Evaluating the Use of Questions and Responses in a Large National Dietary Data Collection Instrument

Lois Steinfeldt, John Clemens, and Jaswinder Anand, United States Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Food Surveys Research Group

1. Introduction

The U. S. Department of Agriculture's Automated Multiple Pass Method (AMPM) Blaise instrument collects 24-hour dietary recall data for the What We Eat In America (WWEIA), the dietary interview component of the National Health and Nutrition Examination Survey (NHANES). AMPM contains over 2,400 questions and more than 21,000 responses about foods with each response determining the next appropriate question. This results in approximately 500,000 possible paths through the questions within the AMPM. Each year it is used in approximately 9,000 interviews which ask individuals to recall the foods and beverages that were consumed the day before the interview. On average, 13 foods are reported for each 24-hour dietary recall and 10 questions are asked about each food. Of the 10 questions, 6 are probes about the description of the food and the amount consumed. During each year of the survey, interviewers use AMPM to ask respondents over a million questions about foods.

AMPM has been in continuous use since January of 2002 and in that time has been used to conduct over 60,000 dietary intake interviews. A large comprehensive update that includes major procedural changes is done for the beginning of each two-year survey period. A smaller update incorporating primarily new foods and portions is done as necessary for the second survey year. Updates are made to questions, response values, and to programming which controls what questions are asked based on previous responses. The purpose of the updates is to assure validity and completeness of the dietary data. These updates result from changes in the food supply and in food consumption patterns, and evaluation of the use of the questions and responses.

2. Qualitative methods

Qualitative evaluation methods are an important part of all phases of data collection. At the Food Surveys Research Group (FSRG), the quality assurance procedures are based on years of experience in identifying and eliminating the source of errors and inconsistencies that occur in food and nutrient intake data. Fortunately some past sources of errors, such as missing responses, have been eliminated with the automation of the collection method. Comprehensive training and monitoring programs for interviewers complement the automated system to help ensure that data collected are of good quality. Data collection oversight is provided for the dietary interviewers by FSRG staff. This includes direct observation of interviews, interviewer questionnaires, review of data collected, and informal feedback to interviewers. New interviewers complete a rigorous training program before starting work. Because of the changing food market and emerging nutrition issues, changes are made to the AMPM each year. These changes, as well as the need to assure standardization of data collection procedures, necessitate periodic refresher training sessions for the dietary interviewers. Training is conducted just prior to the launch of each year's data collection. On-site visits are conducted to observe the data collection and to provide technical oversight. The interviewers are monitored four times each year. During these field observations, interviewers are evaluated on the following:

> Following interviewing protocol for the survey
> Proper use of automated system
> Overall demeanor, rapport, and pace

Randomly assigned interviews are monitored by supervisors for quality assurance purposes.

3. Quantitative methods

Statistical analysis of the use of questions and responses provides a quantitative approach to evaluation which is particularly useful with a large and complex instrument such as AMPM. SAS 9.2 was used to statistically analyze the data. Results from four years, 2005 to 2008, were combined to produce over 2.8 million responses to questions in the AMPM about food details and amounts from more than 35,000 intakes. In 2008, AMPM contained 2,305 questions about food details and food amounts. In addition to these questions, there were 130 questions used to record "Same As" foods. AMPM allows the interviewer to capture and record a food as the same as one eaten previously in the day or the same as a food eaten by another household member. In these cases, AMPM does not prompt for the food detail questions again, but does prompt for the amount of the food eaten. This reduces the number of repetitive questions during an interview for frequently consumed foods such as coffee and when conducting interviews for members of a household who all ate the same dinner.

Table 1 shows the results for how many of the total questions in AMPM in 2008 were used during the 2005 to 2008 time period. Overall, 90% of the questions were used, with a higher percentage of food detail questions used than food amount questions. Overall, this is a very high percentage use of questions in an instrument which obtains detailed information on foods and amounts of food eaten in a large and diverse population with a wide-ranging and constantly changing food supply.

Table 1. Use of AMPM food detail and amount questions

	Count of questions	Questions used in 2005-2008 Number (Percent)
Food detail	950	909 (96%)
Amount	1,355	1,215 (90%)
Total questions	2,305	2,124 (92%)

Most of the questions in AMPM have the option to enter responses not on the list. These are designated as "Other, Specify" (OS) responses. The questions without this option are usually numeric or yes/no type questions. There were a total of almost 44,000 OS responses in the 4-year time period which is an average of about 1.2 OS responses per intake. OS responses are particularly important in a continuing national food consumption survey. They record and document new foods and new package sizes in the marketplace which may then be added to the AMPM. They are also used to collect unusual foods which are consumed infrequently by the U. S. population.

Of particular interest in evaluating the use of AMPM is the percentages of "Don't know" responses. Almost all the question in AMPM allow a "Don't know" (DK) response. There were a total of almost 60,000 DK responses in the 4-year time period which is an average of 1.7 per intake. The questions in AMPM with higher DK counts are food characteristics which are of interest, but may only be known to the person preparing the food. An example of this is the question "What percent lean was the ground beef?" One of the factors that make this question difficult for respondents is changes in labeling over the past decade. AMPM interviewers are trained to accept "Don't know" responses rather than encourage the respondent to guess. Food and nutrient composites are used as appropriate for analysis for foods with DK responses.

Table 2 shows the percent of questions with at least one DK or OS response for each year. Overall the percentages show very little change year to year. There does seem to be a slight downward trend in both DK and OS responses over the 4 years.

Table 2. Percent of questions with at least one DK or one OS response by year

	2005	2006	2007	2008
Don't know (DK)	43.5%	43.5%	43.4%	43.4%
Other, specified (OS)	43.6%	42.4%	41.4%	40.8%

Table 3 shows the percent of all responses that are DK or OS for each year. These percentages show a slight downward trend. The Cochran-Armitage test was used to test for trends in DK and OS responses over the 4 years. The null hypothesis for the Cochran-Armitage test is that there is no trend. One-sided and two-sided p-values are computed for the trend test. The right-sided p-value tests for increasing trend in the proportions and the left-sided p-value tests for decreasing trend. Overall, both DK and OS responses showed statistically significant decreasing trends with p-values <0.001.

Table 3. DK and OS as a percent of all responses by year[*]

	2005	2006	2007	2008
Don't know (DK)	1.9%	1.8%	1.6%	1.6%
Other, specified (OS)	1.4%	1.3%	1.2%	1.1%

[*] Cochran-Armitage Test for Decreasing Trend p <0.001 for both DK and OS.

While the overall numbers show very little change, there could be larger changes in individual questions. With the large number of questions in AMPM, checking the percentages of DK and OS responses for each question for each year would be very time consuming. Statistical tests of differences in the response percentages for DK and OS for the same question over the 4 years would provide indicators of where changes are occurring that need to be evaluated.

Of the 2,305 questions, 1827 were selected that were used in all 4 years. Of these, 713 (39%) had no DK responses over all 4 years and were eliminated from the DK analysis. 828 questions had no OS responses over all 4 years and were eliminated from the OS analysis. Table 4 shows that only two percent of the questions showed significant increases ($p < .05$) in the numbers of responses for both DK and OS. Eight percent of questions showed significant decreases ($p < .05$) in the number of DK responses and twelve percent of questions showed significant decreases ($p < .05$) in the number of OS responses. Most of the questions showed no significant trend for both OS (86%) and DK (90%) responses testing at $p < .05$.

Table 4. Results of trend tests for OS and DK responses

	Questions showing no significant trend	Questions showing significant decrease[1]	Questions showing significant increase[2]	Total number of questions
Don't know (DK)	1,005 (90%)	87 (8%)	22 (2%)	1,114
Other, specified (OS)	864 (86%)	116 (12%)	19 (2%)	999

[1] Cochran-Armitage Test for Decreasing Trend p <0.05.

[2] Cochran-Armitage Test for Increasing Trend p <0.05.

Using the statistical tests reduced the number of questions to be reviewed for DK and OS responses from 2,305 to 22 for DK and 19 for OS. Questions which showed decreasing or no trend, do not warrant evaluation related to these responses. There were no questions with statistically significant increases in both the numbers of OS and DK responses. Of the 22 questions showing increases in DK responses, 16 are food detail questions and 6 are amount questions. Of the 19 questions showing increases in OS responses, 14 are food detail and 5 are amount questions. Although these questions show a statistically significant upward trend in DK or OS responses, the total number of DK or OS responses remains small.

Statistical tests can also be used to look for changes in the percentages of existing response options. For example, the responses to the question "What kind of milk?" over the 4-year time period can be tested for statistically significant changes using Chi Square. Table 5 shows the percent of responses by year. The Chi Square test for differences in the distribution of responses over time is significant at $p < .001$. The percentages in table 5 show a shift to lower fat milk. Of particular interest is that there is very little differences in the percentages of soy, other and don't know responses. In this case the statistical test confirms a known trend in U. S. food consumption patterns. This same approach can be used for other questions, to look for other food consumption and /or food supply patterns in the data.

Table 5. Percent of responses to "What kind of milk?"*

Milk Kind	2005	2006	2007	2008
Whole milk	39	37	32	34
2% milk	37	36	39	39
1% milk	9	11	11	11
Skim/nonfat milk	9	10	10	10
Soy milk	2	2	2	2
Other	2	2	2	2
Don't know	2	2	2	2
Total	100	100	100	100

*Chi Square significant p < .001.

4. Conclusion

Continual review and update of the AMPM is required to reflect important changes in the U.S. food market such as new foods, ethnic foods, and new package sizes. Changes also are directed by the data collected in WWEIA and to address current public health concerns. With over 2,400 questions and more than 21,000 responses, it is not possible, nor is it needed, to review all the questions and possible responses on a yearly or bi-yearly basis. Statistical tests can be used to identify questions where the percentages of all the response options including "don't know" and "other, specified" may be changing over time. While changes in questions may be statistically significant, infrequently used questions and questions with low percentages of DK and OS responses may not be significant in terms of the food and nutrient data produced by the survey. However, as shown in this analysis, this approach can be used with other methods of prioritizing instrument review to substantially reduce the number of questions to be evaluated and focus scarce resources productively.