over time; 3) the difficulty of replicating interventions; and 4) the fact that we are thinking about our programs as prevention "vaccines."

Much of our research and evaluation has been driven by the vaccine model. Research that establishes a vaccine has to look at three things: 1) Does it work, is it efficacious? 2) Is it stable, does it change? 3) Are there side effects? Those are the three principal characteristics of a good vaccine. It's stable. It works. And it doesn't have any side effects.

However, trying to develop a prevention "vaccine" for behavior is much different from developing a vaccine against disease. The vaccine model is very linear. It says we develop vaccines in a case-control setting, test them to establish their validity, and then we establish their replicability in populations. For example, the prevention program is developed in Minneapolis. Then it is "injected" into Chicago. But there is greater variability in the settings for prevention programs than in human bodies. There is variability in the host. Different "hosts"—prevention program settings—can differ dramatically from each other on relevant dimensions, such as environmental influences and characteristics of the people being targeted by the intervention, whereas human bodies are comparatively interchangeable. There is variability over time. There is also variability in the prevention vaccine itself.

One problem with the case-control model is that its constraints do not allow us to modify interventions once they are in place. For example, let's look at a very powerful, interesting study, the 22-community, 4-year Community Intervention Trial for Smoking Cessation (COMMIT). It targeted heavy smokers (more than 25 cigarettes a day), but also looked at moderate smokers. It was a comprehensive, community-wide program with two goals: to ensure exposure to smoking messages and to alert people, particularly heavy smokers, to cessation opportunities in their communities. The results showed no effect in heavy smokers in the intervention communities, and only some effect in moderate smokers. The conclusion drawn was, "now we have evidence that community intervention just doesn't work."

However, a strong secular trend was affecting both intervention and control communities. Change occurred in both communities; everybody was getting better at decreasing smoking rates. The study showed that the intervention did not produce an effect any stronger than a very strong secular trend. Much of what was going on in the intervention group was going on in the control group as well. Because it was a case control program, there were a lot of things that occurred during the 4 years that the interventionists could have changed because they found out they weren't working as well. But they didn't change them because they were testing a vaccine and the vaccine can't be changed in the middle of the test.

These are some problems with the vaccine model: 1) the nature of the intervention itself might change; 2) the amount of time necessary for change may be different from the study time period; 3) expectations may be too high; 4) the importance of secular trends is not factored in; and 5) interventions cannot be adjusted to local needs in research projects.

Research and Evaluation Paradigms

Three distinct research and evaluation questions that we are all confronted with when working on prevention

programs are the following:

1. Does X work? (X usually compares one type of intervention to another)
2. Given that we know X works over there, did it work here?
3. How do I get X to work this time?

The third question is a marketing research question, or it might be described as clinical research for those offended by the marketing vocabulary.

Clinical and marketing models provide better paradigms for integrating research and evaluation than the vaccine model. In the clinical model, a physician looks at two different patients and determines how those patients differ and what treatment each needs. The physician then monitors the effects of the treatment and adjusts it as needed. In the marketing model, the marketer looks at two different consumers, determines how they are different, and develops a program to position the product so that it provides a benefit to each consumer. Then the marketer monitors the program to see how it is working, refining it as necessary.

The vaccine model is linear. It says we hypothesize something. We test it. We evaluate it and then we put it into practice. The clinical/marketing model is circular. It says research and action are interrelated. We need to assess things first, make some plans based on that assessment, test out that planning in real life, go to scale, and then look at and monitor the thing because we're going to be making mistakes all over the place. Then we'll make some adjustments. And when we make adjustments, we're right back at the beginning.

In using these models, start not with the people, but with the behaviors and an understanding of how complex they are. Think about the targets for those behaviors. Then ask yourself three questions: What are my policy options? How can I build better services? And how can I come up with a message that's a little bit clever and persuasive, too? One role for research is to identify what matters for a particular target population in a particular place and then to guide the development of successful interventions.

Key to understanding behaviors is identifying the determinants of the behavior. To identify determinants, look at those who engage in a behavior and those who do not (doers and nondoers), and then compare their answers. For example, in a study of condom use among women in the Caribbean, there was no difference between users and nonusers in AIDS knowledge, perceived severity of the disease, or perceived susceptibility to it. However, there were differences in terms of whether they talked with friends about condoms and whether a sex partner once suggested using a condom. If a program could get sex partners to suggest using a condom, it could probably do something to influence condom use. Having this knowledge totally reframes the issues that the program should address.

Summary

We are facing choices and evaluation and research can help us make those choices. The confounding issues are the complexities of the human behavior we're facing, the instability of those behaviors over time, the difficulty in replicating interventions, and this prevention vaccine myth. Thinking about improving the balance between basic research, evaluation research, and something more like marketing or clinical research may be useful.

II. CHARTING A NEW COURSE

Using Communication and Behavioral Models in Designing Evaluations

Research has shown the importance of a theoretical base for successful nutrition education interventions. Each of the first three sessions in this section covers a theoretical model often used in designing nutrition education interventions today. Besides describing the model, each speaker also presents examples of its application from his or her own work.

The ***health belief model*** has three components: 1) readiness, including individuals' perceptions about their susceptibility to a disease and the disease's severity, their motivation to make any changes, and their subjective estimate of whether or not the regimen is safe and whether it would actually have an effect; 2) factors that modify the disease's perceived threat, such as demographic characteristics, structural variables (e.g., prior contact and knowledge about the disease), and quality of care; and 3) likelihood of taking action to avoid the disease, determined by assessing whether the perceived benefits of preventive action outweigh the perceived barriers.

Social learning theory, or social cognitive theory, posits three overall determinants of behavior: 1) personal factors, such as attitudes and values; 2) environmental influences on behavior (external to the individual); and 3) the behavior itself. Central to the theory are the concepts of reciprocal determinism (the three determinants can influence each other), perceived self-efficacy (a person's confidence in his or her ability to engage in a behavior; the higher the confidence, the more likely the person will be able to do it); reinforcement for the behavior; modeling of the behavior to make it easier to adopt; self-regulation, which concerns how people react when they attain goals; and observational learning.

The ***transtheoretical model of stages of change*** outlines six stages for behavior change based on the individual's current behavior: 1) precontemplation, when the individual is either unaware of or not interested in making a change; 2) contemplation, when the person is thinking about changing, usually within the next six months; 3) preparation, when the individual actively decides to change and plans a change; 4) action, when the person attempts the change; 5) maintenance, when the individual sustains the change for six months or longer; and 6) termination, when the behavior has become so ingrained that the person is no longer in the stage cycle.

The final session in this section describes ***social marketing,*** an often misused and misunderstood term. Social marketing is a disciplined, research-based process of adapting commercial marketing techniques to influence the voluntary behavior of target audiences so that they increase their own or society's well being. Social marketing's key features include a focus on behavior change, an emphasis on formative and process evaluation activities, recognition that the behavior being promoted has competition, segmentation of audiences, and development of products for each segment.

Health Belief Model

Arlene Caggiula, Ph.D.
Associate Professor, Nutrition and Epidemiology
Graduate School of Public Health
University of Pittsburgh

Background on the Model

The health belief model, derived from social-psychological theory, was first delineated by Rosenstock in 1966. Originally developed to explain preventive health behavior or compliance with preventive health behaviors, it was based upon Llewin's goal-setting theory. Llewin essentially described the individual as existing in a life space composed of regions, some of which are positively valued, others negatively valued, and some neutral. Disease would be a region of negative value.

Rosenstock took Llewin's hypotheses further and created a model with three components: individual perceptions, modifying factors, and likelihood of taking action to avoid disease. The individual perceptions refer to 1) perceived susceptibility to the disease—people would have to perceive themselves as personally vulnerable or personally susceptible, and 2) perceived severity of the disease—they would have to believe that the occurrence of the disease would have a moderately severe effect on some component of their lives.

The perceived threat of the disease is modified by demographic and structural variables, cues and action, mass media campaigns, and advice from others. Demographic variables include age, sex, race, ethnicity, socio-psychological variables, personality, social class, appearance, and reference group pressure. For example, if a person has friends that consider the disease a serious problem then he or she may think of it as a more serious problem. Structural variables include prior contact and knowledge about the disease; if a person does not know very much about it, it is difficult to feel much of a threat.

As a result of an individual's original perception as modified by these variables, the likelihood of the individual taking action will be affected by the cost-benefit ratio: whether or not the perceived benefit of the preventive action is greater than the perceived barrier. In other words, taking action would be beneficial and not entail important barriers such as cost, inconvenience, pain and embarrassment.

Rosenstock's model was modified a few years later by Becker and Naiman at Johns Hopkins University, although it remained a three-component model. Individual factors were subsumed under "readiness," and two other elements were added to this component: motivation to make any changes, and the subjective estimate of whether or not the regimen was safe and whether it would actually do something. With nutritional therapy, very often people simply don't want to believe that making dietary changes has any effect. Modifying factors are similar to the old model, with a quality of care characteristic added. This characteristic involves whether or not the intervention program helps patients to reduce the complexity of their regimen, and the relationship between the caregiver and the patient.

Applying the Model

Let's review two studies examining the relationship between this model and adherence to hypocholesterol

eating patterns—those that are low in cholesterol and low in total fat. Both of the studies were done following a group intervention over a fairly long period of time. One study was with hypercholesteremics or people who had a history of early coronary disease. The other study used a group of men involved in the Multiple Risk Factor Intervention Trial (MRFIT), all of whom were at high risk of developing cardiovascular disease based upon their cholesterol, blood pressure, and smoking habits.

Based on the health belief model, for the first study we constructed a series of 20 statements, with 4 different compliance factors represented by 5 statements each. The statements were designed to reflect adherence to dietary regimens. Study participants responded to each statement using a seven-point scale of agreement. We also collected information as to their subjective assessment of their adherence by asking people whether they never followed their eating pattern, seldom followed it, followed it half the time, usually, or always. And we used a food record rating as an objective assessment, which measured saturated fat, polyunsaturated fat, and dietary cholesterol content of the diet.

Our analysis examined how much of a role each factor played in participants' adherence to the regimen. Perception of threat was the greatest predictor of whether or not they complied using their own subjective assessment. Quality of care—that is, the health professional's sensitivity and ability to simplify the regimen to make it possible for more people to adhere—was also important. The cost-benefit and social support factors were least important to adherence in this study.

With the study of MRFIT participants, we analyzed only the objective assessment because we didn't have enough data to analyze the self assessment. Also, we added some additional elements to the model to see if we could improve its predictive ability. We added an element called external environmental media (how much information people get from magazines and groups such as the American Heart Association) and internal and external locus of control (the degree to which a person feels in control of his or her health or feels it is controlled by external factors).

In this population, cost benefit turned up to be the most predictive of their compliance. Cost benefit in this situation meant they felt that following the eating pattern was beneficial and the benefit was greater over time. Quality of care was also important in this study. Threat of disease was not very high and we think it's probably because these individuals had a long period of intervention and no longer felt at great risk of the disease. As for external environmental media, information from groups such as the American Heart Association seemed to be most important.

Social support turned out to be the least important factor in both studies; it may be because both studies had very intensive intervention programs and participants may have felt they had a great deal of support within the program.

Summary

In conclusion, we believe that the health belief model does predict adherence to eating patterns low in cholesterol and total fat. However, the relative importance of the components varies depending upon the type of population, the type of intervention program, and the measure of adherence being used.

Social Learning Theory

Kim Reynolds, Ph.D.

Associate Professor

Department of Health Behavior

The University of Alabama at Birmingham

Overview of Social Learning Theory

Social learning theory had its roots back in the 1940s; however, today's focus will be on the more recent formulations, particularly the work of Dr. Albert Bandura of Stanford University. The more recent term for this theory is social cognitive theory. The theory is very broad and incorporates many different components to explain and predict behavior. It has been used to explain behavior ranging from dietary behavior and physical activity to clinical behaviors, such as phobias. It is widely used for intervention design and very much respected for its ability to guide researchers and practitioners toward concepts that they should bring into their intervention strategies.

The theory assumes that people are active in determining their own behavior, that is, they do not simply act according to how they are rewarded. They continually think about their behavior, different actions they can take, and what the incentive would be for that behavior. Then they select a course of action.

The theory states that there are three overall determinants of behavior: 1) personal factors, such as attitudes and values, that the individual holds; 2) environmental influences on behavior (external to the individual, including the influence of family and friends, physical factors, and availability); and 3) the behavior itself, including behavioral capability and outcome expectancies. Behavior capability involves how-to knowledge and skills—if someone does not have very basic skills-related knowledge and does not have any of the skills required to engage in the behavior, they won't be able to do it. Outcome expectancies are the anticipated results of the behavior; they can be positive or negative, short or long term. In part, assessment of outcome expectancies provides an incentive for behavior.

Principles Underlying Social Learning Theory

Reciprocal determinism, says that attitude and behavior are mutual determinants; personal factors can influence behavior and behavior can influence personal factors. For example, if I hold a positive attitude toward exercise, that may lead me to exercise. As I begin to exercise more and more, that behavior causes me to rethink and strengthen my attitudes toward exercise.

Perceived self-efficacy is a statement about the confidence a person has that he or she can engage in a behavior. The higher the confidence, the more likely the person is to be able to do it.

Reinforcement. Different kinds of reinforcement include extrinsic reinforcement, where someone could be paid for making behavioral change, for example, and intrinsic reinforcement, such as feeling good about goals attained or changes made.

Outcome expectancies. Observing someone else making the behavior change ("modeling" it) and being rewarded for doing so reinforces the value of the action, thereby conveying outcome expectancies to the observer.

Self-regulation. People set goals, watch their performance, and regulate their behavior. Some will reward themselves if they attain their goals; others will adjust their goals and do better next time.

Observational learning, or modeling. People can acquire behaviors and skills, even values, by observing the behavior of others. They can even learn things like perceived efficacy. When people see a similar model successfully doing a behavior, they come to believe that they are more capable because someone like them is engaging in that behavior.

An Example of Social Learning Theory (SLT) in Design and Intervention

The High Five program in Alabama is part of the 5 A Day for Better Health initiative. It consists of a school-based intervention to increase the consumption of fruits and vegetables in 28 schools. The program is being evaluated with a very strong randomized experimental design. The intervention was designed using social learning theory so it focused our attention on multiple determinants of behavior. We sat down with the model and developed strategies that would be linked to the model in the belief that this would produce the strongest intervention. Some examples:

Environmentally focused strategies: We focus on parents and do a number of things to try to modify their behavior, which will, in turn, influence their children's behavior. We also work directly with food services.

Behavioral capability. Children's ability to ask for fruits and vegetables from their parents was an important skill, as was preparation, because kids in the age groups we are working with prepare many of their own meals. Also, skill-related knowledge was needed; for example, some kids were not familiar with a large variety of fruits and vegetables.

Outcome expectancies. The intervention includes stories that relay the positive effects of fruit and vegetable consumption and self-efficacy. It also includes role playing that gets children to practice the target behavior, such as asking skills. This provides children a successful experience in a controlled environment, so they come to believe they are more capable or confident in their ability to use some of these skills.

Reinforcement, self-regulation, and observational learning. Sometimes very small things, like stickers, have great reinforcement value. For self-regulation, we tried to help children set goals. We gave the kids and the parents simple tools for self-monitoring. Observational learning included role playing, skits (including some targeting parents), and things like teachers leaving fruits and vegetables visible for the children to eat.

Linkage Between Social Learning Theory and Evaluation Design

Use of this model can help us with evaluation in several ways. It will:

1. Lead to stronger interventions, which lead to larger effect sizes that are more readily detected in evaluations.

2. Guide measurements, because it involves looking at the environment and examining behaviors and personal factors carefully.

❸ Allow consideration of mediators. What is it between the program implementation and the outcome (dietary consumption) that produced the change? For example, with 5 A Day, mediators might include improved self-efficacy, asking skills and availability—a personal factor, a behavioral skill, and an environmental factor. We would postulate that these mediators account for the change and we would be very careful to measure each of them.

Evaluation Measures

For environment, we look at social influences and measure availability through parents using a measure developed by Baranowski. Behavioral capability is measured by looking at asking skills and key knowledge. Positive and negative outcome expectancies are measured on a 20-item scale. Self-efficacy is measured across meals, by looking at the aggregate self-efficacy for eating five fruits and vegetables per day. Process evaluation measures include how well the reinforcement activities were executed, how self-regulation has been executed, and how some of the observational learning activities were executed.

Summary

Social learning theory is a general theory. It explains many behaviors, but it can be readily adapted to dietary behavior. It is very useful in intervention design and in measurement, through defining mediators and guiding us toward different kinds of measures that we can utilize.

Stages of Change: The Transtheoretical Model

Marci Kramish Campbell, Ph.D.
Assistant Professor
Department of Nutrition
University of North Carolina at Chapel Hill

Background on the Model

Stage models are not necessarily theories in and of themselves; rather, they are a good way to look at how behavior change might occur. While there have been many stage models over the years, this presentation focuses on the Prochaska and DiClemente transtheoretical model. Originally developed to understand smoking cessation, this model has been applied to a variety of other behaviors, including addictions, weight loss, and dietary change.

The stages-of-change model has three basic assumptions: 1) behavior change involves a series of cognitively different steps; 2) common stages and processes of change exist and can be applied across a wide range of health behaviors; and 3) interventions may be more effective if we tailor them to the stage of change that people are in at the time.

The transtheoretical model identifies the following stages, based upon the individual's behavior and, for some stages, some length of time related to the behavior:

1. ***Precontemplation,*** when the individual is either unaware of or not interested in making a change
2. ***Contemplation,*** when the person is thinking about changing, usually within the next 6 months
3. ***Preparation,*** when the individual actively decides to change and plans a change, usually within 1 month; sometimes the person may have tried to change
4. ***Action,*** when the person is attempting to make changes, but has been doing so for less than 6 months
5. ***Maintenance,*** when the individual sustains the change for 6 months or longer
6. ***Termination,*** when the behavior has become so ingrained that the person is no longer in the stage cycle

People do not necessarily move through the stages in a linear progression; they often try to change, relapse, and then try again before succeeding. Relapses can occur at any point in the process before termination is reached.

Moving through each stage involves a different process. For example, the process for contemplators may be seeking information about the behavior change, whereas people in the maintenance stage know how to make the change and are making it; they may have moved on to trying to make policies (e.g., smoke-free flights) and would have entirely different needs. Interventions targeting a particular stage should be tailored to the process appropriate for it; one way to identify the process for each stage is by studying successful self-changers.

Evaluation Using the Stages of Change Model

The stages-of-change model has been used in evaluating several health behaviors, including smoking cessation, mammography screening, and some nutrition behaviors. For smoking cessation, Prochaska's group in Rhode Island looked at 870 people, randomized to four different types of interventions: 1) a standard intervention with American Lung Association manuals, 2) an intervention that used manuals matched to the stage participants were in, 3) a personalized counselor intervention, and 4) interactive computer reports based upon the individual's stage of change.

They found that the standard intervention was the least effective, and the stage-matched individualized reports were the most effective. Similarly, a mammography screening study by Sillette/Skinner & Associates found that giving stage-matched messages, particularly to lower income and African-American women, resulted in the audience being more likely to move forward to action.

On the nutrition front, several really good cross-sectional studies have shown that stage of dietary change correlates with dietary intake, particularly for dietary fat, fiber, and fruits and vegetables. One study we did in North Carolina involved a stage-matched intervention. At the baseline, stage of change was correlated with dietary intake; for example, people in the action stage ate the least fat, the most vegetables and the most fruit. A randomized trial was conducted in which a third of the people received stage-matched messages, a third received a standard intervention, and a third served as a control group.

At followup, after baseline consumption was controlled, people who got the tailored messages decreased their fat intake much more than people who got the nontailored message or the control group. However, social learning theory, the health belief model, and persuasion theory were used as well as stages of change, so we can't really say that all of the effects were due to stages of change.

Practical Uses: Measuring Stages of Change

For program planning, individuals can be grouped according to the stage they are in relative to the behavior to be changed. Then, education can be designed for each stage, based upon the process identified as most likely to ease moving through the stage.

How do we know what stage people are in? First we look at current self-reported behavior related to a targeted objective, to separate doers from nondoers. Then we try to separate them further based on how long they have been doers. We then look at readiness to change. We might measure relapse history if we think past attempts will predict something about future attempts. We may also look at decisional balance, or pros and cons. People in more advanced stages tend to endorse more of the positives of the behavior and people in earlier stages tend to endorse more of the negatives.

An example from 5 A Day:

- How many servings of fruits and vegetables do you eat?
- (For those eating 5 or more) How long have you been eating 5 or more servings?
- (For those eating fewer than 5) Are you seriously thinking about eating more fruits and vegetables in the next 6 months?
- Are you planning to eat more fruits and vegetables in the next month?

Some challenges of measuring stages of change deal with asking these kinds of questions in different populations. Do the questions work with children? With persons of low literacy? With different ethnic groups?

For programs that use stages of change, some questions to ask to evaluate use of the model:

- Did the program appropriately assess stages of change of the participants?
- Was cognitive testing done?
- Did the designers think about what the questions meant?
- Were the intervention stages appropriate?
- Was the research designed so that comparisons could be made between an intervention group and a control or comparison group, using either experimental or quasi-experimental designs?

- Was change in stages looked at, and intermediate outcomes plus behavior change?

Persuasion and Social Marketing

Alan Andreasen, Ph.D.
Professor, Department of Marketing
Associate Dean, Faculty Affairs
School of Business
Georgetown University

What Is Social Marketing?

Social marketing is an adaptation of commercial marketing technologies to programs designed to influence the voluntary behavior of target audiences to increase individual well-being and/or that of the society as a whole. Social marketers are very action-oriented. They are interested in theory, but in terms of what aspects of theory can be used to help design programs.

Key Features of Social Marketing

Focus on behavior changes. How is educating or providing information going to lead to behavior change? Sometimes education has a boomerang effect. For example, many health models encourage program planners to tell people about risk factors so that they know they are at risk; when that was done with breast cancer to encourage mammography, mammography rates actually dropped. This drop occurred when some women concluded they didn't need a mammogram because they didn't have the risk factors.

Insist on marketing research in the formative, pretesting, and monitoring stages. Many programs conduct formative research to develop strategies, then develop materials and put them in the field without testing them. Private sector marketers go out and see whether they work. Monitoring research reveals what's working during the project, so it can be changed if necessary.

Recognize that the behavior change being promoted has competition. When people are being urged to do something, in their mind they have alternatives—and those alternatives have important payoffs and benefits. Sometimes the competition may not be directly related to individual behavior. For example, when trying to get mothers in developing countries to immunize their kids, we found that the competition was taking care of their husbands. Immunizing the kids involved being gone all day, getting home late, and getting hassled by the husband because dinner was late.

Don't develop one product for all markets. In the private sector, General Motors doesn't develop one Chevrolet to fit everybody. They develop lots of different Chevrolets with different options. Yet, often there is one intervention program, even though it makes sense that different members of the target audience are going to respond to different kinds of interventions. Target audience members can be segmented many different ways; stages of change is one; lifestyle research—looking at differences in people's life patterns and the kind of people they are—is another.

Social marketing is more than communications. Social marketing is not social advertising. Behavior is the bottom line and many things have to be put into place before people will undertake the desired behaviors; advertising is just one part of it.

Behavioral Models

Behavior is influenced by four factors:

1. ***Perceptions of expected benefits.*** People have to get something in exchange for engaging in the behavior.
2. ***Perceptions of expected costs.*** People have to pay some costs to undertake the behavior.
3. ***Community-level effect.*** People do things if other people are doing them, even if their personal consequence calculation isn't all that favorable.
4. ***Ability to affect outcomes,*** or self-efficacy, is the confidence people have that they can make the behavior change.

Behavior comes about in stages; different interventions are appropriate for each stage. The stages described below are similar to those discussed in the stages-of-change presentation.

During *precontemplation,* the major issues are education and changing values. Many people are not undertaking the behavior because they don't know about it or because they think it is not appropriate for them.

Contemplation can be divided into early and late stages. Intervention strategies need to differ for the two stages. Benefits are more important early in the contemplation stage—as people think about it, they've got to see a lot of benefits or they're not going to move on and think about it anymore. Later in the contemplation stage, when they know and endorse the benefit, costs become much more important—reducing cost and bringing social influence to bear is very important.

For the *preparation* and *action* stages, self-efficacy is critical. Even if target audience members think the cost/benefit ratio is favorable and there is a lot of group pressure, if they don't think they can do it themselves, they won't do it.

Summary

To summarize the social marketing perspective on intervention programs:

- Monitoring and pretesting (process and formative evaluation) are important parts of an evaluation strategy for any intervention program.
- Emphasize audience segmentation and potentially more sophisticated segmentation in programs.
- Examine the tradeoff between costs and benefits in your target audiences' minds.

- Look at competition (in the target audiences' minds) for the behavior being promoted.
- Intervention programs are dynamic, and ongoing process evaluation is crucial.

Using Formative Evaluations to Identify Target Populations

Elizabeth Howze, Ed.D.
Chief, Health Interventions and Translation Branch
Centers for Disease Control and Prevention

Background

Diet and physical activity patterns are two major causes of death in the United States today. The Centers for Disease Control and Prevention (CDC) began the Nutrition and Physical Activity Communications Project (NuPAC) to see if a communications campaign could be put together focusing on nutrition and physical activity that would work within their existing and planned activities. NuPAC efforts for the past year have focused on formative evaluation, which has provided the foundation for the communications efforts. The behavior changes we are trying to communicate are:

- Choose a diet with plenty of fruits and vegetables.
- Choose a diet low in fat.
- Accumulate 30 minutes or more of moderate-intensity physical activity over the course of most days of the week.

The process outlined in CDC's health communications wheel is being used to develop health communications activities around these objectives. We are almost finished with the first three stages—background research, communication objectives, and target audience. We will be going on to identify message concepts and pretest those to select the communication channels that we think will be effective in reaching our target audiences. We will be developing materials and a plan for dissemination, and we will continually assess what we do.

The Target Audience

People are classified by their stage of change for physical activity and dietary behavior, specifically making a change to lower-fat eating or increasing their fruit and vegetable consumption. For example, with physical activity, they are classified based upon whether they are thinking about making a change in their physical activity, whether they have started to make a change, or whether they have that behavior in place.

Based on a literature search, environmental scan and initial market analysis, the first cut of the target audience is people age 29 to 54 years, with middle income and middle education (those not completing high school or obtaining a doctoral degree or higher are excluded), who are contemplating or preparing to make a change in either their dietary behavior (with respect to fruits and vegetables and fat) or their physical activity. The initial group consists of whites and African Americans, but long-range plans will include other groups as resources are developed.

Focus Groups: Understanding the Target Audience

Sixteen focus groups were conducted in four cities across the United States to get a sense of what the target audience thinks about, to explore the importance of good health and what that means to them, and to explore the costs and benefits of healthy eating and physical activity, their knowledge of the long-term consequences of unhealthy eating (high fat and low fruit and vegetable consumption) and physical inactivity and the health recommendations in those areas, and whether they saw physical activity and nutrition as going together.

There was no variation across the groups by region of the country and very little variation by race. Family was a priority for everyone, and everyone described their lives as busy and stressful. Life stage influenced behavior; for example, people with younger children were much less able to make some of the recommended changes compared to people whose children were grown or who were in a different life stage for other reasons.

Participants saw spiritual, mental, and physical health as very intimately connected, and found being healthy very desirable. Many knew what they should be doing and had tried many times to do it, but putting it into practice was extremely tough. Chief obstacles included lack of time and lack of internal motivation. Participants thought children were very motivating, because the changes would benefit the participants themselves, the children and other people in their lives.

Next Step: Message Development

As for message development, the physical activity recommendation was the most eye-opening. Participants said they did not understand what "accumulate" meant and doubted that they would derive health benefits from less vigorous activity than that previously recommended. The challenge is to work backwards and figure out ways to convey the validity of the physical activity message without portraying it as exercise. Participants saw exercise as going to the gym rather than things one could do around the house, around the yard, or around the neighborhood.

Messages will be developed keeping in mind the holistic health, mind and body connection; the experience the audience has had with trying these behaviors; and the preferences of target audience members. For example, women wanted support and encouragement with concrete tips. Messages may focus on positioning change as a lifelong thing, rather than to attain a short-term goal as some participants think of it now.

The Knowledge-Attitudes-Behavior Model and Defining "Behavior Changes"

Tom Baranowski, Ph.D.

Professor, Division of Behavioral Sciences and Health Education,

Rollins School of Public Health,

Emory University

Assumptions in the Design and Evaluation of Dietary Behavior Change Interventions

- Theory provides the basis for intervention programs. An understanding of human behavior is organized as a theory and its related models will facilitate helping people change their dietary behavior.

- The best theory will account for the most variance in the targeted dietary behavior. But predictive efforts so far fall well short of 100 percent—a review of 21 studies by Stefflaw found that the models used predicted roughly 20 to 21 percent of the variance in the dietary behaviors.

- An intervention based on the best theory will result in the most change and the most readily obtained change. An important issue here is that the intervention does not work directly on the targeted behavior; it affects the mediating variables and those mediating variables are theoretical constructs.

How do we select a theoretical model for use in the design and evaluation of a program? First, it should be applicable to the dietary behavior problem of interest to us; not all theories are applicable. Second, there should be some prior likelihood that the theory predicts the targeted dietary behavior; some do a better job than others. Third, there should be some likelihood that the intervention is based on the model, and that it impacts the mediating variables and the targeted dietary behavior. Finally, the same model should be used to develop and evaluate the intervention.

The Knowledge-Attitudes-Behavior (KAB) Paradigm

There are two major constructs in the KAB model: knowledge affects attitudes and attitudes affect behavior. The model is very interpsychic; it is all based on the individual, and includes no environmental factors. Essentially, the model says that everything is under volitionary control, and that is not always the case. The assumption is that increases in knowledge lead to more positive attitudes, which affect behavior. A variant is that increases in knowledge affect behavior and increases in positive attitude lead to more likely and more frequent targeted dietary behaviors.

Is Knowledge Related to Attitude?

Knowledge is multifaceted. The many different kinds of knowledge include instrumental knowledge and motivational knowledge. Once knowledge is categorized, we move beyond the KAB paradigm because we have gone into motivational issues and different theories that might predict the behavior. A key issue is that what is potentially motivating knowledge for one group may not be motivating knowledge for another. For example, for someone who is 60 or 70 years old, preventing heart attacks and cancer may be particularly salient and motivating, whereas for teenagers those are far away issues and are unlikely to be motivating.

Several studies of nutrition education programs have failed to find a link between knowledge and attitudes. For example, the Teach Well project, which is a nutrition education program within school systems, included a measure of knowledge that combines instrumental knowledge and motivational knowledge. When this knowledge variable was correlated with all the other psycho-social variables, not one correlation was statistically significant—no relationship between knowledge and attitudes. Similarly, with the Gimme Five project, the highest correlation between the knowledge variable and psycho-social variables was about .21; most of the relationships were about .1 to .15. When all of the variables were put into a model predicting fruit and vegetable consumption, the knowledge variable was not predictive of behavior.

Is knowledge related to behavior? Well, probably some knowledge is necessary, but it is certainly not sufficient. We have not clearly defined what necessary knowledge is. A recent article in Health Education Research found that knowledge variability is not particularly important in understanding dietary behavior change. The study examined if knowledge one year predicted dietary behavior change the following year, and found that knowledge accounted for less than 1 percent of the variance in the behavior change.

Is There a Relationship Between Attitude and Behavior?

As early as the 1930s, there were reasons to believe that attitudes are not related to behavior (see Bettinghaus, 1986). When relationships are observed, only a small percentage of the variance is accounted for by the attitudes. Some theories, such as dissonance theory, argue that one needs to induce a change in behavior to change attitudes, rather than vice versa.

Time to Dispense with the Knowledge-Attitude-Behavior Paradigm

Interventions manipulating knowledge or attitudes have not usually resulted in behavior changes. The bottom line is that knowledge-attitudes-behavior provides a poor model for designing or evaluating behavior change programs. The KAB paradigm does not reflect our best understanding of the influence of human behavior; any of the theories presented earlier provides a much more interesting alternative to trying to explain behavior and include a lot more influences than knowledge and attitudes. Also, the relationships implied do not seem to work often. When they do, they account for a very small percentage of the variance in behavior.

Alternatives to the Knowledge-Attitudes-Behavior Paradigm

We need to develop a deeper understanding of the factors that affect dietary behaviors. For example, if we can understand why kids are not eating fruits and vegetables, then perhaps we can design an intervention that affects fruits and vegetables. We have developed a family reciprocal determinism model that examines how family members relate to one another (merging family characteristics) and in turn how the family relates to characteristics of the environment. It's important to include environmental factors but many psycho-social models, particularly KAB, ignore them.

Another aspect of kids' fruit and vegetable consumption is the availability/accessibility paradigm. If fruits and vegetables are not at home, there is no way that kids can eat them. Furthermore, if they are in the home but not in an accessible form, it is unlikely that a third grader is going to go home, open the vegetable drawer, peel some carrots, slice them, get their favorite dip and make a snack. We have tried to identify the factors that affect the availability of fruits and vegetables; an interaction of environmental, personal, and behavioral variables affects each succeeding step in the availability/accessibility paradigm.

Choosing Evaluations That Fit the Intervention and Stage of Development: Breakout Sessions

Doing the Best Evaluation Possible

Anne Murphy

Nutrition Education Evaluation Consultant

University of Michigan - Flint

This session's focus, rather than evaluating print materials as originally planned, is how to do the best evaluation possible while overcoming the many barriers to it, with the least amount of error and at the lowest cost.

The Distinction Between Reporting and Evaluation

Reporting is describing what was done, how much of it was done, what methods were used, how many people participated, and their characteristics. Evaluating is how well the program did, what difference it made, what changed and for whom, to what extent was the change maintained, and at what cost. Evaluation also delves into what caused the change. Was it your information? Was it your program, or another one? Or was it the attention, encouragement and support people got from your program, rather than the intervention itself?

Evaluate Before Developing the Program

The most important time to start evaluating is before the program is developed. Needs assessment is particularly important. We have the best chance of doing a good job if we find out what people want beforehand—and the form they want it in. Evaluating before developing a program or materials is particularly important if resources are limited. Try to save money by not making mistakes in the first place.

For example, a food safety specialist wanted to do food safety training with child care providers. She needed to know what form the training should take—such as videotapes, workshops, a newsletter—and what food safety topics should be covered in it. We created a survey that we sent to a random sample of child care center teachers and family day care home providers. We asked them two primary questions: what do they want to learn and how do they want to learn it?

They wanted information related to safe food within the unique environment of the child care setting. The form they wanted it in was a newsletter. Their second two choices were a booklet and fact sheets; they rated a teleconference, hotline assistance, audiovisual tapes, and workshops low. This is a simple example of how we could have wasted a lot of money on a set of videotapes that wouldn't have had the usage that we would have liked. The child care providers wanted a booklet or a newsletter. That's what we provided, and the response has been tremendous.

Saving Money

The more money spent on evaluation, the more money saved in programming dollars. Perhaps we are so resistant to evaluation because we fear knowing what we are doing wrong. Good evaluation takes the guesswork out of program planning. There are a number of ways to save money when conducting evaluation.

Take advantage of built-in evaluation mechanisms. For example, program staff working on the WIC/EFNEP breastfeeding initiative wanted to know the most cost-effective way to target their program. Should they go in before the baby is born? Should they make one visit? How does that work compared to three visits before and in the hospital and just one followup and then phone contacts? Different instructors were taking different approaches, so they had a built-in experimental design. All they had to do was collect information on the different factors (number of visits, home versus phone contacts, etc.), then separate out the data by those factors.

Use existing instruments. Often, an existing instrument can be used to help save money, if it is tested within your own group. For food safety items, we just got the USDA food quality initiative grant to create a database of items. Custom instruments can be created by accessing the database, pulling the items up by topic, aligning the survey the way you want, and then choosing from different sets of directions or constructing your own.

Cooperate with others. Sometimes, states can create a consortium to conduct evaluation and get more data for less money. For example, the National Food Service Management Institute formed a consortium of 14 States to conduct needs assessment with 11 populations for the Nutrition Education and Training (NET) program. Pooling their resources and developing instruments that they all shared was much less expensive for the States than developing separate instruments. Each state tailored the instruments for their populations (e.g., Mississippi uses the title "food service administrators" rather than "food service directors"), and added a few items unique to their needs.

Do some of the evaluation yourself. Hire out the parts for which you lack the time or expertise. For example, some administrative tasks can be handled in house, including photocopying questionnaires, stuffing and mailing envelopes, and doing some data coding.

Use a random sample. If the population is 5,000 people, nothing is gained by surveying everyone—a random sample of 1,000 or even 500 will provide the same information as surveying the whole group, but it will cost a lot less.

Don't collect or analyze anything extra. People frequently collect too many demographics. Don't collect data unless the information will be used. Asking about income, in particular, makes people uneasy. If a random sample is used, race is not needed unless the data will be analyzed by that factor.

Other Evaluation Tips

Use the most direct measure possible. Sometimes we ask teachers what kids like to learn. Sometimes we ask EFNEP instructors what works best with EFNEP clients. The best way to do it is to get to the actual audience.

Use information gathering first. Talk to the audience, perhaps in a focus group, to find out what language they use, what response categories need to be added, etc. Once everything learned has been incorporated into your instruments, go back to the audience and pretest them. One focus group isn't enough because the first focus group just leads you to everything that needs to change. Ask the second group about the changes made.

Gather data in a format easily used. For example, if the percent of kids whose knowledge improved must be reported, make sure data collection and analysis is set up to answer that question.

Reviews start with you. After an instrument is constructed, sit down and complete it. You'll find all sorts of

things wrong. Then test it with expert reviewers and with the target group.

Think about what you want the final product to be. When the evaluation is finished and you want to talk to others about it, what do you want to talk about? What is it that you want to know? What can you do a good job on?

Sources of error include dishonesty (about what is eaten), inaccurate estimation (even dietitians aren't good at that), inaccurate recall, and misinterpretation of items. After the data is collected, sometimes we make errors in interpretation. Sometimes we find that we did not have enough people to be representative of the population.

Evaluating Materials

Materials development is so expensive that it is important to conduct assessments to make sure the developed materials are what the audience needs. When evaluating these materials, make sure what's being taught is what's being measured. If the program goal is to increase knowledge, don't get mixed up and say, "we taught them knowledge so now we're going to have behavior change." Often we teach knowledge, then assess behavior.

If we want to change behavior, we need to teach people in a way that they learn the skills necessary to change behavior. We also need to teach them the way they want to be taught. For example, kids want to learn with games, demonstrations, experiments, and food preparation. Information presented by the teacher comes out dead last—yet our curriculum development still includes a lot of materials for teachers to teach.

Both the evaluation and the materials need to be consistent with program goals. If they are, evaluation plans and instruments can be developed in tandem with the materials. People are very hesitant to develop an evaluation without looking at the materials to be used because they are afraid it won't be consistent. But, if your materials are consistent with your goals, they teach what you want to teach. If your instruments are also consistent with your goals, they will test what you want to teach.

When evaluating materials, make sure the interest level of your intended audience is assessed. Not everyone is as interested in nutrition as we are. Also, sometimes when we try to be simple we become patronizing. Ask them if anything in the material was insulting or made them feel bad.

Dealing with Varied Audiences and Delivery Mechanisms

How can an evaluation be developed that will work across States and for a program with many audiences and many methods of delivering different materials? Simply document the differences. Then effect can be examined across delivery method, across length of intervention, and across curricula. That analysis will provide information about what is working the best. Just don't ignore the variation. For example, if you're evaluating knowledge and some materials didn't have knowledge in them, that's a little unfair. For this type of evaluation, a database could be developed similar to the food safety database we're developing where instruments can be custom designed. A set of items would be developed that are coded to the different materials, then the appropriate items would be used for each evaluation.

Evaluating Social Marketing Promotions

Craig Lefebvre, Ph.D.
Vice President and Chief Technical Officer
Prospect Associates

Social Marketing in the Context of Health Promotion

There are five strategies for health promotion in the 1990s: 1) strengthen community action, 2) help people develop the personal skills to have a healthy lifestyle, 3) build healthy public policy, 4) create supportive environments in which people can practice their skills, and 5) reorient health services so they become more focused on prevention and health promotion issues, rather than only on disease management issues.

To execute these five strategies, health promoters need to play three roles:

1. ***Enablers.*** Working with individuals, communities, work sites, and different organizations to foster healthier environments and public policies, plus the development of personal skills.

2. ***Mediators.*** The 1990s is the age of coalitions. Health promoters need to mediate and negotiate between very diverse groups of individuals and organizations to form coalitions around achieving healthy objectives for the community.

3. *Advocates* for each of the five strategies described above.

Social Marketing in Health Promotion: What it Is and What it Isn't

Social marketing is not a theory for health promotion, nor is it health education with new words and new enthusiasm. It is not focused only on changing individuals' beliefs, attitudes and behaviors. Social marketing is a framework in which to approach understanding and addressing health and social issues through environmental changes, public policy, and behavior changes for some segment of the population. None of us who do this for a living is naive enough to assume that simply putting posters up in a classroom is going to change anybody's behavior. We're also not naive enough to believe that simply changing a regulation means that everyone then conforms to it. The multi-factorial issues involved in behavior change must be addressed.

Social marketing emphasizes research and audience analysis. It uses a product, price, place, and promotion mix to very specifically and very precisely target the audiences of interest. When you see a public service announcement (PSA), a poster, or a pamphlet, do not think you are seeing social marketing. When you see a comprehensive program in which these things are part of the tactics used to reach a target audience, then you are seeing social marketing.

Three Models of Behavior Change

The following models are useful when developing social marketing programs, because they help us to understand what is likely to happen when a behavior change is introduced into a population.

Diffusion of Innovations

Thousands of studies of a variety of behaviors illustrate that whatever the innovation, most of the time adoption of an innovation follows an s-shaped curve. At the beginning of the process, very few people are aware of the behavior change they are being asked to make. Over time, more and more people start engaging in, or adopting, that behavior. There is usually a critical point on the s-shaped curve around 25 percent. When about 25 percent of the population has adopted a new health behavior, the curve begins to accelerate quickly.

What we try to do is get that curve as high as we can to reach the most people who will adopt that behavior, and then sustain it over time. Contrast that with a "fad" curve, which is when a few people start adopting a behavior, but less than 25 percent do, and then people start dropping off and the curve basically goes away. Until about 25 percent of the population makes the behavior change, there is a slim chance of maintaining the behavior in the population.

Social Learning Theory

Social learning theory encompasses different ways to make people change their behavior, the process for doing so, and some things to think about to help people change their behavior.

When people are at a very low level of awareness about the health behavior, building awareness is the key focus. As more and more people start getting involved in the behavior, the program strategy shifts to providing information. Next, it gets people to try the behavior, then it reinforces the behavior. Many of us think that once we get people engaging in a behavior, we can stop. The whole idea behind social learning theory and behavior change theory is that people must be reinforced. When that step is left out, people relapse.

As we move up the s-shaped diffusion curve, the next step becomes contextual support or environmental and public policy changes. By this point, we are into the population group that diffusion of innovations terms the late majority or the laggards. These are the people who are really tough to change; the environment needs to start changing and their friends and people around them need to change before they will change, too. Finally, sustainability needs to be built in to make sure that the behavior can be practiced repeatedly to avoid relapses.

Stages of Change

Stages of change is another way of segmenting the population based on their behavior. What we try to do with communications and marketing programs is address specific segments of the population. One example is people who might be precontemplators and would need awareness-building strategies to start. People who are in the contemplation stage need to have more information to get them to buy into the value of the behavior. As they move into the action phase, they need to have opportunities to try these behaviors and have their efforts reinforced.

Consumer-based Health Communications (CHC)

A model for creating messages to target audiences is Consumer-based Health Communications (CHC). The CHC model stresses understanding the consumer's reality as well as epidemiological and clinical research. It focuses on consumer research to get a vivid picture of what the audience's values, beliefs, attitudes, desires, needs,

and current behavior are. Then we can better create a message that fits their lifestyle and their psychology, not just our scientific facts.

We use CHC to create message strategies that help us define a specific action that we want people to take. This action will be a step toward the health behavior. In the 5 A Day program, for example, the target audience currently eats roughly 2.5 to 3.5 servings of fruits and vegetables each day. We tell them, "You're eating three and that's great. Just add two more," rather than telling them to eat five servings.

The CHC process also involves looking at the kinds of benefits we promise the target audience; the supports we use for those benefits; and the times, places and states of mind—or "openings"—we can use to reach people. Finally, the process includes identifying the image or tone the campaign should have so that the target audience will react to it in a favorable way.

Evaluation Issues

For social marketing programs, we need to think about what outcomes we are evaluating. With the 5 A Day program, many would try to evaluate the recommended health behavior: eat five fruits and vegetables a day. They would probably measure this in the population as a whole. That's the wrong evaluation point for a social marketing program. We are not trying to get the population to eat five servings of fruits and vegetables a day. We are trying to get a specific target audience to add two servings. We need to be very clear when we are developing social marketing programs about what a marketing communications program can expect to achieve.

Formative research is another important part of developing a social marketing program. The first step is looking at the audience and segmenting it into the most practical and feasible groups given the program's resources and objectives. Once the audience segments are identified, we spend more time trying to understand each segment's needs, beliefs and attitudes. We also compare the target segments with those already engaging in the desired behavior. We then spend time trying to understand the channels that people will listen to and respond to. Which ones are credible? Authoritative? Which ones are most accessible? We also look at the environment to figure out what can help send the message and reinforce it.

Concept testing is another part of formative research. Once we have some program ideas, we try them out with some target audience members before developing them. That way we know in advance if the materials or services are something the audience is likely to use. Market testing on the products to be developed is another step in the formative research process. When we think we have come up with a great new product, we write out a paragraph to describe it and test it with audience members to see whether they are even interested in it. We test out promotional strategies in a similar fashion.

Product pricing is another area of market testing. Pricing is not necessarily how much we will charge for the product. It includes other costs, such as social costs, time and effort, etc. The other side of pricing is benefits. What are the positive things that people can receive from engaging in this behavior, and how do we make them tangible and real? All the obstacles, or barriers, can be removed but if people aren't motivated to do something, they won't. It's important to focus on benefits and positive motivations when working on product pricing.

Program monitoring data is most helpful when it is available quickly—within a few weeks. That way, the program can be changed to meet new needs. A program monitoring system tallies targets and message types.

To whom is the message being delivered? What is being said to them and at what intervals? Program monitoring should be time based. This chronological point of view allows you to understand, step by step: 1) how your program is rolling out, 2) how it is being refined, and 3) what impact that has on how you are reaching and effecting changes in your target audiences.

Program monitoring can include going out to the message receivers and getting feedback from them—whether they are intermediary gatekeepers or target audience members—through customer satisfaction surveys and other techniques. Many studies also look at indirect and unintended effects. For example, the early cholesterol education messages were along the lines of: "Eat low-fat diets. Low cholesterol levels are best." Many parents reacted by putting infants and very young children on very low-fat diets, raising concern among pediatricians. In response, the message was modified to emphasize that the behaviors were not recommended or necessary for children under 2 years of age.

Obstacles to Evaluation

There are five common obstacles to conducting evaluation in social marketing programs:

1. ***Research*** is reserved only for the "big" decision, which is usually about money. But most of the decisions made—about who will be targeted, how they will be reached, and what the message is going to look like—are all big decisions in social marketing programs.

2. ***Survey myopia.*** Research does not have to be done by surveys; qualitative research has a role to play as well. Surveys and randomized designs keep many people away from doing this kind of evaluation; it doesn't have to be that way.

3. ***Research is expensive.*** There are lots of inexpensive ways of getting research done. For example, college or university students often need practical experience doing research, and are available free (for class credit) or at relatively low cost.

4. ***The sophisticated researcher myth.*** Program staff can do solid research to answer some basic questions and pick up very good leads about how to structure a program.

5. ***Most research is not real.*** We commission these big studies that result in voluminous reports that sit on the shelf. Social marketing research needs to be action-oriented. It has to lead to very specific and very tangible options and recommendations for modifying the program to make it more effective. Doing this type of research requires quick turnaround.

Examples

The *Know Your Cholesterol* campaign, part of the Pawtucket Heart Health Program, began in 1985. It included a screening education program, dietary recommendations and a follow-up protocol (mail and telephone surveys of people identified at screenings as having high cholesterol levels). The campaign products included a nutrition self-help kit that all participants received, tip sheets about nutrition, the "rate your plate" score, a restaurant program with menu labeling, and a grocery-shelf labeling program. The campaign also included outreach to physicians, telemarketing, newspaper and radio public service announcements, a 4-week series of news

paper columns, targeted direct mail, various presentations, and community events including cholesterol screening, counseling, or referral events (SCORE).

Every person who went through events that included a behavior change component filled out a process tracking form. It included the date, the person's name, address, telephone number, date of birth, and gender. We collected them on more than 100,000 people. It allowed us to do many interesting market tests and market tracking studies.

The tracking system was used to monitor what was happening in the campaign and with participants. For example, the SCORE events were promoted to 500 35- to 55-year-old men through direct mail. We tracked how many of them came in for a screening event during the month. Three hundred physicians received mail pieces educating them about the new consensus development guidelines, 165,000 people were getting behavior-change messages through the media, and about 1,500 people got screened in the first 4 weeks of the program.

How people fall in the different categories can also be tracked, which helps assess how many people are changing behavior. For example, we found that more people lowered their blood pressure level than increased it during the first 4 weeks (and during the subsequent 2-month followup period). We also tracked participation by site, which told us where to concentrate our efforts as we rolled the program out, and provided demographic information about participants.

Another example of tracking comes from the 5 A Day program, where media coverage and public awareness are tracked. Are the media covering the 5 A Day message the way it is being put out to them? Or are they putting their own spin to it? How many people are aware of 5 A Day and the 5 A Day messages?

For the media content analysis, a sample of 1,100 news clips appearing between July 1992 and October 1993 was coded for key program messages and the tactics used to disseminate the message. Eighty-six percent of the clips had 5 A Day core messages in them. Forty-one percent mentioned specific health benefits for eating five fruits and vegetables a day. Twenty-eight percent included healthy eating strategies. Five percent talked about barriers to eating five fruits and vegetables a day. From this we learned that our core messages were getting good coverage, and health benefits were receiving a fair amount. But relatively few stories were telling people how to eat five servings of fruits and vegetables a day.

Thinking about the media as a target audience, the challenge is to get them to start giving people more ideas about ways they could stay healthy. We would prefer they spend less time on specific health benefits because the consumer audience isn't worried about health benefits.

Audience tracking studies have been conducted for the last 4 years for 5 A Day, beginning with a baseline in 1991. At baseline, only 8 percent of people in the U.S. population knew that five or more was the number of servings of fruits and vegetables a person should eat each day for good health. By the following September, roughly 22 percent knew; 29 percent knew in 1993 and it leveled off at 27 percent in 1994.

Going back to diffusion theory, we're at the point where something may start happening if 5 A Day doesn't become a fad. As we saw on the earlier graph, 5 A Day has reached the critical mass of 25 percent of the population. If it's a fad, we will begin to see decreases in awareness. Otherwise, increases will continue.

Conclusion

People also look at these numbers and say, "You're not doing a whole lot out there; it's only 29 percent." Twenty-nine percent represents more than 60 million people who now know that they are supposed to be eating five fruits and vegetables a day. Social marketing often gets held to accounting measures that people are used to using with small-group research (20-60 people). If we only get 20 percent of 60 people that change, we say, "That was bad work, bad study, bad design." If we get 20 percent of 260 million people, it's a whole different number. The 5 A Day campaign saw almost a 400 percent increase in the number of people aware in 2 years. Population-based behavior changes take time.

Program Evaluations in the Community

Adrienne Paine-Andrews, Ph.D., Associate Director

Kari Harris, M.A., Research Associate

Work Group for Health Promotion and Community Development

University of Kansas

Background on Coalitions

Coalitions take a community approach to addressing complex problems, such as cardiovascular disease. Schools, people from different sectors throughout the community, health organizations, and businesses all come together to identify what they can do in their sector to make a difference. What are the specific programs, policies, and practices that they put in place to effect the program mission? A nutrition-related coalition could be looking at preventing cardiovascular diseases or some cancers, enhancing birth outcomes or child development, and improving nutrition for older adults. Dietary risk factors play a part in all these issues.

Examples used in this presentation are from our work with Kansas LEAN, which is a statewide coalition initially started with the Kaiser Family Foundation Project LEAN efforts. They found local funding from the Kansas Health Foundation. From Kansas LEAN came the Kansas LEAN School Health Intervention projects in Dighton and Salina, Kansas. The projects focused on a community coalition effort to support components such as nutrition education, modifying school lunches so they are healthier, and some physical activity opportunities in the community. These two projects were so well received that the funding agent, along with the state health department and other agencies, found funding to replicate these initiatives in six more Kansas sites.

Evaluating Coalition Activities

There are many things to consider in selecting your evaluation questions: 1) What do members and leaders of the community partnership want to know? 2) What are you interested in learning about the effects of your project? 3) What are the requirements from grant makers and others? 4) What are their interests—are they interested in a process, or the outcome? Also consider the resources that are available to address the questions.

When we approach evaluating a community coalition or a partnership, we need to consider a number of questions around process measures, intermediate outcomes, and more distal outcomes. We look at the process of an

initiative from its beginning to its end. For example, was the community mobilized to reduce risks for cardiovascular disease (CVD)? Did the people who represent the different sectors (schools, health organizations) take action toward their mission? Did they take steps toward implementation? Were members satisfied with the partnership? Were they satisfied with the planning? The leadership? Were resources used appropriately?

For our projects, the community had key questions for intermediate outcome measures. First, we asked what changes in the community resulted from the initiative. What new programs, policies, and practices took place in the initiative that were consistent with the mission? It takes a while to get to the bottom line, for example reducing cardiovascular disease or cancer. Intermediate outcomes help illustrate how much progress is being made. Other questions assessed whether the changes were important to reducing risk for CVD, and what critical events seem to spur rates of community change.

Questions for more distal outcome measures included the following: Is there a change in behavior related to risks for CVD? Does the initiative have a community-level outcome related to the risks for CVD? Is the community-level outcome related to changes facilitated by the initiative?

We used a monitoring system to get at the process in some intermediate outcomes, such as community change and community action. We used surveys to measure member satisfaction and outcomes, such as how important the community changes are. Most of what is presented refers to some time-series designs, such as looking at change over time or growth over time. When possible, we used group designs to look at comparison communities.

Detail on Measurement Techniques

Monitoring system. The monitoring system looks at process and intermediate outcome assessments. It addresses two primary questions: Was the community mobilized to reduce risk for CVD? What changes in the community resulted from the initiative?

The monitoring system is a way of tracking major events and accomplishments in a coalition. It is useful for understanding the initiative, deciding where to focus efforts, promoting awareness of accomplishments, recruiting support, and securing grants. It's a log system where people involved in the initiative record what they are doing. Log entries are clarified and categorized by the people in charge of monitoring.

The data is then graphed and provided to coalition leaders. For example, in the category of community change, changes in programs, policies and practices related to the program mission would be recorded, such as changing school lunch menus to reduce fat, adopting a nutrition assessment with students, and changing supermarket practices. The graphs provide evidence for the funders and community members of the types of changes that are being made in the community to support the effort.

Assessment of how important changes are to reducing CVD. We sent out a survey to coalition members and experts in the field to assess how important each community change is for the goal of reducing cardiovascular disease. The experts can use the literature to assess the importance of each change, and the community members often know the kinds of things that need to change in their community. What's important in one community might not be important in another. Each change is rated on a five-point scale (five is the highest rating). The changes are then multiplied by the rating and graphed to get a sense of what changes are most important. In our communities, most of the changes were rated "4" or "5."

Satisfaction surveys. It's important to learn if coalition members are satisfied with the initiative, but getting coalition leaders to ask members how things are going is often hard. We used a survey to assess member satisfaction, thereby informing the partnership of the strengths and weaknesses of the coalition. The survey results provide a good opportunity to celebrate the coalition's successes and make corrections to address problems.

Community-level indicators for CVD. These indicators are ways to obtain information about the more distal effects of the prevention effort (and some measures of a community's health potential), by counting opportunities in the community for low-fat eating, physical activity, smoke-free living, etc. They provide a strong picture of the initiative's effects on cardiovascular health in the community as a whole. Categories of indicators include the following: 1) information and skills building; 2) point of purchase information; 3) healthy menus in schools and work sites; 4) environmental changes, such as healthy alternatives in vending machines; 5) policy and regulation, such as looking at publicly funded food programs that follow nutrition guidelines; 6) formal work-site policies that support healthy eating.

We can also look at the relationships between intermediate outcomes, community change, and community-level outcomes. For example, as community changes increase, we see a decrease in a community-level outcome: the percentage of adults who are overweight. As the community changes level off, there is less of a decrease in that outcome.

Interviews with key participants. This instrument looks at process, intermediate outcome, and the more distal outcomes. It asks, "What critical events seem to spur rates of community change?" It provides a way of learning about important events in the life of the initiative, helps identify factors that affect the initiative's success, helps identify negative side effects, and provides a history of the initiative. We ask people to list each critical event. Then we ask them why it was important, what the context was, what the consequences were, and so on. We also ask people to reflect on the lessons they learned about the initiative, and to look at future directions.

Lessons Learned

1. The primary purpose of the evaluation is to support movement, not to judge success or failure.
2. Evaluation should begin early, and be an integral part of the development process.
3. Evaluation should be participatory and collaborative. The monitoring system can help establish and maintain effective functioning of the initiative; it provides a picture of where energy is being spent and whether the initiative is producing the desired effects.
4. Feedback to coalition members and funders should be provided at regular intervals, especially early in the initiative's development. We provide data monthly early on, and then quarterly after a couple of years.
5. Community leadership can use evaluation information to attract and maintain support and resources.
6. Grant makers can use evaluation information to encourage productivity and accountability.

Choosing Appropriate Dietary Data Collection Methods to Assess Behavior Changes

Alan Kristal, Ph.D.

Department of Epidemiology, Fred Hutchinson Cancer Research Center

University of Washington

It is important to separate measuring dietary change from measuring diet. There is a big difference between the two. The science of measuring dietary change is in its infancy. Most of the research on measuring diet is in three areas:

1. Epidemiology (particularly trying to measure a relationship between diet and health outcomes)
2. Nutritional sciences (mostly trying to understand the relationship between diet and some kind of underlying biological mechanisms)
3. Public health (to describe broad-scale trends in large populations)

But the science of dietary change per se, especially in the context of intervention trials, is relatively underdeveloped.

Questions for Evaluating Diet Intervention Studies

What kind of intervention is it? Is it a clinical intervention, with multiple groups over a long period of time, or is it a public health intervention where the intensity per individual is relatively limited? The effect size for clinical interventions is large and hopefully fast. In a public health intervention, it's small, and at best, it's gradual.

What do you expect to happen because of it? For example, consider 30 percent energy from fat. In 1990, we started at 38 percent. Do you expect a relatively fast decrease where the change is maintained, or is it going to be some squiggly line where over time it gets down to 30 percent? The underlying process that you expect to take place is going to affect how you evaluate.

How can expected outcomes be operationalized in real measures? Often, we think of outcomes in very global, nonspecific, and meaningless ways. We think of dietary adequacy, food security, and nutrition knowledge. At the community level, we look at availability of healthful foods, availability or existence of nutrition programs, or media coverage. Individual measures can be aggregated to get community measures, but they need to be kept separate conceptually. A survey on a random sample of the population can be aggregated, but it still needs to be based on individual assessments.

What we can really measure when doing nutrition interventions falls into four groups: 1) physiologic measures; 2) behavioral measures—usually based on self-report and including nutrient intake, food use or dietary habits; 3) cognitive outcomes—changed knowledge, attitudes, attention or stage of change; and 4) environmental measures, such as measurement of supermarket shelf space.

Measurement considerations include:

1. Participant burden—if the burden is high, they won't cooperate and the sample will be biased. Fatigue will be an issue for those who do complete it.

2. Complexity of administration—sometimes self-administered instruments won't work; sometimes a trained dietitian is needed to handle data collection.

3. Complexity of analysis—Know how the study will be analyzed. Sometimes it's just too complex.

4. Scientific measurement issues in terms of validity and reliability.

5. Responsiveness—does the instrument measure the behavioral target and is the behavior going to change enough that statistically the measure can pick it up. Sometimes, shorter instruments do a better job than longer ones. It depends on what each is measuring. For example, in one study using a 98-item food frequency questionnaire and an 18-item diet habits questionnaire, the responsiveness for the diet habits questionnaire was higher.

How will it be assessed? Separating the effect of the intervention on the measure versus the effect of the intervention on behavior is very difficult. It could be done with objective measures, but with self-report, bias is always an issue. Problems with validity and reliability abound with the current dietary intake measures.

Comparison of Standard Dietary Intake Measures

A number of measures can be used when trying to evaluate how an intervention has changed people's usual eating habits. Major considerations when making a choice among the following measures include cost and bias—to what extent do people tell you what you want to hear, what they think they should be eating, and/or what your intervention told them?

24-hour recalls provide nutrient information and are useful with an intervention-control design where the outcome can be measured by comparing the mean of one group to the mean of the other. If individual outcomes are of interest, the measure is problematic because intraindividual variability is so high. Multiple 24-hour recalls would be needed, which is very expensive. Unannounced 24-hour recalls have relatively low bias: People may preferentially forget what they ate, but they can't change it. However, they can't always say exactly what they ate. They may not remember details. Compliance is high with this measure, because people don't have to prepare anything—just spend 20 minutes answering questions.

A state-of-the-art system for collecting 24-hour recall information is the University of Minnesota's Nutrient Data System. This computer system prompts for information about each food, providing standardization and eliminating the need for clinical dietitians to do the interviews.

Four-day diet records. This method is very expensive and time consuming for participants and evaluators. Also, the bias is extraordinary because people do what they think the interventionist wants them to do for the four days. Other problems include a potential lack of understanding about what was eaten (e.g., was the chicken breaded?) and literacy. Compliance is horrible.

Food frequency questionnaires are inexpensive and useful because they generate information on nutrients

and usual food use. However, people don't really know the answers, i.e., if they are asked how often over the last 6 months they ate broccoli, they cannot retrieve the answer from memory; they have to construct it. The most easily biased dietary assessment tools are food frequencies. Other limitations are the extent to which people can describe what they are eating and literacy. Compliance is fair; the questionnaires generally take 20-40 minutes to complete.

For intervention research, standard food frequency questionnaires may not be useful because they will not be sensitive to the behavioral target of the intervention. For example, if the intervention is fat modification, most food frequencies don't have enough specificity to detect if people did what the intervention told them to do unless they are customized. Also, a summary section is needed to allow adjustment of the data; otherwise, the more foods asked about in a category, the higher consumption will be.

An alternative to traditional food frequency questionnaires is the short food frequency questionnaire; these work well if the nutrient of interest is concentrated in a few foods, such as calcium or beta-carotene. They do not work well with macronutrients, except for alcohol. And they do not control for total energy.

Diet habits questionnaires typically include 18 to 25 items asking people about their dietary habits (using a four-point response scale); they can be used to scale dietary intake. A moderate amount of bias results from people providing "socially desirable" answers rather than true answers. Compliance is pretty good, and they validate well for percent energy from fat and for measuring fruits and vegetables. However, they are tricky to develop, requiring some expertise in psychometric theory and a lot of expertise in evaluation research. Another issue with these questionnaires is that the results aren't easily interpretable—instead of energy from fat, the result is a number like 1.36.

Measuring Stages of Change

Stages of change correlate well with other measures of diet. For example, some results from the *Working Well* trial on fat intakes showed those in precontemplation to decision stages had fat intakes of 39.3 percent. When they got into the action stage, it dropped to 37 percent. In maintenance, it dropped to 30.5 percent. The same thing is true for fiber, and for fruits and vegetables: The percentage is flat until the action stage, where it increases.

Closing Thoughts

Don't make up a measure and think it's going to work. Spend time pilot testing and understanding how you are going to analyze it. It's very difficult to come up with items that are meaningful, reliable, and valid. The measures used need to be appropriate to your population, in terms of language and in terms of the administration method. Make sure your measures have some responsiveness. They must be sensitive to the intervention both in terms of the effect size and whether they fit with the behavioral or cognitive target of your intervention.

Avoid science by democracy. Look at some experts' work rather than having everyone working on a project vote on what they like or don't like.

Finally, consider the money it costs and the time it takes to hire a consultant. Do some pretesting, and do some thinking to make sure the tool is going to work. Making certain your measures work when evaluating these studies is essential.