FNS USDA
Food & Nutrition Service

NUTRITION EDUCATION: PRINCIPLES OF SOUND IMPACT EVALUATION

Office of Analysis, Nutrition and Evaluation **September 2005**

Background

The Food and Nutrition Service (FNS) is frequently asked, by a variety of nutrition education partners, how it defines a sound impact evaluation. The principles introduced here describe the characteristics of strong impact assessments of nutrition education. They are also consistent with the Government and Performance Results Act and the Office of Management and Budget's guidance for clear demonstration of program effects.

The principles are neither unique to evaluating nutrition education nor to assessing nutrition education in a particular FNS program. Instead, they reflect a set of standards that are generally regarded as pre-requisite to drawing credible conclusions about the impact of many types of educational, economic and social initiatives.

The information provided here first distinguishes impact assessments from other kinds of evaluation and then introduces the characteristics of strong impact studies. This document is not a hands-on guide for designing impact evaluations and may include a few unfamiliar terms. For those interested in operational tools for conducting impact evaluations, an annotated reference list is provided.

Further, the statement of principles is not intended to establish policy directing nutrition educators to conduct impact evaluations. The principles do, however, suggest when an impact evaluation should be considered and what research features are optimal for learning which interventions improve eating habits and support healthy lifestyles.

FNS is committed to helping to build a sound body of nutrition education information. We offer these principles as a tool for educators, researchers, and policy makers to support our joint efforts toward that objective.

Impact Assessments and Other Evaluation Types

Evaluation research takes a variety of forms that serve different purposes and address different questions. Because time, dollar, and technical resources are often limited, it makes sense to consider what is most important to learn and how results will be used.

Key types of evaluation include formative research, process or implementation studies, outcome assessments and impact evaluation. *Formative research* typically occurs up front, as an intervention is being developed. It may determine if the target audience understands the nutrition messages or test the feasibility of implementing a previously developed intervention in a new setting. Formative research results are used to shape the features of the intervention itself prior to implementation.

In contrast, *process evaluation* involves tracking actual implementation. Examples include documenting the number and characteristics of people who hear a radio spot or the amount of time taken to deliver a package of nutrition education classes. The information collected answers the question of whether or not an intervention is delivered as intended and may also identify implementation barriers and strategies for overcoming them.

Outcome assessments address the question of whether or not anticipated group changes or differences occur in conjunction with an intervention. Measuring shifts in a target group's nutrition knowledge before and after an intervention is an example of outcome evaluation. Such research indicates the degree to which the intended outcomes occur among the target population. It does not provide definitive

evidence, however, that the observed outcomes are due to the intervention.

Surveillance studies monitor the status of a population or subgroup on outcomes of interest, like attitudes toward healthy eating or dietary intake. Data from this research document how well or poorly the population is doing. As such, it may indicate whether or not action is needed. Surveillance studies do not explain, however, what causes the observed status or what will make a difference.

Only *impact evaluations*, allow one to conclude authoritatively, whether or not the observed outcomes are a result of the intervention. In order to draw cause and effect conclusions, impact evaluations incorporate research methods that eliminate alternative explanations. When it is important to know whether or how much difference an intervention makes, then an impact evaluation is indicated.

Evaluation planning begins with identifying who the audience is, what questions are essential to answer, and how results will be shared. The importance of particular questions varies over time and across different audiences. While the field of nutrition education can benefit from each sound impact evaluation, available expertise, limited time and resources or other more critical information needs may preclude an impact evaluation. If, however, cause and effect conclusions are critical to intervention planning or funding decisions, serious consideration of an impact evaluation is important.

Impact Evaluation Principles

To facilitate sound impact evaluations, the principles are organized chronologically. Some demand early consideration; others may be addressed appropriately after the intervention is implemented.

First Things First

1. Make certain that the nutrition education intervention can be evaluated.

Answers to each of the following questions are pre-requisite to any impact evaluation:

- *What are the specific objectives of the intervention?* For example, is the intervention expected to increase awareness of the link between diet and certain diseases, develop skill at understanding food labels, improve menus in school cafeterias, and reduce consumption of high fat snacks or something else?

- *How large of an impact can be expected reasonably?* Is the intervention sufficiently intense to bring about measurable changes in behavior? Is the nutrition knowledge or eating habits of the target audience at a level that leaves room for measurable improvement?

Clear and complete answers guide impact evaluation plans. Defining nutrition education objectives and considering the possible magnitude of the impact results in focusing the evaluation on the appropriate outcomes, measures and sample sizes.

- *Why and how is the intervention expected to achieve its objectives?* For example, are facts expected to change individual motivation which then influences behavior? Do multiple channels of communication affect the probability of attitude shifts? Is personal goal setting considered pre-requisite to dietary improvement?

- *Will the intervention be implemented as intended?* For example, are all components of the intervention delivered in the prescribed order, intensity and setting? Do nutrition educators receive the planned type and degree of training?

Understanding the underlying theory and dynamics of how an intervention is expected to work should help to identify key implementation features. That information can, in turn, be used to assess whether or not the intervention is being carried out as planned. Specifying and measuring the sequence of education impacts

makes it possible to capture both proximal and distal results. Learning that an intervention is not implemented as intended, may allow an early decision to postpone a costly impact evaluation and support a change in the nutrition education features. In general, detailed data on the intervention contributes to a better understanding of the results.

2. Build on available research.

Become familiar with what is already known and act on it -- not only when planning an educational initiative, but also when making decisions of whether and how to evaluate its impact. Check the record -- your own prior work, input from colleagues, publications and research data bases -- to see if key elements of the proposed intervention have been tested before. Use the available information to make modifications that are likely to contribute to successful implementation and positive impacts.

Assess the presence of any serious flaws in the prior research which limit the conclusions that can be drawn. Examples include, but aren't limited to, weak comparison designs, small or otherwise limited samples, inappropriate statistical analyses, measures without documented consistency and meaningfulness, and/or conclusions that are not supported by the methods used. Since few people are experts on every technical dimension, rely on others to share their knowledge – both for assessing prior research and your own evaluation plans.

Get the maximum return on your investment in impact evaluation by making choices that add to the body of knowledge. If the success of an intervention is well established and credible, it may not be the best candidate for additional assessment. Weigh carefully the information return on the resources required.

Alternatively, an intervention with a strong performance record could be a good evaluation vehicle for learning what specific conditions or features contribute to success. This may involve examining the relative importance of different intervention components, the duration of observed impacts or the variability of impacts across different populations.

3. Hold out for research designs with random assignment but use them selectively.

Do you want to conclude, authoritatively, that a nutrition intervention does or does not have an impact? If yes, it is essential to (1) randomly assign members of the target audience to groups who get educational services or do not and then (2) compare the groups. Anything short of this design rigor leaves room for alternative interpretations.

While technically superior for establishing cause and effect, experimental research in real world settings may be difficult or prohibited. For example, the WIC Program requires a specified nutrition education package to be an integral part of the benefit. Setting up a control group, in which educational services are withheld from WIC participants is not an option.

Even when one can legally and ethically withhold nutrition education, experimental design should be used selectively and strategically. Invest when it can add to the body of knowledge -- for example, when an intervention is new, has changed substantially or is being delivered to a new audience. Invest in experimental designs when the other elements of evaluation are sufficiently sound -- e.g., there are adequate sample sizes, replicable research settings, along with valid and reliable measures.

Finally, be clear about the question that needs to be addressed. Sometimes the key objective is to get an overall picture or monitor trends. If so, a surveillance study may be more appropriate than an experimental design. However, if the goal is to identify nutrition education that makes a difference for those served, the most definitive conclusions about what works for whom come from random assignment to treatment and control groups.

4. Choose impact measures that fit the intervention and approach existing standards for credible assessment.

Measures should be compatible with an intervention's purpose and its likely magnitude of change. If an intervention is about promoting healthful food choices, the appropriate measurement domain is behavior. If the intervention emphasizes fruit and vegetable consumption, pass up measures that cover a broad range of behaviors. If the intervention is relatively short in duration or intensity choose a measure that is sensitive enough to pick up small behavior changes – like trying more vegetables in a subsequent tasting event.

There are standards for assessing a measure's reliability, that is, its capacity to produce consistent responses. There are also benchmarks for judging a measure's validity, i.e., its ability to capture what it intends to measure.

A reasonable starting place is to review the pool of available instruments. The first screen includes compatibility with the intervention, research questions and target population. Second, consider a measure's technical credentials for reliability and validity. Then, examine prior variability in a measure's scores to determine how many persons or other units are needed to observe impacts that are statistically significant.

Because sound behavior measures in the nutrition education field are limited, some effort to develop or modify an existing instrument is likely. At a minimum, establish the internal consistency of items in a new or revised measure and obtain expert judgment about the appropriateness of item content. Using multiple instruments to assess an impact also adds weight to the measurement process and results. Careful documentation of the steps taken to develop and test a behavioral instrument will contribute to the inventory of stronger tools.

5. Observe standards for the fair treatment of study participants.

Many agencies and universities have formal requirements and clearance procedures to ensure that the rights of study participants are respected. While the specifics may vary somewhat across organizations, standards typically seek to:

- ensure informed consent,
- guarantee the confidentiality of participant data,
- limit the burden associated with study participation, and
- make sure that no one is denied essential services.

Planning an impact evaluation begins with identifying and then observing relevant ethical standards.

Although the food and food resource benefits from FNS programs may not be withheld for the purpose of creating research control groups, there is more flexibility with respect to the nutrition education component of the Food Stamp Program. Currently, food stamp nutrition education is a State option. While all States currently provide educational services, they are not provided to all program participants. As FSNE initiatives are being developed and assessed, it is appropriate to incorporate a non-treatment or no education condition for the purpose of evaluating impacts.

For other Programs, like WIC, a specified nutrition education package is an integral part of the benefit. Withholding educational services is not an option. Here, the experimental model requires random assignment of individuals, clinics, or other units to different kinds of educational interventions.

As the Intervention Begins

6. Collect impact data after start-up problems get resolved but before implementation rolls out.

The goal is to avoid testing nutrition education effects too soon. Many new initiatives go through a transition before reaching a steady-state level of operations. Before measuring impacts, check that nutrition education services are being delivered as intended and that the initial bugs have been eliminated.

At the same time, don't wait so long that the intervention is widely delivered. Once broad implementation takes place, it may be difficult to ensure that control individuals or groups have not been influenced by some components of the intervention being evaluated.

Timing matters if the goal is to capture the full impact of a nutrition education initiative.

Post-Evaluation

7. Report both positive and negative results – but do so accurately.

Adding to the body of scientific knowledge demands not only strong research design, but also thorough analysis and thoughtful reporting. While experimental studies support probability statements about an intervention's impact, the findings will be more or less applicable to other settings and populations.

In order to generalize results with confidence and replicate interventions with success, consider the specifics that make a difference. First, are the features of a model intervention sufficiently detailed so that they can be replicated? Even when an intervention is carefully described, ask whether or not its components can be implemented routinely. For example, what are the implications of a successful intervention in which providers are highly educated professionals and participants are obligated to attend all lessons?

There are a number of other indicators to check when interpreting and reporting results. Among the questions one should address: Did persons assigned to the control group inadvertently receive any of the nutrition education services intended for the intervention group? Does participant attrition occur differentially across intervention and control groups? Do the predicted intermediary impacts occur in the expected order and magnitude? Are the observed effects similar to findings reported elsewhere?

Attention to alternative explanations and existing trends is part of sound evaluation whether the results are positive or negative. This means examining the plausibility of impacts in terms of intervention features, relevant theory, existing research and the rigor of the research design and its execution. When analysis and reporting tackle such questions, the usefulness of a study increases measurably.

8. Share results to maximize their value.

The purpose of sound evaluation is to inform future decisions and choices. So, sharing evaluation results with stakeholders -- nutrition educators, other researchers and policy makers -- should be built into the research planning process. Even when the intervention results are modest or non-existent, the results of a well-executed impact evaluation are informative. They may help identify more or less desirable policy alternatives, provide a different perspective on an issue, and contribute to the body of knowledge that will have an impact on future decisions.

Information should be clear and tailored for what each audience wants to know. Attention should be given to reporting the research, results and what findings mean. A full research report that captures the following key components should be available:

- the theoretical context or explanation for how an intervention will change behavior;
- a clear picture of the intervention that identifies messages, channels of delivery, including the expertise of delivery agents; target audience; intensity and duration;

- a description of the evaluation features that covers research design, selection of study participants, impact measures and analysis;

- an examination of results that addresses their plausibility and generalizability; and

- lessons learned and recommendations.

Finally, reporting requires a plan and follow-through to get information to the intended audiences. Publication in peer-reviewed journals is an important mechanism for sharing research but other channels should be considered both to reach specific audiences and make results more generally available. These channels may include posting reports or abstracts on organizational web-sites, announcements to relevant list serves, and presentations at variety of meetings, teleconferences or newsletter items.

PRINCIPLES OF IMPACT EVALUATION

1. **Make certain that the nutrition education intervention can be evaluated.**
2. **Build on available research.**
3. **Hold out for research designs with random assignment but use them selectively.**
4. **Choose impact measures that fit the intervention and approach existing standards for credible assessment.**
5. **Observe standards for the fair treatment of study participants.**
6. **Collect impact data after start-up problems get resolved but before implementation rolls out.**
7. **Report both positive and negative results – but do so accurately.**
8. **Share results to maximize their value.**

Links to Technical Resources for Impact Evaluations of Nutrition Education

Federal Emphasis on Random Assignment Studies

- Department of Education. *Scientifically Based Evaluation Methods; Notice.* Washington, D.C., Federal Register, pp. 3586-3589; January 25, 2005.

  Public Notice in which the Secretary of Education (1) acknowledges that random assignment is not the only design capable of providing estimates of program effectiveness, but (2) concludes that it is most defensible method in that it reliably produces an unbiased estimate of effectiveness.

- Office of Management and the Budget. *What Constitutes Strong Evidence of a Program's Effectiveness?* May 2004. www.whitehouse.gov/omb/part/2004_program_eval.pdf

  Guidance for the Program Assessment Rating Tool (PART) that endorses randomized controlled trials as the preferred method for measuring program effectiveness and well-matched quasi-experimental studies as a possible alternative when randomized designs are not feasible.

- Senate Appropriations Report. *Departments of Labor, Health and Human Services, Education and Related Agencies Appropriation Bill, 2006.* Congressional Record, Report 109-103, pp. 284-289, July 14, 2005.

  Report to accompany House of Representative bill 3010 that notes the importance of scientifically based research to increase the effectiveness of education interventions and calls for greater focus on the use of randomized trials.

Databases for Searching Nutrition Education Research

- AGRICOLA
- Ageline
- CDC's Preventing Chronic Disease
- Combined Health Information Database (CHID)
- Elton B. Stevens Company (EBSCO) Sociological Collection
- Educational Resources Information Center (ERIC)
- Federal Research in Progress (FEDRIP)
- Institute for Scientific Information (ISI Web)
- Ovid Healthstar
- PubMed

General Evaluation Guides

- American Association for the Advancement of Science. *Taking Stock: A Practical Guide to Evaluating Your Own Programs.* Washington, DC: 1997. www.horizon-research.com/publications/stock

  Very basic guide that is targeted to community-based organizations planning their own program evaluations. Research emphases are on defining intervention goals and objectives, using different

data sources, and telling the evaluation story. Includes glossary of evaluation terms and example of evaluation report. *Does not* address research design or analysis issues.

- Canadian Health Services Research Foundation. *Knowledge Transfer and Exchange.* Ottawa, Canada. www.chsfr.ca

  Material on effective dissemination of research.

- Hersey, James and S. Daugherty. *Evaluating Social Marketing in Nutrition: A Resource Manual.* Report prepared by Research Triangle Institute and Health Systems Research, Inc. Alexandria, VA: Food and Nutrition Service, USDA, 1999.

  General guide to formative, process and impact evaluations of social marketing campaigns with attention to the special challenges of assessing such interventions.

- National Science Foundation. *User-friendly Handbook for Project Evaluation.* Washington, D.C.: Directorate for Education and Human Resources; January 2002. www.nsf.gov/pubs/2002/nsf02057nsf02057.pdf

  Very basic guide that provides special emphasis on conceptual description of interventions or logic models, alternation types of data collection, report preparation, and culturally responsive evaluation strategies. Includes a glossary of evaluation terms. *Modest treatment* of research design and analysis issues.

- Treasury Board of Canada, Secretariat. *Program Evaluation Methods: Measurement and Attribution of Program Results.* Ottawa, Canada: Pubic Works and Government Services; 1998. http://www.tbs-sct.gc.ca/eval/pubs/meth/pem-mep_e.pdf

  Book-length guide that focuses on design alternatives, data collection methods and analysis approaches.

- U.S. Department of Agriculture. *WIC Evaluation Resource Guide.* Alexandria, VA: Food and Nutrition Services; September 1991. www.fns.usda.gov/oane/MENU/published/WIC/FILES/WICEvaluationResourceGuide

  Brief guide that addresses why and when to evaluate; links between evaluation questions and evaluation design; types of data collection instruments, sampling choices, and basic descriptive data analysis procedures.

- U.S. Department of Agriculture. *The Community Nutrition Education (CNE) Logic Model.* Washington, DC: Cooperative State Research Education and Extension Service, December 2002. http://www.ces-fsne.org/ Click on logic model.

  Detailed model for conceptualizing potential links between nutrition education inputs, outputs and outcomes or impacts.

- U.S. Department of Education. *Identifying and Implementing Education Practices Supported by Rigorous Evidences: A User Friendly Guide.* Washington, DC: Institute of Education Sciences, December 2003. www.ed.gov.rschstat/research/pubs/rigorousevid/rigorousevid.pdf

Evaluation users' guide that defines randomized controlled trials and their relative advantage for producing strong evidence on intervention effectiveness. Features that are pre-requisite to well-designed and executed randomized design trials. Discussion of circumstances in which well-matched comparison group studies can provide possible evidence of effectiveness.

- U.S. Department of Health and Human Services. *Framework for Program Evaluatio.n* Atlanta, GA: Centers for Disease Control and Prevention; September 1999. www.cdc.gov/mmwr/PDF/RR/RR4811.pdf

  Very general overview of steps to conducting program evaluation – from engaging stakeholders to sharing lessons learned. Discussion of criteria for judging usefulness of evaluation.

- U.S. Department of Health and Human Services. *Physical Activity Evaluation Handbook.* Atlanta, GA: Centers for Disease Control and Prevention, 2002. www.cdc.gov/nccdphp/dnpa/physical/handbook/pdf/handbook.pdf

  General guidance that applies evaluation framework (see above) to interventions to promote physical activity. Includes worksheets for applying each evaluation step and special sections on writing objectives, theory applications, and measurement resources. *Modest attention to design and analysis.*

- U.S. Department of Health and Human Services. *Program Managers Guide to Evaluation.* Washington, DC: Administration on Children, Youth and Families. www.acf.dhhs.gov/programs/opre/other_resrch/pm_guide_eval/reports/pmguide/pmguidetoc.html

  Evaluation guide that addresses both technical and management issues. Topics cover justification for evaluations; steps and tips for hiring an outside evaluator; preparation of an evaluation plan; data collection sources and tools; links between the evaluation questions and data analysis, and results reporting. Includes a glossary of evaluation terms, multiple checklists and worksheets.

- U.S. Department of Health and Human Services. *Guidance for Comprehensive Cancer Control Planning. Part II, Section 8 – Conduct Evaluation.* Atlanta, GA: Centers for Disease Control and Prevention. http://cdc.gov/cancer/ncccp/guidelines/part2/section8.htm

  Brief evaluation guide that addresses technical and administrative issues. Specific attention effectively using evaluation to support intervention planning.

- U.S. General Accounting Office. *Designing Evaluations.* Washington DC: U.S. General Accounting Office, Program Evaluation and Methodology Division; 1991. www.gao.gov then search for report PEMD-10.1.4

  Evaluation guide that focuses on evaluation questions and designs. Examines different designs associated with descriptive, normative and impact evaluations. Includes a glossary of evaluation terms.

- University of Wisconsin. *Enhancing Program Performance with Logic Models – On-line Course.* Madison, Wisconsin: Extension Service. 2002. http://www.uwex.edu/ces/lmcourse

  Web-base course on logic model basics and community nutrition education logic models.

Measures of Nutrition Education Impacts and Outcomes

- Contento, I.R., J.S. Randell, and C.E. Basch. *Review and Analysis of Evaluation Measures Used Nutrition Education Intervention Research.* Journal of Nutrition Education and Behavior, V. 34 (1), January-February, 2002; pp.2-25.

  Review of nutrition education evaluation measures used over a twenty year period. Measures categorized by population subgroups and focus (e.g., dietary intakes, other behaviors, physiological measures). Summary describes adequacy of validity and reliability associated with different measures.

- Hartline-Grafton, H., R.Nyman, R. Briefel, R. Cohen. Mathematica Policy Research. *Prototype Notebook: Short Questions on Dietary Intake, Knowledge, Attitudes, and Behaviors.* Washington, D.C., USDA/Economic Research Service, September 2004. http://www.ers.usda.gov/publications/efan04010/

  Compendium of survey questions used in previous research to assess dietary knowledge, attitudes, and behaviors among low-income adults. Information on item reliability, validity sensitivity to change and other dimensions is included.

- Hersey, J.C. (Guest Editor). *Evaluation of Nutrition Education with Low-Income Families.* Journal of Nutrition Education, Volume 33 (Supplement 1), September, 2001. http://www.ers.usda.gov/Briefing/Food Stamps/readings.htm

  Review of tools to measure nutrition education impacts in five general areas: environment, food resource management, food safety, diet quality, and food security.

- Logan, C., M.K. Fox, and B-H. Lin. *Effects of Food Assistance and Nutrition Programs on Nutrition and Health, Volume 2: Data Sources.* Washington, D.C., U.S. Department of Agriculture, Economic Research Service, September 2002. http://www.ers.usda.gov/publications/fanrr19-4/

  Assessment of data bases with respect to their potential for analyzing the impacts of USDA's nutrition assistance programs. Items from some of the associated surveys may be applicable to evaluating nutrition education interventions