

Evaluating Nutrition Education: Steps for Moving Forward Together

Moderator: Carol Stiller, MS, RD, Senior Nutritionist, Supplemental Food Programs, Food and Nutrition Service, U.S. Department of Agriculture, Alexandria, VA

New Items for the Evaluation Toolbox

Carol Olander, PhD, Branch Chief, Family Programs Evaluation Branch, Office of Analysis, Nutrition, and Evaluation, Food and Nutrition Service, U.S. Department of Agriculture, Alexandria, VA

Good afternoon. I share one of your interests and responsibilities, and that is telling the story of nutrition education and how it works. A lot rides on how well we do this job. Evaluation can improve our education delivery and provide accountability to the larger public policy audience.

So why are we still talking about evaluation rather than doing it? Some of the conference speakers have already spoken about evaluation challenges encountered. Evaluations require resources—dollars, time, and technical expertise—that are often in short supply. They also take a certain amount of control over the educational setting in order to implement a strong study design and produce sound results. Given the effort required to deliver nutrition education, we are routinely challenged to strike a balance between service delivery and the assessment of that delivery.

While I don't have a blank check from USDA to meet all of your evaluation interests and needs, I can share information from a technical assistance tool that was just posted to the FNS website. This tool is called, "Nutrition Education: Principles of Sound Impact Evaluation." It is intended to help sort out the evaluation choices we each face and use our resources efficiently.

Evaluation research takes a variety of forms—each serving different purposes and addressing different questions. While the research literature includes alternative naming conventions, the pyramid shown here offers a uniform way to compare and contrast evaluation types.



At the base of the pyramid are formative evaluations. They typically occur upfront in the planning and development process; although, they may be used to revise existing approaches for use with a new audience or setting. Formative evaluation questions include things like: “Who is the appropriate target audience?”; “Are some messages going to be better understood than others?” and “Are some activities going to be more meaningful to my target audience than others?”

If you move up the pyramid, we are talking about something called process evaluation. The questions shift here. They focus on documenting in a systematic way the features and the scope of nutrition education activities. “What messages are actually delivered and how are they delivered”; “Who is receiving the nutrition education services?” and “How many people are reached by these services?” Answers indicate whether or not we have delivered the intervention in the intended way. They may also identify barriers and steps for getting over them.

Process information also provides one form of accountability. It allows us to document that we are using resources as planned and to demonstrate that we’ve used scarce resources carefully.

At the next level of the pyramid, we have the question of effectiveness: “Does nutrition education work or, specifically, does our nutrition education work?” This question is deceptively simple. An authoritative answer, however, requires research methods that rule out alternative explanations. And there is the crunch—being able to rule out those alternative explanations for the changes that are observed. This kind of study control is challenging to execute, but when you need to know whether or not a particular form of nutrition education makes, then the rigor of an impact evaluation is indicated.

On up toward the top of the pyramid. Because of the control and associated resources that an impact evaluation requires, many nutrition education assessments examine a related, but different, question: “Do clients change after participating in our nutrition education classes?” Are people in the area with a social marketing campaign different from those who aren’t exposed to campaign messages? Comparing the eating behavior, or nutrition knowledge, of students before and after they receive nutrition education is an example of an outcome evaluation. Outcome assessments provide information on whether or not the intended change occurred in relation to the educational services provided, but do not establish that the observed change is solely attributable to the intervention itself.

At the top of the pyramid is surveillance assessment. Politicians and other senior policymakers may be focused on the bottom line. Have we achieved the desired outcome? Is the problem less prevalent? For this group of stakeholders, broad population changes and various health indicators may be sufficient. Surveillance studies document how well, or how poorly, a population is doing. As such, they indicate whether or not further action is needed. They do not typically explain the observed status or changes in the target population.

The second half of my remarks focus on the evaluation principles themselves; there are eight of them. I want to say thank you to my colleagues at the Economic Research Service, and at the Extension Service who provided very thoughtful and useful comments on an earlier draft of this document. They are not responsible, of course, for anything with which you may disagree!

Why did FNS produce such a document? The Agency is occasionally asked what we mean when we describe or refer to impact evaluations, and what constitutes a strong impact evaluation. The principles in this document are intended to provide our perspective, our answer to that question.

It is worth pointing out, however, that the principles are neither unique to nutrition education, nor are they specific to a particular FNS program. Rather, they draw from a set of standards that are generally considered prerequisite to drawing credible conclusions about the impact of many types of educational, social, and economic initiatives. The principles are also consistent with the Government and Performance Results Act, and the Office of Management and Budget's Guidance for Demonstrating Program Effects. This may be more directly relevant to folks who are wearing a Federal hat, because these are the standards by which our programs are judged as being effective or not effective.

It is important to note that the principles are a technical tool. They do not direct nutrition educators to conduct impact evaluations, but rather discuss when it is most appropriate to conduct an impact evaluation. They also describe FNS' view on the optimal study features for determining whether or not an intervention improves eating habits, supports healthy lifestyles, or generally produces the desired results.

The document also contains an annotated bibliography that provides links to more detailed and topic-specific evaluation guidance on the Internet. That might be measurement options, databases to search for other relevant research, design issues, sampling options, or statistical techniques.

Now for more details on the principles themselves. The starting place is to make certain that the intervention to be evaluated can be. That means you have to think about the evaluation upfront as the intervention is planned.

Begin by answering the important basic questions. What are the objectives of the intervention? How large of an impact can you reasonably expect to observe when you implement? Why and how is that intervention expected to produce the change you anticipate? Will you be able to implement the education in a way that you intend to? Clear answers not only guide evaluation plans, but they also contribute to a stronger intervention.

Next, build on available research. Become very familiar with what is known and act on it. The goal here is to maximize our evaluation dollars. We are never going to be in the position to conduct strong impact evaluations of every activity that is implemented. So make sure

you are familiar with what is out there both as a guide to choosing the features of your intervention, but also choosing where to target impact evaluation resources.

Three, hold out for research designs with random assignment, but use them selectively. If it is important to conclude that a nutrition intervention does or does not have an impact—that is, alternative explanations for the results have been eliminated—then it is essential to use random assignment. Assign members of your target audience to groups who get the services or do not, and compare the differences or changes in them. Anything short of this leaves room for alternative explanations that you need to address in your analysis and interpretation of the findings.

While technically superior, experimental research in the real world may be difficult or even prohibited—so choose selectively. The WIC Program, for example, is one where the rules require that nutrition education be provided to all adults who receive WIC benefits. Even when you have the opportunity to use random assignment, you may decide not to invest in impact evaluations unless other features of your evaluation are sufficiently sound. For example, you need reasonable sample sizes, good outcome measures, and a replicable research setting.

Four, choose impact measures that fit the intervention and approach current standards for credible assessment. Five, observe standards for the fair treatment of study participants. Many agencies and universities have formal requirements for such procedures, and they are there to ensure that the rights of participants are protected. Six, carefully time the collection of impact data; that is, after start-up problems get resolved, but before implementation rolls out. You don't want your comparison or control group subjects to be exposed to the intervention inadvertently.

Seven, report all results—both positive and negative—accurately. Our knowledge of what works requires not only strong designs, but very thorough analysis and careful reporting. Any presentation of results needs to address how well the study controls for alternative explanations. Even with an experimental design, it's important to ask, “Did the control group folks inadvertently receive any nutrition education services?” and “Are the observed effects similar to what has been reported by other people?”

Eight, share the results in order to maximize their value. The purpose of evaluation is to inform future decisions and choices. So a commitment to sharing study findings ought to be built into the research planning process. This includes all results, even those that are modest or negative. Use a variety of communication channels. Publications are certainly desirable, but don't overlook posting to organizational websites, using listserves, making presentations, and using newsletters to share your experiences.

To wrap up with a few key points, remember that different types of evaluation address different kinds of questions. Evaluation planning starts with considering who needs to know what. But, if the question you are facing is, “Does this intervention work and does it make

a difference?” then seriously consider the principles for producing an authoritative impact assessment. Thank you very much.

Making Theory Work for You

Helen Chipman, PhD, RD, LN, National Coordinator, FSNE, Cooperative State Research, Education and Extension Service (CSREES), U.S. Department of Agriculture, South Dakota University, Brookings, SD

It is good to be here with you this afternoon to address theory and how it can work for you. When I first got out of school my eyes would glaze over when people started talking about theory. With experience in both practice and research, I have seen how these can come together, and the notion of incorporating theory into practice has become quite exciting for me.

Although, this may seem a difficult topic for some of you, as it was for me early in my career, I hope that as we talk this afternoon you can get as excited about recent developments that are occurring with respect to evaluation and theory as I am.

What Theory Is and What it Is Not

To prepare for this presentation, I started with Google and looked up the range of definitions associated with “theory.” I picked two, one at the level of a sixth grader, and the other at the level of the National Aeronautics and Space Administration (NASA). They are quite similar. One is a “general principle that explains or predicts facts or events.” The other is an “explanation for some phenomena based on observation, experimentation, and reasoning.” Common among the different definitions—and these are points that I would like to emphasize—are that THEORY IS an attempt to explain one or more facts or phenomena. Theory is plausible or scientifically acceptable. Also, theory is dynamic; it gains acceptance through testing in different types of circumstances.

Sometimes, we tend to think that there are only certain types of theories that are appropriate—those that have been around for awhile. But, it is important to remember that theory is constantly evolving. It is a model or idea that predicts what will happen under different circumstances. Again, words to keep in mind when thinking about theory are: explanation, plausible, dynamic, and potential for prediction.

At the same time, there are things that THEORY IS NOT. Theory is not fact. Theory is not reality, and sometimes we forget this. A given theory is not applicable in every situation. When considering different theories it is important to think through what the theory is and how it might apply in a particular situation. Is it appropriate for what we are trying to do?

Relevance – Why Theory Is Important

Why should we talk about theory? In recent years, program accountability and performance-based budgeting have been emphasized in government. However accountability is termed, it is not going away. Accountability is absolutely dependent upon strong evaluation and strong evaluation depends on good theory.

Program Theory and Evaluation

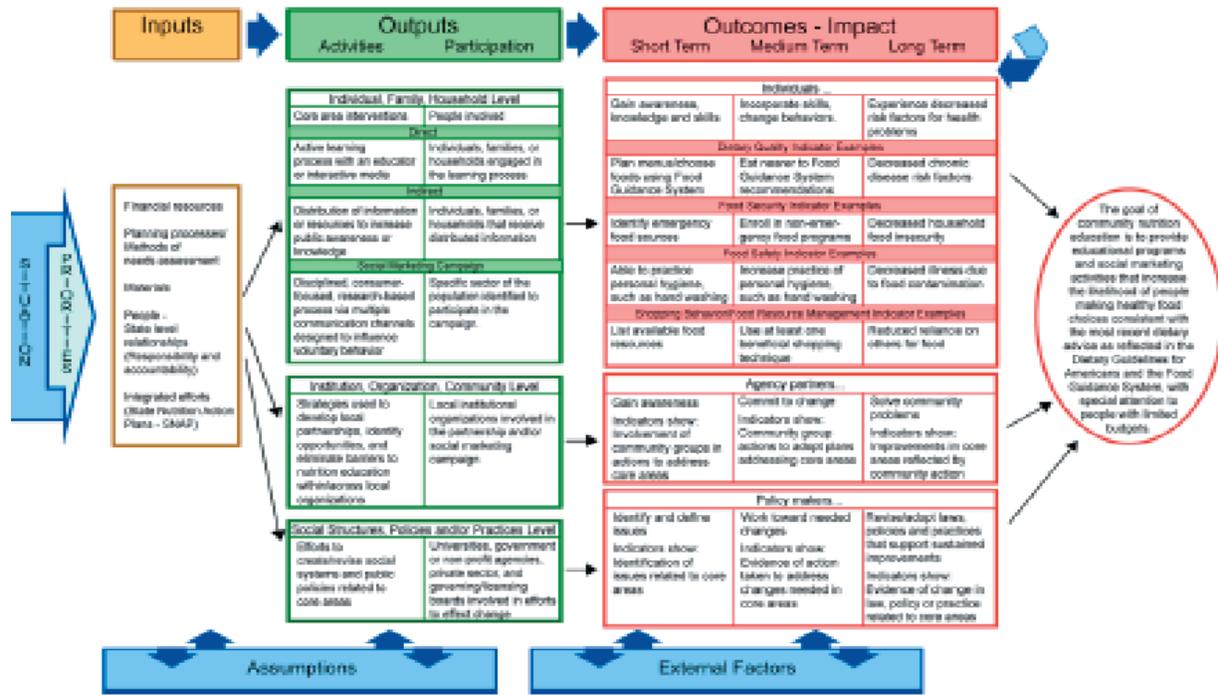
In reviewing the recent literature on programs and evaluation, I was intrigued with some of the terms that I found: program theory, theory of change, and theory-driven evaluation. This literature points out that program theory can provide a logic framework for planning data collection and can claim a reasonable approximation of a causal explanation. Stated more directly, in complex programs where it is very difficult to attribute cause, by developing a sound program theory one can achieve a reasonable explanation of causality. Program theory can help evaluators and others understand and tell why and how a program works, and be able to articulate that to others.

As I said, this is a very rapidly evolving field. In research published just this month [September 2005], I came across the following working definition, by Bledsoe and Graham: “Synthesis of both stakeholder program logic and social science theory to define what a program does, in what manner, and how much of an effect each goal and objective can have on the outcome.”

Logic Models

I am guessing most, if not all, of you have heard of logic models. Here, I am going to show two logic models. I do not have time to discuss them in detail, but will identify some key components of each which strengthen their usefulness and value to us. The first is the Community and Nutrition Education (CNE) Logic Model that I have had the good fortune to be involved with developing and testing over the past 5 years. The draft version you see here is where we are at with the second iteration. We are continuing to refine this model. Version 2 will be posted on the www.ces-fsne.org website in January 2006. You can also learn of the initial development of this model in the September 2005 issue of the Journal of Nutrition Education and Behavior.

The Community Nutrition Education (CNE) Logic Model – Overview



The Community Nutrition Education (CNE) Logic Model – REVISED DRAFT, March 2008; Helen Chapman and CNE Logic Model Workinggroup

For details on initial development see *J. Nutr. Educ. Behav.* 2005; 37:197-202

What I want to point out regarding this model is that it consciously incorporates social science theory. Embedded within the model is the socio-ecological model, reflecting a direct relationship between individuals, communities, and social structures/policies and the assumption that all are needed for sustained change to occur. The *CNE Logic Model* has also been used with the trans-theoretical model to explore outcomes based on participant messaging and stages of change, the end goal being to match messages to participants for realistic expected outcomes.

The power of logic models is in the arrows. They represent the chain of connections between investment and outcome. They define the program theory. Note that this logic model is multidirectional. When certain outcomes occur, it is important to check if and how those outcomes are connected to the audience involved, the types of activities conducted, and the investment made.

As Carol [Olander] pointed out with respect to evaluation in general, it is critical to clearly articulate the assumptions and external factors. For example, we know that in the past 5 years, the discussion around obesity has changed markedly in industry, the media, and entertainment. To assume that there is a certain pattern resulting from nutrition education independent of what is happening in our current social, economic, and political environment would not give us the robust theory that we desire. Thus, we see a place for assumptions and external factors clearly identified in the *CNE Logic Model*.

Some of you may be using other logic models, such as this one from the Centers for Disease Control and Prevention (CDC), created for the VERB™ Campaign. Here, too, there has been careful consideration of social science theory and program theory. Additionally, social marketing concepts have been applied, as they are used in many education and outreach programs.

Logic Models and Evaluation

In both of these examples, the flow of the model needs to be tested to determine if the models are reasonable and sound. To illustrate, the appropriateness of the outcomes listed in the *CNE Logic Model* needs to be confirmed through research and testing of primary and secondary data. Some concern has been expressed that the evaluation methods and measures (i.e. the tools) are not reflected in these models. Actually, in both examples given here, evaluation undergirds the entire model. At every level, it's important to determine if the different components of the model work using methods and measures that are reliable and valid. In essence, we are saying that we need evaluation for accountability and we need strong theory to support that evaluation. At the same time, we need to make sure our theory is sound and to test that theory to make sure that it leads to good evaluation and effective accountability.

What Can You Do?

The question that I was asked to address for this presentation was “what about you?” How can you (and I) include or strengthen our use of theory in our own programming efforts here and now? We need to start with where we are because good theory begins with local knowledge and wisdom. If we try to develop theory at a university, State agency, or in Washington, D.C., without taking into account the local application, our theory is going to have problems.

| What about You?

- **Start where you are. Theory comes from:**
 - **Local knowledge and wisdom**
 - **Research and evidence base**
 - **“Best” or “Promising Practices”**
 - **Evaluation studies**
 - **Other lessons from the field**
 - **General social science theories of change**
- <http://www.uwex.edu/ces/lmcourse/>

Additionally, good theory builds upon the research and evidence base. It takes into account best or promising practices. I want to stop on this point for just a moment. Sometimes, when we come to these conferences, we get ideas that excite us and want to go home and check them out immediately. I would hope that before implementing such practices, we pause and ask ourselves, “Why did I pick that particular best practice? What do I expect to accomplish? How will that particular practice fit within the context of what we (our organization or agency) are trying to achieve?”

We need to also think through why, how, and to what extent we are effective at what we do, and to review evaluation studies and lessons from the field in other disciplines to enhance our understanding. For example, today, I touched on some of the recent evaluation literature and how social science theories and evolving program theories might inform the way we look at evaluation of nutrition programs.

What else can we do? We can contribute to the development, refinement, and testing of what is in place. Carol [Olander] mentioned a piece that the Food and Nutrition Service (FNS) has done. We will hear from the Economic Research Service (ERS) about some of the work they are doing. We, in the Cooperative State Research, Education and Extension Service (CSREES), working with the Land-Grant University System, are furthering the development of the *CNE Logic Model*. We are going to beta-test this model with our Land-Grant institutions. Your assistance in asking questions, providing data, and contributing to any national effort will strengthen what we are doing, as we try to move towards sound program theory.

What about You?

- **Contribute to the development, refinement, and/or testing of what is in place**
 - **Use insights gained at professional meetings to strengthen your program; build theory into your program**
 - **Assist with testing (Example - CNE Logic Model beta testing)**
 - **Raise questions with intent to build knowledge**
 - **Work collaboratively with others; share knowledge gains**

Importantly, our goal is to work collaboratively in national and State efforts. That is one reason that we wanted to make sure that this would be an interactive session. With the focus and emphasis on needing evaluation, and needing it now, we can't wait. We must learn from each other.

Program Theory - Cautions

There are some cautions to be considered in talking about program theory. One is the tendency to be drawn immediately towards outcomes. We want the outcomes, we want the methods, we want the measures, and we want to put them into action. As a result, there can be a tendency to focus too much on outcomes and too little on actually testing the theory and making sure that our thinking is sound.

Also, there is the potential for perfecting the key to the wrong lock. We never get to the goal that is what a particular program is supposed to accomplish.

Or, we remain static. We might develop a wonderful model, but neglect to consider that programming is contextual; it is ever changing. It's essential to stay abreast of such changes if we are to develop sound program theory.

Acknowledgments

In closing, I want to mention that development of program theory and evaluation are not accomplished in isolation. For example, development of the *CNE Logic Model* includes two workgroups on a national level across the country that have contributed significantly to the richness of what we are learning. Wisdom from inter-agency and intra-agency colleagues who ask: “What about this, it doesn’t quite make sense to me” or “Are you missing that piece” has enhanced our own understanding. Administrative support has been absolutely critical, as have the insights shared by instructional technology specialists. Lastly, those of you who are willing to work with us to raise questions, help test, and provide data are playing an essential role in efforts to develop sound program theory. Whether you have worked with us specifically on the *CNE Logic Model*, have worked with other Federal agencies on other evaluation efforts, or are working within your States to incorporate sound theory and evaluation practices, you are contributing to the greater body of understanding to help predict, or explain in a reasonable and plausible way what we do. You are making a difference.

Thank you.

References

- ◆ Bledsoe KL, Graham JA. The Use of Multiple Evaluation Approaches in Program Evaluation, *American Journal of Evaluation*, September 2005;26:3, 302-319.
- ◆ Medeiros LC, Butkus SN, Chipman H, et al. A Logic Model Framework for Community Nutrition Education, *Journal of Nutrition Education and Behavior*, 2005;37:197-202.
- ◆ Philliber S. The Virtue and Specificity in Theory of Change Evaluation: Practitioner Reflections, The Aspen Institute, 1998. www.aspeninstitute.org. Accessed August 9, 2005.
- ◆ Taylor-Powell E. Logic Models: A Framework for Program Planning and Evaluation, *Nutrition, Food Safety and Health Conference*, Baltimore MD, March 31, 2005. [Presentation] www.uwex.edu/pdande/evaluation/evallogicmodel.html. Accessed September 5, 2005.
- ◆ Weiss CH. Where We’ve Been and Where We’re Going: Experts Reflect and Look Ahead – On Theory-based Evaluation: Winning Friends and Influencing People. *The Evaluation Exchange*, Winter 2003/2004; IX:4. www.gse.harvard.edu/hfrp/eval/issue24/theory.html. Accessed September 8, 2005.
- ◆ Community Nutrition Education (CNE) Logic Model. See www.ces-fsne.org for original model. Accessed September 5, 2005. Version 2 added to website, January 2006.

- ◆ Enhancing Program Performance with Logic Models; Web-based course. www.uwex.edu/ces/lmcourse/. Accessed September 5, 2005.
- ◆ VERB™ Campaign Logic Model. See www.cdc.gov/youthcampaign/research/.
- ◆ www.eval.org

So, You Need To Measure Behavior Change?

Eileen S. Stommes, PhD, Senior Sociologist, Food Assistance Branch, Food and Economics Division, Economic Research Service, U.S. Department of Agriculture, Washington, DC

Joanne Guthrie, PhD, MPH, RD, Assistant Deputy Director for Nutrition in the Food Assistance and Nutrition Research Program (FANRP), Economic Research Service, U.S. Department of Agriculture, Washington, DC

The views expressed in this article are those of the author and not necessarily those of ERS or USDA.

DR. STOMMES: Good afternoon. Thank you for being with us this afternoon. Many of the things I am going to say you have already been referred to thematically by both Carol and Helen. The title of my presentation is, “So You Need To Measure Behavior Change? Progress in the Food Stamp Nutrition Education Evaluation Project.” This project is being carried out within ERS, the Economic Research Service, but in collaboration with FNS, CSREES, ARS, and the Society for Nutrition Education.

This particular project is being led by the Food and Nutrition Assistance Research Program of ERS. The overall program was initiated in 1998 and includes both intramural and extramural research—the research project that I will discuss is a combination of both. It has been jointly supported with funds from FNS and ERS.

The focus of our current activities is the development of a specific measure of dietary behaviors targeted by FSNE and, before getting into details, I want to mention a couple of things. One is that any effort of this type requires resources, in terms of people and funding resources, and it requires much time and patience. The first step in our process towards developing a dietary behavior change measure began with publication of the *Journal of Nutrition Education Supplement* on “Evaluation of Nutrition Education with Low-Income Families” which was published in 2001. But, the project actually began much earlier than that and the details of its inception and progress are available on the ERS website.

The *Journal of Nutrition Education Supplement* represents the results of the first step in our process. It reviews what was known about evaluation of nutrition education with low-income families, according to the five priorities of the FSNE Program. Our collaborators—an

expert group of nutrition researchers and educators--went through the published literature and found the evaluation measures that had been done in each of these key areas and compiled them into five separate articles. So, this was a starting point in terms of what is the best of evaluation measures that we found in the literature in each of these five topic areas.

One of the things that came out of this publication, because it involved many people both in its development and in its publication, was the development of collaborative relationships. I would like to mention three of them that are very important. One is an interagency working group on Food Stamp Nutrition Education Evaluation. That includes FNS; CSREES, otherwise known as Cooperative Extension; ERS; and ARS. Following work on the Journal Supplement, ERS provided a grant to the Society for Nutrition Education for a postconference workshop in July 2003 to come together and try to identify the most important next steps for improving FSNE evaluation. I will get into that later. It has also involved a continuous collaboration during each of the workshops that I will describe to you briefly. All of our collaborative efforts have centered around developing better ways to answer the questions "What is the impact of nutrition education? How do you measure it?"

During the series of workshops, throughout this whole process, we came closer and closer to agreement. I won't say consensus, because there are still many differences. But, we began to talk more and more the same language of what we meant by evaluation. And, Carol, by the way, I really like those principles you discussed in your presentation, because I think if we had those to begin with, it would have been easier from the very beginning.

In January of 2003, we had a workshop which was a very small format, essentially a preplanning workshop for the larger July 2003 workshop to be held at the SNE Annual conference. We brought in representatives of the Interagency Task Force, Society for Nutrition Education, members of Federal and State agencies, and I think there are several people in the audience who were at that workshop. At this workshop, we were attempting to, and it sounded very general at the time, identify the research needs to support program evaluation. As a result of that workshop, we did several things.

We agreed that we needed to further define what we meant by outcomes. We needed to work on, or look at, developing information systems which would support both research and evaluation; develop meaningful process measures; develop strategies for improving collaborative efforts in research and evaluation; and then maintain ongoing dialogue.

The following July, we held the larger workshop at SNE. We focused on automated, web-based data collection systems, and how program reporting systems, or administrative reporting, can support evaluation. Then, we wanted to link process strategies and outcomes. It is okay to know if you have a good outcome, but if you don't know how you got there, then you really don't know much about that outcome. Though many needs were identified, the identified priorities led us, at least for our next steps, to narrow our focus.

We decided to put our efforts into developing a standardized set of outcome measures for nutrition education, again, with low-income audiences.

As a followup to that workshop, we commissioned a review of all of the available measures of dietary behavior. Some of you are going to think back to one of my first slides and say, wait a minute, didn't you already do that? Yes, but no. That review had failed to identify an appropriate questionnaire for our uses, but we thought that examining questions within questionnaires would be useful as a starting point for development of a new measure. We contracted with Mathematica Policy Research, Inc. to cull the literature on dietary behavior change and to give us the questions that had actually been used in surveys that were reported in the literature. They gave us all the questions that they deemed to pass muster according to a set of criteria that included considerations of the evidence on their reliability, validity, and appropriateness to the FSNE audience. We had 500 rejected questions as a result of that and 100 questions that passed muster.

We then went to another workshop in April of 2004, and we had a very small group of people. I think some people are here in the audience, who looked at what we call the "prototype notebook." They identified candidate questions for further testing. In other words, they said, "We think this is a question about dietary behavior that should be followed up on and should be tested for possible inclusion in our outcome measure." Attendees also suggested additional questions to cover gaps in existing questions, or to cover topic areas not addressed. Weight management, or healthy weight, was one of those issues. I believe the whole-grains area was another one that people felt the literature didn't represent.

Everyone agreed that the questions needed refinement because they had come from a variety of instruments and were often not developed with the low-income audience in mind. We asked them to update it up through 2003. So, this is the best the literature had to offer us. There were a variety of instruments. There were different reference periods—some were asking about weeks. Do you do this once a year, once a month, twice daily? There were different response categories and, importantly, they were all developed before the 2005 *Dietary Guidelines*.

A group of volunteers reviewed the questions after the workshop and agreed that we needed to move towards cognitive testing of the questions if we were ever to develop a measure of dietary behavior change. After several discussions, ERS developed a contract with Abt Associates, a consulting firm, to revise the questions for cognitive testing, and to develop a protocol with standardized probes, training materials, and a video—a video on how to do cognitive testing, and another video on how not to do cognitive testing.

Now, what is the next step? We have the protocol, we have the training materials, but we don't have the questions yet. We have a set of about 75 questions that need further work. We have just signed a contract with Mathematica Policy Research, Inc, and FNS is part of that effort, providing review and financial support. The key activities of this will be to do a content review of the draft and bring it up-to-date with the 2005 *Dietary Guidelines*,

removing outdated questions and adding new questions. Then, the remaining questions will be cognitively tested. The third stage is to do field testing of these questions to develop a short dietary behavior measure. It will, essentially, consist of a very short set of questions that will contain core questions on important dietary outcomes targeted by FSNE, and then modules that will address selected topics in more detail, such as fruit and vegetable consumption, dairy consumption, and whole grains.

We want a questionnaire that will be capable of being administered by telephone, paper and pencil, and by electronic means. But, the test is whether or not it can be administered by telephone, because that is the most difficult standard to pass. The question length should be no more than 15 minutes. The long-term plan for this dietary measure is to validate this very short measure against established measures of dietary quality. We are working with FNS to see if we can incorporate the set of questions that we are developing into a future survey that includes Food Stamp Program participants. We are still working on the coordination aspect.

Validation possibilities include testing it against the Healthy Eating Index, based on food consumption data collected by repeated 24-hour recalls with the appropriate audience or against some other criterion. Then, the final measure will be suitable for use by national, State, regional, and local nutrition educators in evaluating the effect of their program.

For continuing information, ERS has a FSNE website. What you need to do is go to the ERS website and go to the Food Stamp page and scroll down a little bit. All of the information has been published thus far on FNS evaluations is available there. We also have a series of maps that indicate FNS funding for States and FNS funding per food stamp recipient. You may want to go there, we don't have the 2005 data up yet, but we will shortly. We also will feature links to ERS data and other relevant data sources. We link with Helen's site that she mentioned.

Another followup is that we have a set of three articles on evaluation and our plans to develop a measure of dietary behaviors targeted by which are forthcoming in the Journal of Nutrition Education and Behavior. I hope you will read those articles when they are published in winter of 2006.

QUESTION FROM AUDIENCE: I have a question for USDA. Are there efforts going into evaluating sales data from food stamp redemption in terms of outcomes for our program? We can do individualized diet assessment until we are blue in the face, but it is really how those benefits are being used at the grocery store that really is, I think, the gold standard of whether we have succeeded or not. So I would be interested in any research initiatives from USDA in that area.

MS. GUTHRIE: I am a co-author on the presentation with Eileen. That is a very good point. I just talked about how great I thought it was that Brenda was using non-self-report measures, so it is a logical question. We are interested in using more sales data.

At the moment, our current source of sales data is data that we have purchased from A.C. Nielsen which, as you know, is a major market research firm and they have national data on food purchases in America. We have used that for some analyses that look at purchasing behaviors of low-income and other kinds of consumers. But, they don't identify that people are food stamp recipients. So they can't, at this moment, talk about the purchasing behaviors of food stamp recipients.

If we have funding through our current data development initiative, A.C. Nielsen might add questions on Food Stamp Program participation if we made it worth their while, just as they might for any other client. That would certainly be something to consider; particularly, if we were to have the available funding. That, of course, would not be linked in any direct way with any kind of nutrition education. We talked at other times of doing things that are more specific evaluations, for example we have funded one with WIC where we have looked at redemption of low-fat milks by WIC participants using WIC vouchers after implementation of an education program focused on drinking low-fat milk.

Translating Research Into Practice

Brenda L. Dobson, MS, RD, Nutrition Services Coordinator, Iowa Department of Public Health, Des Moines, IA

Good afternoon. It is a pleasure to be here this afternoon to share some information from our WIC Special Project Grant. My objective this afternoon is to share information about our evaluation design. I would love to share the results, but that wasn't the mission I was given. I will refer you, however, to the abstracts for the conference where you will see a few selected highlights of the results of the project.

I will start by describing our nutrition education intervention, which is grounded in two different theoretical models – the trans-theoretical model and critical thinking skills model. We staged parents regarding their willingness to offer vegetables to young children. A lot of work with the trans-theoretical model has focused on an individual's own intent to change. What is different about this project is that we focused on the parent's willingness to offer vegetables to young children.

The staging tool that we used was developed by the NC 219 Regional Project, which involves community nutrition faculty from a number of land grant universities, including Iowa State University, our key partner in this particular project. We also identified barriers that parents expressed that they had in terms of offering vegetables to their young children, and then we provided messages targeted to the stage and the barrier, or barriers, using activities to foster critical thinking skills.

The formative research that supported this project was the focus group work that supports the Pick A Better Snack project, the social marketing project developed by the Iowa Nutrition Network several years ago. It also builds on some formative research conducted in other States on similar topics.

I will share with you some elements of the research design. In terms of selecting the local agencies, we used a stratified selection process to identify a sample of WIC agencies in the State. The selection was multitiered, with selection first occurring at the agency level. After stratifying by primary language used in the home (which is information that we collect in the State WIC data system) and the presence of EFNEP in the WIC agency's service area, we then identified one intervention and one control agency using a random numbers table.

We have 20 local agencies in Iowa, but our starting point had to be 19 in this process because two agencies share the services of one dietitian. So, we had to take that into account when we did this random selection. Randomization wasn't possible at the clinic level because the dietitians work in a number of clinic sites each month. With only 20 agencies covering 99 counties, you can see that many of them are multicounty agencies. We randomly selected six clinics per agency. Some agencies had fewer clinics than six because of the nature of their service delivery schedule and area, but that was our starting point. This gave us a representative sampling of our WIC participation and caseload across the State.

In selecting project participants, we determined what the sample size was based on data from a Missouri project called, "The Impact of WIC in Farmer's Market Nutrition Programs on Participant Dietary Intake of Fresh Fruits and Vegetables." We used the most recent data that was available when we wrote the proposal (the 1996 BRFSS [Behavioral Risk Factor Surveillance System] data), ending up with a desired sample size of 534 parents of children ages 2 to 5, and with the same sample size for controls.

The next step was to determine how many parents needed to be enrolled at each of the clinic sites. We used a 6-month enrollment cycle in this project, so we ended up with a clinic goal of 22 enrolled every 6 months.

Staff at the local agencies were trained to use a random numbers table to identify the first parent to be asked to participate on a given day. For example, our first enrollment cycle started on June 15th. The starting point was the clinic's schedule for the day—that was the sampling form. The random numbers table was then used to identify which child would be invited first to participate in the study. Every subsequent child, between the ages of 2 and 5 after that, whether it was a walk-in or an appointment, was then invited to participate in the survey. That is how we accounted for the walk-in phenomenon that all of you who work in WIC know that we deal with.

This was a systematic sampling, but the starting base was the certification appointment schedule. We only enrolled one child per family, because doing otherwise would muddy the

data in terms of parental knowledge, and attitudes, and practices. This ended up having an interesting effect. We didn't realize that many of our local agencies had informal scheduling policies where they listed the oldest child first on their appointment schedule. We had more children "age out" of the sample than we had anticipated at the beginning; but we also ended up with a much larger sample size.

We collected a variety of enrollment data. We wanted to know, based on that appointment schedule, did the people come, were they invited, did they no-show, and how many declined to participate? We had some who were ineligible—a few who were ineligible for WIC services, but also a few who were ineligible because of major language barriers. The project was developed using English and Spanish materials, but only those two languages. We also had a few children who were not eating solid foods and were being maintained by two feedings, therefore, this intervention was clearly not appropriate for them.

For outcome evaluation, we focused on measurements that were done at baseline (or enrollment) and at each certification contact. We looked at parents' knowledge, attitude, and practices. We also looked at children's intake of vegetables; both their frequency of intake and the variety of vegetables that they consumed. We anticipated that some children would age-out of the population during this intervention, so we expected only a small percentage of children to continue to be seen for the full duration of the intervention. Therefore, we used multiple cross-sectional studies at those 6-month intervals, and compared those results to baseline. That gave us information about how WIC clients as a population changed.

The regression analysis that we completed includes data from children, actually, parents, who had at least two certification contacts. We have a wealth of data, some of which we haven't had a chance to use yet because it extends beyond the general purpose of this project. We have the typical things you would expect to find in the WIC electronic data system.

We needed some additional data, so we developed a single-page questionnaire. This additional questionnaire, with eight questions, was for collecting information on the relationship of the parent or caregiver who brought the child for services, the caregiver's age, the place of the caregiver's birth, and whether or not the caregiver was a current WIC participant. If they were also receiving nutrition education, we needed to consider that and whether or not they were participating in an EFNEP program.

We also collected information about the parent's knowledge about vegetables and asked: How many servings a day should young children have? How often do you need to offer new foods to children before deciding a young child doesn't like them? We had seven questions related to knowledge and seven questions related to attitude. Then we calculated scores for the responses to these questions. We used those scores in the data analysis. Each of the responses had a numeric value and then we calculated the scores.

We also looked for associations between barriers and the parents' stage and between barriers and children's consumption. Among those who were already consuming vegetables at least three times a day—we had some children whose parents reported they were consuming vegetables at least three times per day at baseline—we also looked for factors to predict common barriers.

We had several data elements related to parents' practices that we also evaluated. The staging algorithm collected the usual number of times per day that vegetables were offered to the young child and how long they had been offering vegetables. Knowing how long they had been consuming three or more servings per day helped distinguish whether they were in the action or the maintenance stage. If the answer to the initial question was "less than three," then what was their intention to change within 30 days, or within 6 months? That determined their stage of change on that continuum.

We also developed a brief household inventory related to vegetables that were in the home on the day that they came for WIC services. We compared this information to their food intake measures in terms of variety consumed in the past week and specific vegetables consumed in the past week. This has been very interesting information to review.

One of our hypotheses was that we would see Farmers' Market Nutrition Program check redemption rates increase in our intervention counties. We matched check registers with participant IDs to identify checks that were issued to project participants. Then we compared claimed checks to redeemed checks to determine redemption rates.

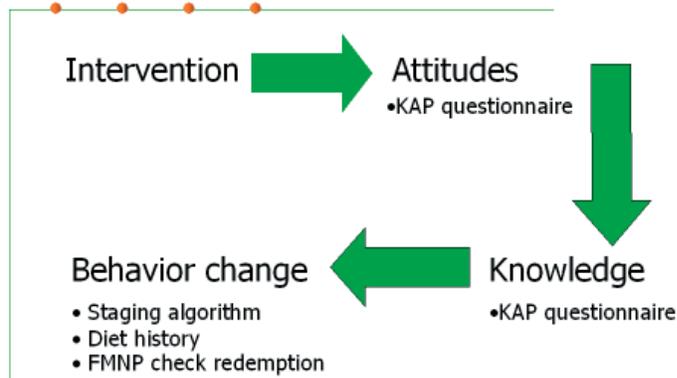
We also looked at children's intake. We were using a 7-day food frequency during the project, so we could look at the number of times vegetables were offered per day. Then, we looked at the number of different vegetables consumed per week. We also have other kinds of diet and history information, such as number of meals per day, snacks per day, how many meals a week they eat away from home, and similar information that is helpful to explain what we saw with the project results.

The Knowledge, Attitudes, and Practices (KAP) questionnaire collected a lot of the data for the knowledge and attitude scores. Related to children's intake, it asked about the frequency of offering vegetable snacks. That data was important as well. We completed additional analyses including an analysis of WIC staff attitudes, using pre- and post-KAP questionnaires. We did onsite clinic observations to determine and make sure that the intervention was delivered in the way it was intended to be delivered.

The intervention was structured to have a strong collaboration with the farmers' market and the EFNEP Program. We also collected information about that and did an extensive cost analysis related to program development, the research component, and program operations. It was challenging to separate those costs.

We measured attitudes and looked at knowledge and behavior change. It was very important to note that we measured attitudes and knowledge. If we had only measured behavior change, we would have been disappointed in our results.

Evaluating Nutrition Education



We saw significant changes in stage and shifts in attitude and knowledge, all of which are precursors to behavior change. This was an important lesson for me, as a program person, to be sure to look at the different parts of the behavior change process and cycle, and make sure that you evaluate in each of those areas.

The nutrition education materials are available on the web as well as tips for community collaboration. The final project report is also posted at www.idph.state.ia.us/hpcdp/vg_home.asp.

In closing, I want to recognize and thank the project team members from Iowa State and from the EFNEP Program. Without their collaboration, we could never have done this kind of evaluation from the State Health Department. It was key to the success of this project. Thank you.