Using Cognitive Interviews to Improve Survey Instruments

Heather Haeger, Indiana University
Amber D. Lambert, Indiana University-Bloomington
Jillian Kinzie, Indiana University-Bloomington
James Gieser, Indiana University

Indiana University Center for Postsecondary Research
1900 East Tenth Street
Eigenmann Hall, Suite 419
Bloomington, IN 47406-7512
nsse@indiana.edu
Phone: 812.856.5824

Abstract

This research features results from cognitive interviews and focus groups conducted as part of the psychometric testing to improve questionnaire design and inform revisions to the National Survey of Student Engagement (NSSE). Findings are specific to NSSE, but include suggestions for adapting the approaches for use at institutions, as well as consider implications for other surveys of college students.

Introduction

As the pressure to assess educational practices increases, it is crucial to ensure that the instruments used for that assessment accurately measure educational practices and experiences. More specifically, it is critical that questionnaires measure what they intend and that respondents understand and correctly interpret items. The National Survey of Student Engagement (NSSE) is a survey often used by colleges and universities to assess the time and effort students dedicate to educationally purposeful activities. This paper provides an overview of the process of evaluating and making revisions to the NSSE survey instrument and includes a discussion of findings from cognitive interviews and focus groups.

Conceptual Framework

Researchers have long focused on problems related to the wording of survey questions. However, only in the past three decades have survey questions and instruments been assessed in a more systematic way, aided by theoretical frameworks and methods based in cognitive and social psychology (Collins, 2003; Pressser, Couper, Lessler, Martin, Martin, Rothgeb, & Singer, 2004). The cognitive interview method has come to be viewed as an important means to ensure the quality and accuracy of survey instruments and is used to identify and analyze sources of response error in survey questionnaires (Willis, 1999; 2005).

Cognitive interviewing is meant to identify and analyze sources of response error in survey questionnaires by focusing on the cognitive processes respondents use to answer questions on a survey or questionnaire. Specifically, the purpose of the method is to understand whether subjects understand the question, both consistently across subjects and in the way intended by the researcher (Collins, 2003). Because cognitive interviewing focuses on the cognitive processes respondents use to answer questions on a survey or questionnaire, both overt/observable and hidden processes are studied (Willis, 1999; 2005). The purpose of cognitive interviews is to focus on the survey question, not the person answering the question. The researcher is not focused on the subject's cognitive processes for their own sake, but on the soundness of the survey question.

Theoretically, there are four actions that the subject enacts when answering a survey question: they must comprehend the question, be able to retrieve information, make a judgment as to its relevance and accuracy as an answer to the question, and respond to the question (Collins, 2003; Daugherty, Harris-Kojetin, Squire, & Jael, 2001; Jobe, 2003; Willis, 1999; 2005). First, the subject must understand the survey question and the response options provided. If the subject interprets the meaning of a question or the response options in a way other than intended by the survey designers, drawing accurate conclusions based on the subject's answer may be in jeopardy. Once the question is understood correctly, the subject must be able to accurately perform primary survey tasks. These include the accurate retrieval of information related to the question, drawing conclusions about this information, and possibly conducting mental arithmetic computations. The retrieval of information is a complex, multi-step process that demands several steps, some of which may be conscious and some unconscious; some retrieval processes will be automatic, while others may go unnoticed by the subject (Willis, 1999; 2005). The concern is

whether the subject is able to make a judgment and formulate their answers to the information being requested (Collins, 2003).

Finally, once the subject has formed a judgment to a given question, s/he enters the response stage of the cognitive interview model (Collins, 2003). In this stage the subject must formulate a response according to the explicit response options provided. The choice of pre-specified response options may influence the way the subject forms a judgment and chooses a response. Moreover, at this stage the subject may edit his/her response in order to provide a more socially accepted or desirable answer. Because the interviewer cannot know precisely what is going on inside the subject's head throughout this process, the interviewer employs a variety of techniques to prompt the subject to verbalize the process of arriving at a response to provide insights into his/her cognitive processes.

Methodologically, the practice of cognitive interviewing falls into two general techniques: think-aloud and verbal probing (Daugherty, Harris-Kojetin, Squire, & Jael, 2001; Willis, 1999; 2005). Think-aloud refers to an explicit activity in which the subject is encouraged to "think aloud," or verbalize, his/her thought processes as s/he answers survey questions. This can take place concurrently, in which the researcher reads the question to the subject, then observes and records as the subject "thinks aloud" about the question as s/he answers it, or retrospectively, when the subject is asked to verbalize his/her thoughts about the questions at the conclusion of the survey (Redline, Smiley, Lee, & DeMaio, 2001). The interviewer is primarily passive in either version of this technique, aside from providing encouragement for the subject to "tell me what you're thinking" if s/he hesitates or pauses.

Advantages of the think-aloud method:

- Freedom from bias imposed by frequent interviewer interjections

- Minimal interviewer training requirements

- Open-ended design

Disadvantages of the think-aloud method:

- Subject usually requires training in the method

- Subject may resist the technique

- Possibility for subject to stray from the topic at hand

- Subject may bias his/her description of his/her decision processing

The verbal probing method can either be concurrent or retrospective (Willis, 1999). Concurrent verbal probing is characterized by additional, specific questions that are designed to elicit further information after the subject has provided a response. These probes can be scripted or spontaneous. Retrospective probing is when subjects are asked at the end of the interview to verbalize their thoughts about questions they answered earlier when taking the questionnaire. Whether concurrent or retrospective, scripted or spontaneous, verbal probing is interested in finding out more about a particular aspect of the question-answering process (Daugherty, Harris-Kojetin, Squire, & Jael, 2001). Probes have become increasingly favored by researchers.

Advantages of the verbal probing method:

- Interviewer maintains control of the interview

- Relative ease of training the subject

Disadvantages of the verbal probing method:

- Artificiality; that is, this technique is not reflective of a typical survey interview in which the researcher simply asks questions and the subject answers them

- Potential for bias through poor selection of probes

Although cognitive interviews have been identified as an integral part to survey design, researchers are typically concerned with standardizing survey questionnaires and procedures, without taking time to be sure that subjects understand the questions (Collins, 2003). Attending solely to questionnaire design or simply pilot testing a questionnaire are insufficient methods to ensure the quality and accuracy of the instrument. The inclusion of cognitive interviews during the development of a survey instrument is an important and necessary component of survey research. In support of this, this study illustrates an approach to cognitive interviews for testing surveys conducted with college students and introduces a hybrid "think-aloud" and probing technique. In addition, the study highlights some of the survey issues that could go unknown without the use of cognitive interviews.

Methods

In-depth testing of the NSSE instrument using cognitive research testing has been conducted several times since the surveys inception in 1999 (Kuh, Kinzie, Cruce, Shoup & Gonyea, 2006; Ouimet, Bunnage, Carini, Kuh, & Kennedy, 2004). The most recent series of cognitive tests were conducted as part of the planned update to the NSSE survey (scheduled to launch in 2013). Specifically, cognitive testing was undertaken to explore the cognitive processes that students use to answer survey questions and to identify items that are not well understood by respondents. Several techniques to evaluate students understanding of questions were used, including cognitive interviews, using a combination of both the think-aloud and verbal probing techniques described above, as well as focus groups. These approaches were used to determine how well students understood survey items, as well as to test new and modified survey items and different response options. In-depth cognitive interviews were conducted with 120 students and ten focus groups were conducted with a total of 79 students at 12 different Midwest campuses.

Campuses selected to test the instrument represented a range of institutional characteristics including size, control, religious affiliation, and Carnegie classification.

During the first wave of interviews in 2010, a variety of techniques were used in the student interviews. In some instances, students were asked to answer a specific question from the survey and then to describe how they interpreted that question. This interview process was designed to inform the development of specific survey items being tested on the first pilot instrument in 2011. Concurrent verbal probing was used in this first round of interviews. This method involves the interviewer asking a question, having the subject answer, and the interviewer asking more specific questions designed to elicit further information about the response (Willis, 2005). These probes can be scripted or spontaneous. Retrospective probing was also used where subjects are asked at the end of the interview to verbalize their thoughts about questions they answered earlier when taking the questionnaire. The advantages of this method are that the interviewer maintains control of the interview, and the relative ease of training the subject. The disadvantage of this type of cognitive interview is in its artificiality. This technique is not reflective of a real survey interview, in which interviewer simply asks questions and respondent answers them. The risk is a potential for bias resulting from poor selection of probes. In addition to concurrent and retrospective probing, participants were also asked to "think-aloud" while answering some questions. Think-aloud refers to an explicit activity in which the subject is encouraged to verbalize his/her thought processes as s/he answers survey questions. The interviewer is mainly passive in this process, aside from providing encouragement for the subject to "tell me what you're thinking" if s/he hesitates or pauses. The advantages of this method are the freedom from bias imposed by frequent interviewer interjections, minimal interviewer training requirements, and open-ended design. The disadvantages are that subjects usually

requires training in the method, subjects may resist the technique, possibility for subjects to stray from the topic at hand, and subjects may bias their descriptions of their decision processing.

The next wave of data collection occurred in 2012 and was aimed at both testing specific items, and evaluating the instrument as a whole. Concurrent probing, retrospective probing, and think-aloud were again used to test specific items, but participants in the cognitive interviews were asked to take the full, online survey and to provide feedback about specific questions identified in the interview protocol (these questions were new or modified questions on the NSSE instrument) and any questions that were ambiguous or confusing. The participants were instructed to stop after completing several questions on the online survey and were then asked follow-up question on specific items. Students were also asked to think-aloud as they answered the questions in order to understand how students understood those questions. During focus groups, students were given a paper copy of the survey and instructed to read and answer specific question. Concurrent probing was also used to ask specific follow-up questions about items on the survey. Which items to address in focus groups was determined for each campus. Items that were more challenging for students at that type of institution were tested further in the focus group. In addition to face-to-face interviews and focus groups, the final wave of interviews also involved phone interviews with students who were enrolled at an institution that is completely online. The students took the survey online while speaking to an interviewer on the phone. These students were asked the same questions about specific items and were also instructed to point out items that were difficult to answer as a student taking all their classes online.

During the cognitive interviews and focus groups, an interviewer and a note taker were present. The interviews and focus groups were also recorded. Analytic memos were created

based on the notes and digital recordings from the cognitive interviews and focus groups. These

memos contained a summary of each student's response to specific items and the investigators

impression of the quality of the survey question based on those responses. These memos were

coded into the following categories: a) no problem noted on the item, b) minor misunderstanding

or problem, c) significant problems. The assignment of these codes was done independently by

two reviewers to assure inter-rater reliability. A third reviewer analyzed the coding done by both

researchers and addressed any differences in coding. A summary of the different ways that

students interpreted the questions was created to inform question writing for items that students

understood, but answered in different ways. The team of researchers who conducted the

cognitive tests jointly analyzed findings and thoroughly vetted interpretations and conclusions

across items. Findings from this wave of data collection and analysis, along with other

psychometric testing, were used to inform the development of new and modified survey

questions and response options.

Findings

Results from the cognitive interviews and focus groups demonstrated that students

interpreted and responded to many of the tested survey items in the intended ways and largely

confirmed past evaluations of items (Kuh et al., 2006; Ouimet et al, 2004). This was true even

for more complex items that involved calculating the time, frequency, or number of occurrences

of various activities. Students employed cognitive strategies such as thinking about their course

syllabi to count the number of papers they had written or calculating the average number of

papers per class and generalizing from that figure to all of their classes. Students were also able

to easily recall how often they do common behaviors like asking questions in class or worked

with other students on a course project. This research also revealed areas for improvement on the

survey. The analysis of the cognitive interviews revealed error in the following categories: 1) language problems – not knowing the meaning of words/phrases, 2) inclusion/exclusion problems – determining whether certain concepts are to be considered within the scope of an item, e.g. does personal reading online apply to work completed for classes 3) temporal problems – these involve the time period to which the question applies, e.g. how many years or high school did the student consider when answering the question about most of his high school grades 4) logic problems – how do students interpret phrases like 'and' and 'or' in survey questions and 5) computational problems – difficult mental arithmetic (Conrad & Black, 1996; Willis, 1999; 2005).

The first type of error to emerge from the analysis was with students not understanding some of the terminology introduced in new survey items. For example, a number of terms used by practitioners and researchers in higher education, such as "learning communities" or "faculty-supervised research," were confusing for students. Multiple versions of these questions were tested to determine what wording would be clearest to students. A "not applicable" option was also added to questions where students struggled with terminology so students wouldn't have to skip the question or guess a response if they didn't understand who the question was referring too.

The language problems in the survey were also related to some inclusion/exclusion problems. For example, when using the term "student affairs professional" in the 2010 survey and 2011 cognitive interviews, students were including anyone they had interacted with on campus including campus security, dining hall staff, and janitorial staff, but were often leaving out staff in their residence hall or student support center. To combat these issues, the wording was changed to "student services staff" and a parenthetical with examples was added. This item

was tested with the new working in the 2012 cognitive interviews. Some students still included interactions with staff not included in the parenthetical, but the focus was much narrower.

The specific item level testing did not yield any temporal problems within an individual item, but temporal problems emerged while testing the survey as a whole. While individual items were worded clearly, there was a section of the survey where items shifted from asking students to report experiences within the last year and asking them to report experiences throughout their college experience. Asking students to repeatedly change back and forth between the time frames required more cognitive energy for students and led some to misreport information because they were thinking of the wrong time frame. This problem will be addressed in the final version of the survey by reordering items to be more consistent in the time frame students were looking at.

The only item that had errors caused by a logic problem was an item trying to measure how often students had meaningful interactions with people who were different from them in various ways. The question asked students to rate how often they "had serious conversations with people who differ from you in the following ways…" including people who differ in their political views or in their racial/ethnic background. Students often answered one part of the question but did not make the connection between serious conversations and people who were different from them. Students answered this question in one of three ways. Few students answered the way the question was intended, which was based on having serious conversations with people who were different from them. Many students answered the questions thinking about having serious conversations about the listed issues, but did not consider who they were having the conversation with (e.g. a number of students thought about discussing issues of race relations in the US with someone they knew, but did not think about whether that person was of a different race from themselves or not). This resulted in students including experiences that did

not involve interactions with people who were different from themselves. Finally, some students answered in close to the way intended, but only included serious conversations with someone who was different from them if they were specifically talking about that difference (e.g. these students would not count a serious conversation with someone of a different race unless they were talking about race). This resulted in students under reporting diverse interactions because they were limiting they're answer by the subject matter. To address this problem, the question was reworded to make it a less complex cognitive task and so that it emphasizes the people you interact with, not what you talk about.

Most of the items that required students to perform some form of computation to answer the question were easy for students to understand and formulate a response. One item that students had difficulty computing was how many pages they read per week. Students generally thought about how long a typical chapter is, how many chapters were typically assigned per class, and how many classes they were taking. Students were able to perform this part of the computation with some effort, but greater error was created in trying to count less standardized readings. For example, students had a difficult time determining how many pages they read in formats without page numbers (e.g. readings online, reading on an e-reader, or books on tape). Students also had a difficult time calculating how much they actually read as opposed to what was assigned because it was more difficult to remember and add together parts of readings instead of generalizing how many chapters were assigned.

Conclusion

The findings from this research have been invaluable in the revision of the NSSE survey instrument, and have implications for practice for researchers and institutions interested in creating and/or using surveys on their campuses. Furthermore, these approaches can be adapt by

institutions that want to understand more about the meaning their students make of survey items

and can also help institutional research staff and others working with their NSSE data better

contextualize their results. It can be valuable for institutions to conduct their own in-depth

explorations of survey results by asking their students to explain what survey questions are

asking using specific examples from their campus context. Furthermore, sharing this

contextualized information about students' views of items and what they have in mind when they

respond to survey questions can help enliven campus discussions about survey results by

providing concrete examples of student behaviors and institutional practices. To facilitate this

work, NSSE created a toolkit for campuses to implement modified cognitive testing (see:

www.nsse.iub.edu/pdf/cognitive_interviews_facilitation_guide.pdf) In the session, we will share

a guide for conducting modified cognitive interview and focus groups on college campuses and

will also discuss the implications of this work for other college student surveys.

References

Conrad, F. & Blair, J. (1996). From impressions to data: Increasing the objectivity of cognitive interviews. In *Proceedings of the Survey Research Methods Section of the American Statistical Association,* Alexandria, VA.

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research, 12*, 229-238.

Daugherty, S. D., Harris-Kojetin, L., Squire, C., & Jael, E. (2001, August 5-9). Maximizing the quality of cognitive interviewing data: An exploration of three approaches and their informational contributions. Proceedings of the Annual Meeting of the American Statistical Association.

Jobe, J. B. (2003). Cognitive psychology and self-reports: Models and methods. *Quality of Life Research, 12*, 219-227.

Kuh, G. D., Kinzie, J., Cruce, T., Shoup, R., & Gonyea, R. M. (2006, July). Connecting the dots: Multi-faceted analyses of the relationships between student engagement results from the NSSE, and the institutional practices and conditions that foster student success: Final report prepared for Lumina Foundation for Education. Bloomington, IN: Indiana University, Center for Postsecondary Research.

Ouimet, J. A., Bunnage, J. B., Carini, R. M., Kuh, G. D., & Kennedy, J. (2004). Using focus groups to establish the validity and reliability of a college student survey. *Research in Higher Education, 45* , 233-50.

Presser, S., Couper, M.P., Lessler, J.T., Martin, E., Martin J., Rothgeb J.M., & Singer E. (2004). Methods for testing and evaluating survey questions. *Public Opinion Quarterly, 68*(1):109-130.

Redline, C., Smiley, R., Lee, M., & DeMaio, T. (2001, August 5-9). Beyond concurrent

     interviews: An evaluation of cognitive interviewing techniques for self-administered

     questionnaires. Proceedings of the Annual Meeting of the American Statistical

     Association.

Willis, G.B. (2005). Cognitive Interviewing: A Tool for Improving Questionnaire Design.

     Thousand Oaks, CA: Sage Publications.

Willis, G. B. (1999). Cognitive interviewing: A "How To" guide. Research Triangle Institute.

     1999 Meeting of the American Statistical Association. Research Triangle Park, NC:

     Research Triangle Institute.