

Operation Frontline: Assessment of Longer-term Curriculum Effectiveness, Evaluation Strategies, and Follow-up Methods



Continuing Education Questionnaire available at www.sne.org/ Meets Learning Need Codes for RDs and DTRs 3020, 4010, and 4040.

Sara Swindle, MS; Susan S. Baker, EdD; Garry W. Auld, PhD, RD

ABSTRACT

Objectives: To determine the effectiveness of a traditional pretest versus a retrospective pretest, the stability of self-reported behavior changes at 3 or 6 months post-class series, and the most effective method for longer-term follow-up assessment of Operation Frontline's *Eating Right* class series.

Design: Longitudinal study; subjects surveyed at some combination of traditional pretest, retrospective pretest/posttest after the last class, and at 3 or 6 months.

Setting: Communities in the metropolitan Denver area.

Participants: Of the 53 participants, 90% were women, 70% were of Hispanic descent, 49% were between the ages of 20 and 29, and 64% had at least a twelfth-grade education or General Educational Development diploma.

Main Outcome Measures: Contrast results from a traditional pretest with a retrospective pretest; assess temporal stability of self-reported behavior changes; and contrast response rates for multiple follow-up methods (mail, telephone, or reunion class).

Analysis: Item and scale scores were compared across various time points using repeated-measures analysis of variance.

Results: No significant differences were found for 6 of 7 variables between the traditional pretests and retrospective pretests. Most self-reported behavior changes were retained at 3 and 6 months post-class. Mail had the best response rate (62%) for this limited-resource population.

Conclusions and Implications: The maintenance of behavior changes at 3 or 6 months post-intervention supports the effectiveness of the *Eating Right* series. Mail follow-up achieved higher response rates than telephone. The lack of significant differences (except 1 variable) between the traditional pretest and the retrospective pretests does not support the retrospective pretest as being more accurate. However, the retrospective pretest may be more desirable for educators, as it eliminates the need to test participants twice.

Key Words: nutrition education, evaluation, retrospective pretest, limited-resource

(*J Nutr Educ Behav.* 2007;39:205-213)

INTRODUCTION

Limited-resource audiences often are targeted by nutrition education programs seeking to improve dietary intake and quality and increase physical activity. However, the effectiveness of nutrition education has not been well re-

searched, especially on populations from a lower socioeconomic status. Because the goal of many programs is to change behavior, evaluation of behavior is essential to measure the effectiveness of these programs. As McClelland et al observe, "There is no gold standard for dietary evaluation of free-living people and little consensus concerning appropriate methods for assessing the impact of nutrition education on limited-resource populations."¹ Ammerman et al found that interventions were more successful at achieving desired behavior changes among those at risk of disease than among healthy populations, supporting the concept that participants' motivation is a critical factor influencing behavior change.² Similarly, in a meta-analysis

Department of Food Science and Human Nutrition, Colorado State University, Fort Collins, Colorado

Address for correspondence: Garry W. Auld, PhD, RD, 210 Gifford, Department of Food Science and Human Nutrition, Colorado State University, Fort Collins, CO 80523-1571; Phone: (970) 491-7429; Fax: (970) 491-7252; E-mail: auld@cahs.colostate.edu.

©2007 SOCIETY FOR NUTRITION EDUCATION

doi: 10.1016/j.jneb.2007.03.003

to determine if nutrition education improves dietary practices and changes behavior, Contento et al found that the most successful programs use strategies to affect behavior change in environmental interventions, as opposed to simply disseminating large amounts of information.³ Although this review confirmed that programs do have an impact on behavior change, most of the studies conducted only short-term evaluations.

Several national programs that have made efforts toward longer-term evaluation include the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)⁴ and the Expanded Food and Nutrition Education Program (EFNEP).⁵ Few local programs have made efforts to perform longer-term evaluations, although the Stanford Nutrition Action Program⁶ and La Cocina Saludable—The Healthy Kitchen,⁷ have conducted longer-term evaluations.

Share Our Strength, a national nonprofit organization, established Operation Frontline (OF), a community nutrition education program, to address hunger among limited-resource populations.⁸ Operation Frontline's *Eating Right* curriculum is a 6-week class series offered to limited-resource adults in both English and Spanish.⁹ The curriculum includes topics such as planning balanced meals, low-fat and nutritious cooking, smart shopping, and appropriate food-safety practices, which are taught by a volunteer nutritionist, accompanied by a cooking demonstration and food preparation skills, taught by a volunteer chef.⁹ According to Flesch-Kinkaid's analysis, printed handouts and materials are written at a fourth-grade reading level.¹⁰ Participants also receive a grocery bag of food to take home, providing the food resources needed to replicate the recipes demonstrated in class.

The *Eating Right* curriculum incorporates constructs of the Theory of Experiential Learning, in which individuals are active participants in the learning process and participants' prior experiences inform the development of the learning environment and methodology. The instructor's role is to facilitate the learning experiences to meet the needs of the learner, and the learners' prior experiences and reflections on those experiences contribute to the process. This approach provides a comfortable motivational environment in which participants interact with one another as well as the instructor during lesson activities.¹¹

Past evaluation efforts by OF have consisted of only immediate post-evaluations with the use of a retrospective pre/post test given on the last day of class. A retrospective pre/post test asks respondents to report on their behaviors at 2 points in time, prior to the onset of classes and at the completion of classes, ie, they complete a posttest and "retrospective" pretest at the same time on a single instrument. As an example of an OF question, the survey asks, "How often do you use olive oil or canola oil while cooking?" The answer choices include "Never," "Rarely," "Sometimes," "Most of the time," and "Almost always," and the subject answers for the 2 time points, "Before class" and "Now."

This type of evaluation assesses the short-term changes in behavior, knowledge, or skills that are reported at the end of the class sessions.¹² Operation Frontline evaluations have consistently demonstrated improvements in self-reported behaviors after participating in *Eating Right* classes. Over an 18-month period, 63% to 91% of almost 2200 participants showed improvement, depending on the behavior measured.¹³ On the 5-point Likert scale used by OF, average improvements are 0.57, 0.74, and 0.78 on their Food Resource Management, Eating Habits, and Nutritional Practice scales, respectively.

Although a pre/post test methodology is common in many programs, response shift bias is a potential shortcoming.¹⁴ There is an assumption in the traditional pre/post test protocol that "a person's standard for measurement of the dimension being assessed will not change from pretest to posttest. If the standard were to change, the posttest ratings would reflect this shift in addition to the actual changes in the person's level of functioning."¹⁵ Thus, the response shift could lead to an apparent difference between pretests and posttests that is not real or that masks a real difference. Rohs (2001) concluded that the retrospective pre/post test provided a more accurate reflection of behavior change. Although it acknowledged potential shortcomings of a retrospective pretest, such as a social desirability bias or problems with recall, OF judged the retrospective pre/post test approach as more appropriate for their clients and program because the information could be gathered at 1 time point after their clients were familiar with the subject matter terminology.^{12,13}

In addition to their interest in a longer-term follow-up to track the sustainability of self-reported behavior change, the developers of OF also wanted to identify the method of follow-up (phone call, mail, reunion class) with the highest response rate. The 3 objectives for this study were to: (1) compare results of a traditional pretest with a retrospective pretest; (2) assess the nutrition behavior changes (and the temporal stability of those changes) among *Eating Right* participants at a 3 or 6-month follow-up; and (3) track the overall response rate of multiple follow-up methods (mail, telephone, and reunion class) and determine the most effective method of contacting limited-resource groups for future OF follow-up.

DESCRIPTION OF THE EVALUATION

Study Design

Approval for this project was obtained from the Human Research Committee of Colorado State University. All study participants completed the *Eating Right* retrospective pre/post test. Participants who enrolled in OF during the 4-month data collection period also completed a traditional pretest. Individuals who completed the *Eating Right* classes up to 6 months prior to this 4-month data collection period were potential subjects. These individuals were surveyed at 3 or 6 months after completion of their *Eating Right* class

series; their participation rate is reflected in the response rates for the follow-up surveys. Participants were randomly assigned by class to receive the 3- or 6-month follow-up surveys by mail, telephone, or in person at a reunion class.

The limited total pool of applicants, small window of time for data collection, limited resources, and sporadic class schedule made it unrealistic to have a control group. These factors also precluded the collection of data from all participants at all time points (traditional pretest, retrospective pre/post test, and 3- or 6-month follow-up).

Subjects

The study involved subjects who participated in the 6-week OF *Eating Right* classes, which are offered to limited-resource families in the Denver metropolitan area. Operation Frontline participants in Colorado are referred by over 30 collaborating human service agencies, such as WIC and food banks. We attempted to survey subjects from 17 *Eating Right* class series (ranging from 6 to 12 participants). These class series included new groups enrolled during the 4 months of data collection, as well as those groups who completed the *Eating Right* series up to and including 6 months prior to the start of data collection.

Instruments

The retrospective pre/post test is the standard OF survey. All participants enrolled in *Eating Right* during the data collection period completed a traditional pretest and retrospective pre/post test. Behavior questions were divided into scales based on their similar constructs and prior combinations used by OF (Table 1). The scales were titled *Eating Behaviors*, *General Behaviors*, and *Shopping Behaviors*. Cronbach α was used to confirm the scales' internal consistency (α values ≥ 0.68) using data from a different sample of OF participants. Items on the food safety scale did not have

high internal consistency and were not used as a scale in this study. However, 3 items from the proposed food safety scale were analyzed individually, as was an item on the frequency of eating breakfast, which did not load highly on the scale. The Eating Behavior Scale items were not asked on the traditional pretest.

Grouping similar, correlated survey questions to produce scales provides several benefits in evaluation. Using multiple items to assess a single, underlying construct may provide a better estimate of that construct, and a more reliable assessment, than would a single item.¹⁶ The Eating Behavior, General Behavior, and Shopping Behavior scales were reliable according to Cronbach α . The lack of internal consistency (reliability) of the food safety scale might be attributed to several factors. The food safety content of the curriculum is limited in contrast to the proportion of the evaluation tool devoted to food safety issues; thus, the link between lesson content and the evaluation items might have been weaker. Also, it was speculated that people may be less willing to admit to some inappropriate food safety practices (washing hands) than others (leaving out leftovers). Thus, a variation in responses because of a social desirability bias could contribute to a lower internal consistency.

Demographic variables included ethnicity, age, and education levels. Participants indicated whether they were currently participating in WIC, the Child Nutrition Program (CNP), or food stamp program.

Data Collection

Data were collected between April 2004 and January 2005. Surveys were read aloud to OF participants by the nutritionist leading the sessions. For those participants recruited for the 3- or 6-month follow-ups, surveys were either read over the phone by a bilingual caller, self-administered when mailed, or read aloud at a reunion class. Surveys were

Table 1. Reliability of Scale Data for Operation Frontline Evaluation Study (N = 68)*

Scale Name	Cronbach α	Items	Item Correlation
Eating Behavior [†]	0.71	How often do you . . . use olive oil in cooking?	0.44
		eat 2-4 fruits per day?	0.39
		eat 3-5 vegetables per day?	0.59
		drink low-fat milk?	0.54
		prepare foods without salt?	0.39
General Behavior [†]	0.72	make meals that include a variety of foods from Food Guide Pyramid?	0.59
		think about healthful choices for family?	0.54
		have healthful snacks available?	0.48
Shopping Behavior [†]	0.68	read food labels when shopping?	0.50
		use a grocery list when shopping?	0.60
		compare prices when shopping?	0.60

*This sample of Operation Frontline participants was not part of the study.

[†]Survey questions based on a 5-point Likert scale, where 0 indicates never and 4 indicates almost always.

mailed to participants in their preferred language (Spanish or English). Participants did not complete surveys at every time point; rather, each participant was contacted at 2 of the 4 time points—at the beginning (traditional pretest) and end of class (retrospective pre/post test), or at the end of class and at 3 months or 6 months following the class. Those completing the follow-up surveys received coupons for free yogurt.

Data Analysis

Item and scale scores were compared across time points using repeated-measures analysis of variance (ANOVA) and the Fisher protected least significant difference (LSD) test for the following combinations: the traditional pretest, retrospective pre/post test; and the retrospective pre/post test and the 3-month or 6-month follow-up. A significant difference was indicated by a *P* value of less than .05. Responses from each method of data collection (telephone, mail, and reunion class) were tracked to determine which method produced the greatest response rate.

RESULTS OF THE EVALUATION

Demographics

The demographic data were collected by OF staff. The majority of the 53 subjects were women (> 90%) of Hispanic descent (70%). Thirteen percent were African American, and 11% were non-Hispanic whites. Forty-nine percent were between the ages of 20 and 29, and 32% were between the ages of 30 and 39. Over a third (36%) had less than a high school education, whereas 35% had a twelfth-grade education or a General Educational Development diploma. A third or less were participating in federal food assistance programs: WIC, 33%; CNP, 13%; or food stamps 19%.

Outcomes According To Time of Survey

Each participant was measured at 2 of 4 time points. All 53 participants completed the retrospective pre/post test (Table 2). Although 19 individuals completed the traditional pretest, only 12 of the 19 came to the last class and completed the retrospective pre/post test; these 12 were included in the 53 (Table 3). Twenty-seven of the 53 completed a 3-month follow-up, and 14 completed a 6-month follow-up (Table 4). Although 53 individuals completed a 3-month or 6-month follow-up (Table 5), only 41 of these subjects attended the final class in the series, at which the retrospective pre/post test was collected.

Retrospective pretest and posttest (n = 53).

Table 2 compares the scores from the retrospective pretest and posttest, for all participants. Scores improved from the retrospective pretest when compared to the posttest on all

Table 2. Outcome Scores According to Time: Retrospective Pretest and Posttest (n = 53)

Variable	Time of Evaluation†	
	Retrospective Pretest	Posttest
	LSmean (SEM)§	
Scales‡		
General Behavior	2.0 (0.1)	3.3 (0.1)***
Shopping Behavior	1.7 (0.1)	3.0 (0.1)***
Eating Behavior	1.8 (0.1)	3.0 (0.1)***
Items		
Eat breakfast	1.7 (0.1)	2.9 (0.1)***
Wash hands	3.3 (0.1)	3.8 (0.1)**
Thaw at room temp	3.2 (0.2)	3.3 (0.2)
Leave leftovers out > 3 h	3.5 (0.1)	4.3 (0.1)***

****P* < .001, ***P* < .01

Significance according to analysis of variance (ANOVA):

†Retrospective Pretest: given at the end of last class; Posttest: given at the end of last class with the retrospective pretest.

‡Survey questions based on a 5-point Likert scale, where 0 indicates never and 4 indicates almost always.

§LSmean indicates least square mean; SEM, standard error of the mean.

||Behavior items that should decrease from retrospective pretest to posttest because of the nature of the question but were reverse-coded to be consistent with other items, ie, a higher number reflects a more positive change.

3 scales, general behavior (2.0 vs. 3.3, respectively), shopping behavior (1.7 vs. 3.0), and eating behavior (1.8 vs. 3.0). In other words, there was a significant positive change in reported behaviors after the class was completed, when participants were performing learned class behaviors “most of the time.” Three of the 4 behavior items were significantly different, favoring the posttest.

Traditional pretest to retrospective pre/post test (n = 12).

Table 3 compares the reported behaviors of a subset of participants before the class started (traditional pretest survey) to behaviors reported after completion of the class (retrospective pre/post test). Higher scores on the 4-point scales indicate improved behaviors. General behavior scale scores were higher at the posttest than at the pretest and retrospective pretest (3.4 vs. 2.7 and 2.4, respectively), and shopping behavior scales were significantly higher at the posttest when compared to the pretest and retrospective pretest (2.9 vs. 2.5 and 1.9, respectively). The only significant difference on the scales between the pretest and the retrospective pretest was in shopping behaviors (2.5 vs. 1.9, respectively). No differences were seen for individual items across the 3 times.

Retrospective pretest, posttest, and 3-month (n = 27) or 6-month follow-up.

Table 4 compares 2 different subsets of participants to see if self-reported

Table 3. Outcome Scores According to Time: Traditional Pretest, Retrospective Pretest, and Posttest (n = 12)

Variable	Time of Evaluation†¶		
	Traditional Pretest†	Retrospective Pretest† LSmean (SEM)‡	Posttest†
Scales§			
General Behavior***	2.7 (0.2) ^a	2.4 (0.2) ^b	3.4 (0.2) ^{ab}
Shopping Behavior***	2.5 (0.1) ^{ac}	1.9 (0.1) ^{bc}	2.9 (0.1) ^{ab}
Items			
Eat breakfast	2.3 (0.3)	1.9 (0.3)	2.9 (0.3)
Wash hands	3.6 (0.1)	3.7 (0.1)	3.9 (0.1)
Thaw at room temp #	2.4 (0.4)	3.7 (0.4)	2.9 (0.4)
Leave leftovers out >3 h	3.7 (0.3)	3.5 (0.3)	4.2 (0.3)

*** $P < .001$, # $P < .06$

Significance according to analysis of variance (ANOVA):

†Traditional Pretest: given before the first class; Retrospective Pretest: given at the end of last class; Posttest: given at the end of last class with the retrospective pretest.

‡LSmean indicates least squares mean; SEM, standard error of the mean.

§Survey questions based on a 5-point Likert scale, where 0 indicates never and 4 indicates almost always.

¶Superscript items with the same letter are significantly different.

||Behavior items that should decrease from traditional pretest and retrospective pretest to posttest because of the nature of the question but were reverse-coded to be consistent with other items, ie, a higher number reflects a more positive change.

behavior changes were maintained. The 2 groups of participants were those who completed the retrospective pre/post test and the 3-month follow-up and participants who completed the retrospective pre/post test and the 6-month follow-up. There was no significant difference between the posttest and the 3- or 6-month follow-ups on general behavior (3-month, 3.2 vs. 3.1, respectively; 6-month, 3.5 vs. 3.0, respectively) and shopping behavior scales (3-month - 3.0 vs. 3.0 respectively; 6-month, 3.1 vs. 2.9, respectively), or the 4 individual items. These results suggest that participants' self-reported behavior changes were stable over time.

At 3 months and 6 months, scores on both scales and 3 of the individual items were lower (less preferred) on the retrospective pretest than the posttest and follow-up. The items were eating breakfast, washing fruits and vegetables, and not leaving leftovers out for more than 3 hours. These results suggest that self-supported behaviors prior to the *Eating Right* series had changed.

Response Rates

Survey follow-up by telephone, mail, or reunion class. Previous *Eating Right* participants (n = 135) were randomly assigned (by OF class) to 1 of 3 follow-up methods: telephone, mail, or reunion class (Table 5). Of 40 participants selected to be contacted by telephone, only 22 had telephone numbers that were still in service (54.6%). Those 22 were called up to 5 times before they were dropped; 12 participants were reached, yielding a 30% response rate via telephone.

Mailings were sent to 58 participants on up to 3 sepa-

rate occasions. The initial mailing of 58 surveys resulted in 16 returned surveys (27.5%) and 2 invalid addresses. The second mailing (40 surveys) resulted in only 3 returned surveys (7.5%) and an additional invalid address. The third (final) mailing consisted of 36 surveys, of which 17 (43%) were returned, yielding the best overall response rate. In total, 36 of 58 surveys were returned after 3 mailings, yielding a 62.1% response rate.

Of the 37 participants that investigators attempted to contact by telephone to invite to a reunion class follow-up, 17 (45.9%) were actually contacted, and 5 attended the reunion class. This result yielded a response rate of 29.4% from the 17 who were contacted, but only a 13.5% response rate for the 37 participants.

DISCUSSION

Longer-term Outcomes

This study examined self-reported behaviors as a result of a 6-week nutrition education program and measured the retention of behavior changes on a longer-term basis. By using constructs of the Theory of Experiential Learning,¹¹ the *Eating Right* curriculum provided the opportunity to influence behavior within an educational setting.^{3,17}

The results suggested that most self-reported behaviors were retained at 3 or 6 months after the class series. These results followed a pattern similar to findings from other longer-term follow-up studies conducted by WIC,⁴ EFNEP,⁵ and evaluations of other, smaller programs such as La Cocina Saludable.⁷ The WIC follow-up took place at 2 months and 1 year post-intervention; significant differences

Table 4. Outcome Scores According to Time: Retrospective Pretest to Posttest and 3-Month (n = 27) or 6-Month Follow-up (n = 14)

Variable	Time of Evaluation†¶		
	Retrospective Pretest	Posttest LSmean (SEM)§	3-Month Follow-up
Scales‡			
General Behavior***	2.0 (0.1) ^{ab}	3.2 (0.1) ^a	3.1 (0.1) ^b
Shopping Behavior***	1.7 (0.1) ^{ab}	3.0 (0.1) ^a	3.0 (0.1) ^b
Items*			
Eat breakfast***	1.9 (0.2) ^{ab}	2.9 (0.2) ^a	2.9 (0.2) ^b
Wash hands**	3.4 (0.1) ^{ab}	3.9 (0.1) ^a	3.9 (0.1) ^b
Thaw at room temp	2.7 (0.3)	3.3 (0.3)	3.4 (0.3)
Leave leftovers out > 3 h*	3.7 (0.2) ^{ab}	4.3 (0.2) ^a	4.2 (0.2) ^b
6-Month Follow-up			
Scales‡			
General Behavior***	1.7 (0.2) ^{ab}	3.5 (0.2) ^a	3.0 (0.2) ^b
Shopping Behavior***	1.4 (0.2) ^{ab}	3.1 (0.2) ^a	2.9 (0.2) ^b
Items			
Eat breakfast**	1.1(0.3) ^{ab}	2.8 (0.3) ^a	2.6 (0.3) ^b
Wash hands**	2.8 (0.2) ^{ab}	3.9 (0.2) ^a	3.8 (0.2) ^b
Thaw at room temp	3.7 (0.5)	3.5 (0.5)	4.1 (0.4)
Leave leftovers out > 3 h*	3.1 (0.4) ^{ab}	4.5 (0.4) ^a	4.5 (0.4) ^b

****p* < .001, ***p* < .01, **p* < .05

Significance according to analysis of variance (ANOVA):

†*Retrospective Pretest*: given at the end of last class; *Posttest*: given at the end of last class with the retrospective pretest; *3-month Posttest*: given 3 months after the class; *6-months Post-test*: given 6 months after the class.

‡Survey questions based on a 5-point Likert Scale where 0 indicates never and 4 indicates almost always.

§LSmean indicates least squares mean; SEM, standard error of the mean.

||Behavior items that should decrease from retrospective pre-test to post-test due to the nature of the question but were reverse coded to be consistent with other items, i.e., a higher number reflects a more positive change.

¶Superscript items with the same letter are significantly different.

Table 5. Survey Response Rates Based on 3 Different Follow-up Methods

Telephone: n=40 potential contacts	
Invalid Phone Numbers	18 (45%)
# of People Called	22 (55%)
# of People Contacted and Surveys Completed	12/22 (54.5%)
Overall Response Rate	12/40 (30%)
Mail: n=58 potential contacts	
Initial Mailing	58
# Returned	16 (27.5%)
1 st Follow-Up	40
# Returned	3 (7.5%)
2 nd Follow-Up	36
# Returned	17 (43%)
Total Invalid Addresses	3 (5.2%)
Overall Response Rate	36/58 (62.1%)
Reunion Class: n=37 potential contacts	
# of People Contacted	17 (45.9%)
# of People Who Attended and Surveys Completed	5/17 (29.4%)
Overall Response Rate	5/37 (13.5%)

in fat and fruit and vegetable intake seen between treatment and control groups at the posttest (2 months) remained at the 1-year follow-up, although the magnitude of the difference was less.⁴ In the EFNEP study, when measures of food practices (thawing, storing leftovers, using grocery lists, etc.) were compared over 3 time points, significant differences were found between the pretest and the posttest. The 1-year follow-up measures were similar to the posttest measures, but significantly better than the pretest, which suggests that behavior change was maintained.⁵

In the La Cocina Saludable study by Taylor et al, improvements in participants' self-reported behaviors followed a somewhat different pattern.⁷ When scores were compared to those at a pretest, participants demonstrated improvement in 3 of 5 behavior categories at the posttest and 5 of 5 behavior categories at the 6-month follow-up. In addition, scores in 4 of 5 behavior categories improved between the posttest and the follow-up. These results indicate that participants continued to make positive behavior changes after the class series ended. The previous findings and those found in this study support the view that nutrition education may have longer-term impact and that changes seen at a posttest are often sustained, or even improved, over a longer period of time.

Short-term Outcomes

Results from the self-reported behaviors on the retrospective pretest indicate that there was a positive behavioral change noted by participants. Participants reported significant improvements in general, shopping, and eating behaviors, and 3 of 4 individual behavior items (eating breakfast, washing hands, and leaving leftovers out for more than 3 hours). These findings are consistent with the review conducted by Contento et al on the efficacy of nutrition education, which showed that interventions using educational methods directed at behavioral change as a goal were more likely to be successful in actually changing behavior.³ The structure of the *Eating Right* class is a hands-on approach, in which participants are able to perform the behaviors, such as washing their hands and preparing breakfast items, themselves.

Traditional Pre-then-post Method Vs. Retrospective Pre/Post Test

The findings contrast with those of Nussbaum, who supported the use of the retrospective pretest as the evaluation method of choice, arguing that the traditional pre-then-post method was less effective in showing change.¹² Others contend that a retrospective pretest provides a more accurate assessment because the pretest may contain unfamiliar information or terminology, leading to guessing or misrepresentation of actual knowledge, which would create a response shift bias.^{15,18} A counterargument is that the retrospective pretest may create a “motivational” bias, that is, social desirability, because respondents now “know” the proper response.¹⁹ When comparing the 2 different types of survey methods (traditional pretest vs. retrospective pretest) relative to their ability to find differences or changes (sensitivity), no significant difference was found between scores on the pretest and the retrospective pretest on 5 of 6 variables. However, the small sample size ($n = 12$) and 1 additional variable (thawing at room temperature) with substantial but not significant differences between the traditional and retrospective pretests limits confidence in these results. Although these results neither prove nor disprove the equivalence of a traditional and retrospective pretest as a means of program evaluation, they do indicate the need for additional investigation with larger sample sizes.

One advantage of giving a retrospective pretest is that it essentially captures 2 time points simultaneously,¹⁹ as the survey is administered on the last day of classes and asks participants to report on their behaviors prior to the beginning of classes and their current behaviors. Operation Frontline recognized that the retrospective pretest might be easier for participants to understand, because they have been exposed to the terminology and content, and it may be more efficient, as less class time is required for data collection. Participant discomfort might also be lessened,

because the program does not start with a “test.”^{12,19} There is also some gain in efficiency relative to data entry, because it is faster to enter (or scan) data from a single survey versus separate pretest and posttest surveys, even though the total number of variables is the same.

Although the number of participants who took the traditional pretest was quite small ($n = 12$), these similar results suggest that there is no substantial difference between the traditional pretest and the retrospective pretest. The factors of convenience, time, money, and participant comfort, however, support the retrospective pretest as a more efficient method of evaluation to meet the needs of OF.

Follow-up Method Outcomes

The other main finding of this study was that follow-up by mail was the most effective method of reaching participants after the class series. Participants had expressed interest in a reunion class, but reunion classes were difficult to organize, and many participants were unable to attend.

Conducting a follow-up evaluation with limited-resource clients presents challenges owing in part to the more transient nature of this population.⁷ Mailing was the most successful way to reach participants in this study, with a 62.1% response rate, followed by telephone (30%), and then by use of a reunion class (13.5%). These findings do not coincide with the results of other investigators. For example, Dillman’s work with the general public revealed that a mailed survey can yield nearly a 60% to 75% response, whereas telephoning can yield a higher response rate, close to 85%.²⁰ Although this study was limited to low-income adults, who differ from the general public, the telephone seemed to be ineffective. For some, the telephone number was no longer in service, but even with a working phone, many people were not at home, even though they were called 5 different times during the day. Salant and Dillman recommend attempting to call 6 or more times,²⁰ suggesting that this response rate might have improved with additional phone calls. In this study, more participants responded to the mail method. The relatively high response rate with mailing may have been in part a result of the previous rapport developed between the OF volunteers and staff and the participants.

The low response rate of the reunion class was unexpected, as many participants had expressed a desire for another class. However, the response to a reunion class was confounded by using the telephone to contact participants for the reasons noted above. Other means of organizing a reunion class should be explored, for example by mailing invitations, before ruling this out as a viable follow-up method of contact.

When choosing a survey method, Dillman suggests a decision can best be made within the context of the survey topic and the population to be studied.²⁰ Mailing may be most beneficial for follow-up contact by OF with its limited resource class participants.

Limitations

Ideally a study such as the one described in this paper should include a control group and assess all participants at all time points. Operation Frontline offers classes a limited number of times in the year, with various and unpredictable start times. Instead of following these participants from a traditional pretest through a 6-month follow-up, it was more practical to survey a larger group of *Eating Right* participants across a shorter time frame. Although this approach is not ideal from a research perspective, it is practical for OF to follow given its limited staff and resources, as well as the aforementioned challenges in reaching a transient population over a period of several months. Having limited resources for research is not unique to private nonprofit organizations; it is also a common situation with government programs such as EFNEP and Food Stamp Nutrition Education (FSNE), where little money is available for “research,” but evaluations are expected.

The 3- and 6-month follow-ups were challenging to complete. Some of the subjects did not have the means to receive mail or telephone calls. Others did not have the time to devote to a reunion class, as it may have been difficult for them to get off work or pay for child care. The small sample size at the traditional pretest ($n = 12$) and 6-month follow-up ($n = 14$) reduces confidence in the statistical results, as small sample sizes could introduce a type 2 error (ie, not seeing a difference that would be apparent in a larger sample). Lastly, typical OF participants are transient, and the individuals the investigators were able to reach for the 3- or 6-month follow-up might have been less transient and, therefore, not as representative of the entire OF population.

Implications for Research and Practice

The evaluation of Operation Frontline’s *Eating Right* class series demonstrated positive behavior change immediately after the class series and also suggested that those behavior changes were sustained at 3 or 6 months post-intervention. Because the ultimate goal of any nutrition education intervention is to establish healthful habits over the long term, it may be limiting to the actual goals of an intervention to focus only on short-term changes.¹ This evaluation was the first to test for longer-term outcomes of OF, and the reported maintenance of self-reported behavior changes provides support for the effectiveness of the program. Although this study design did not allow a direct comparison across classes, the strong, consistent results, in spite of class series being delivered by different personnel (nutritionists and chefs), suggest that the 6-week program is robust.

This study identifies a promising method for conducting longer-term evaluation efforts. Specifically, with this transient population, mailings produced the best return rate. Although the investigators did not find differences in out-

comes between the traditional and retrospective pretests on 5 of 6 variables, more research is needed with limited-resource populations to assess their levels of comfort with these 2 approaches to program evaluation, as well as the accuracy of both methods (traditional and retrospective pretests) for effective program evaluation.

ACKNOWLEDGMENTS

The authors acknowledge and thank the administrators of Operation Frontline, Denver, Colorado, as well as the administrators of the national Operation Frontline program at Share Our Strength, for their support in conducting this project, specifically for facilitating the data collection and for providing access to program participants and demographic data related to the participants. In addition, the authors acknowledge the Nutrition Assessment and Evaluation Center in the Department of Food Science and Human Nutrition at Colorado State University for assistance with research design and data analysis.

REFERENCES

- McClelland JW, Keenan DP, Lewis J, et al. Review of tools used to assess the impact of nutrition education on dietary intake and quality, weight management practices, and physical activity of low-income audiences. *J Nutr Educ*. 1995;33:S35-S48.
- Ammerman AS, Lindquist CH, Lohn KN, et al. The efficacy of behavioral interventions to modify dietary fat and fruit and vegetable intake: a review of the evidence. *Prev Med*. 2002;35:25-41.
- Contento IR, Balch GI, Bronner YL, et al. The effectiveness of nutrition education and implications for nutrition education policy, programs, and research. *J Nutr Educ*. 1995;27:312-328.
- Havas S, Anliker J, Greenburg D, et al. Final results of the Maryland WIC Food For Life program. *Prev Med*. 2003;37:406-416.
- Arnold CG, Sobal J. Food practices and nutrition knowledge after graduation from the Expanded Food and Nutrition Education Program (EFNEP). *J Nutr Educ*. 2000;32:130-138.
- Howard-Pitney B, Winkleby MA, Albright CL, et al. The Stanford Nutrition Action Program: a dietary fat intervention for low-literacy adults. *Am J Public Health*. 1997;87:1971-1976.
- Taylor T, Serrano E, Anderson J, Kendall P. Knowledge, skills, and behavior improvements on peer educators and low-income Hispanic participants after a Stages of Change-based bilingual nutrition education program. *J Community Health*. 2000;25:241-262.
- Share Our Strength: Operation Frontline. Available at: <http://www.strength.org/what/-operationfrontline/>. Accessed February 27, 2007.
- Share Our Strength: Eating Right. Available at: <http://www.strength.org/what/-operationfrontline/curricula>. Accessed February 27, 2007.
- Kincaid JP, Fishburne RP Jr, Rogers RL, Chissom BS. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease formula) for Navy enlisted personnel. Research Branch Report 8-75, Millington, Tenn: Naval Technical Training, US Naval Air Station, Memphis, Tenn; 1975.
- Rogers CR, Freiberg HJ. *Freedom to Learn*. 3rd ed. Upper Saddle River, NJ: Prentice Hall; 1994.
- Nussbaum RJ. Operation Frontline program evaluation overview. *Share Our Strength*. June 2004;1-8.
- Ambrose K. An evaluation of Operation Frontline: Serviced community and impact. Available at: <http://www.strength.org/what/>

- operationfrontline/pdf/Ambrose2006_OFL_Evaluation.pdf. Accessed February 27, 2007.
14. Sprangers M, Hoogstraten J. Pretesting effects in retrospective pretest-posttest designs. *J Appl Psychol*. 1989;74:265-272.
 15. Rohs FR, Lanone CA, Coleman RK. Response shift bias: a problem in evaluating nutrition training using self-report measures. *J Nutr Edu*. 2001;33:165-170.
 16. Johnson RA, Wichern DW. *Applied Multivariate Statistical Analysis*. 2nd ed. Upper Saddle River, NJ: Prentice Hall; 1988.
 17. Lipski U, Igani E, for the Prevention Unit, Division of Preventive Oncology, Cancer Care Ontario. Report #3: Evaluation of nutrition interventions: a background document. November 2001;1-26.
 18. Rockwell SK, Kohn H. Post-then-pre evaluation: measuring behavior change more accurately. *J Extension* (serial online). 1989;27:1-7. Available at: <http://www.joe.org/-joe/1989summer/a5html>. Accessed February 27, 2007.
 19. Hill, LG, Betz, DL. Revisiting the retrospective pretest. *Am J Eval*. 2005;26:501.
 20. Salant P, Dillman DA. *How to Conduct Your Own Survey*. New York: John Wiley and Sons; 1994.

SNE's e-Catalog of DGA and MyPyramid Materials Released July 2006

Thanks to a grant from the Grain Foods Foundation, SNE has developed and released a searchable, online catalog to help nutrition educators find and market materials promoting the 2005 Dietary Guidelines for Americans and MyPyramid food guidance system.

What? The e-Catalog contains peer-reviewed materials ranging from 1 page leaflets to multi-media curriculum.

Where? On the SNE web site at www.sne.org/mypyramid

Why? To help SNE better serve its membership and the nutrition education community.

How to Participate?

To submit materials for potential review and inclusion in the online catalog or volunteer as a reviewer, visit the web site at www.sne.org/mypyramid.

Please share this information with your colleagues