

Exploring a Method to Evaluate Survey Response Scales

René Bautista¹, Lisa Lee¹

¹NORC at the University of Chicago, 55 East Monroe, Suite 2000, Chicago, IL 60603

Abstract

We present results from a qualitative study of two alternate scales used in a survey of consumer attitudes toward health information technology. Respondents were presented with a 4-point scale for questions on privacy and security of electronic health records (*very—somewhat—not very—not concerned at all*). Across these questions, 25% to 50% of respondents expressed the highest level of concern (*very concerned*). It was unclear whether this was a ceiling effect or whether this reflected the need for an additional option between *very* and *somewhat concerned*. We conducted cognitive interviews in English and Spanish to test an alternative 5-point scale that included *fairly concerned*. Respondents first answered the questions using the 4-point scale and then decided whether they would have chosen a different answer on the 5-point scale. Respondents also visually represented on a line the distance between options on the 4- and 5-point scales. Distances between response options were not always perceived as equidistant and some respondents switched the ordering of two options, *fairly* and *somewhat*. The cognitive interviews also highlighted issues that were unique to the Spanish translation.

Key Words: cognitive testing, response scales, measurement error

1. Introduction

In the theoretical paradigm of Cognitive Aspects of Survey Methodology (CASM), questionnaire features—including question stem and options—play a critical role in the cognitive and communication process behind the answering task (Schwarz, 2007; Tourangeau, 2003, Willis, 2008). As important as how questions are asked in a survey is how respondents provide their answers. Most surveys use closed-ended formats, allowing respondents to choose from among a set of answer choices that the questionnaire designer has determined in advance. However, in order to gather the most accurate and precise measurements of respondent behaviors and attitudes, the response categories that are provided must allow respondents to select an answer that closely matches the answer they want to give. In particular, an issue that can arise when providing respondents with a scale is providing appropriate categories on the scale to allow for sufficient gradation in measurement of response across questions and across respondents. In multiple situations, even when the question stem has been appropriately formulated (e.g., a question asking for a single concept, presenting an adequate period time, using simple terms), the *respondent's interpretation and use of a response scale* may represent a challenge in the response process.

The survey methodology literature has widely adopted a four-stage cognitive model to understand psychological mechanisms around the survey response process (Tourangeau, et al., 2000). Namely, the respondent engages in a comprehension phase, performs a memory retrieval process, makes a judgment decision, and maps judgments and information onto a response option. These mechanisms help identify likely causes of errors from a cognitive perspective in a survey item. A pretesting technique that explores how the four-stage model relates to specific survey item features is *cognitive interviewing* (Willis, 2005).

Typically, cognitive interviews tend to focus more on the *stem* of the question and less on the *response scale*. However, the four-stage cognitive model—particularly the comprehension and judgment stage—can be helpful to shed light on how respondents understand and use *response scales*. The literature on questionnaire design that relies on standardized interviewing, suggests that core concepts and language used in the instrument need to be *consistently understood across respondents* (Fowler, 1995; Fowler, et al., 2008). While response scales are core components of a survey item, they are usually assumed to have the same interpretation among respondents. In practice, however, response scales are subject to different interpretations; such differences may constitute a source of measurement error in the response process.

The present study was conducted by NORC under subcontract to the MITRE Corporation for the Office of the National Coordinator for Health Information Technology. In this study, we explored a cognitive interviewing method to investigate how individuals understand response scales. Particularly, we examined two alternative categorical response scales (i.e., a four-point scale and a five-point scale to measure concern). We used a cognitive approach that presents the response scale as a continuum and asks respondents to place scale categories on the continuum. We utilized this method to evaluate both the *order* in which respondents place response categories, and *distance* between response options on the continuum.

2. Research Question

In this methodological research we investigated individuals' interpretation of two response scales that measure concern for privacy and security of electronic health records; a four-point scale and a five-point scale (Figure 1). Both scales share similar characteristics but differ in the fifth added category ("*fairly concerned*"). Examples of the questions in the survey are presented in Figure 1.

Privacy means you have a say in who can collect, use and share your medical record. How concerned are you about the privacy of your medical record?

Security means having safeguards to keep your medical record from being seen by people who aren't permitted to see them. Safeguards may include technology. How concerned are you about the security of your medical record?

Figure 1: Example survey questions on Privacy and Security of Electronic Health Records

The research question that guided our investigation was: How do respondents interpret and use the terms “Very concerned / *Fairly concerned*/ Somewhat concerned / Not very concerned / Not concerned at all” in the context of questions on privacy and security of electronic health records. Figure 2 shows the concern scales that were tested.

Four-point scale	Five-point scale
<input type="checkbox"/> Very concerned <input type="checkbox"/> Somewhat concerned <input type="checkbox"/> Not very concerned <input type="checkbox"/> Not concerned at all	<input type="checkbox"/> Very concerned <input type="checkbox"/> <i>Fairly concerned</i> <input type="checkbox"/> Somewhat concerned <input type="checkbox"/> Not very concerned <input type="checkbox"/> Not concerned at all
Note: The response option that is different in the five-point scale is shown in italics.	

Figure 2: Two Concern Scales for Questions on Privacy and Security of Electronic Health Records

Further, we investigated how Spanish speakers interpret and use those terms. Figure 3 displays the scales in Spanish.

Four-point scale	Five-point scale
<input type="checkbox"/> Le preocupa mucho <input type="checkbox"/> Le preocupa algo <input type="checkbox"/> No le preocupa mucho <input type="checkbox"/> No le preocupa nada	<input type="checkbox"/> Le preocupa mucho <input type="checkbox"/> <i>Le preocupa bastante</i> <input type="checkbox"/> Le preocupa algo <input type="checkbox"/> No le preocupa mucho <input type="checkbox"/> No le preocupa nada
Note: The response option that is different in the five-point scale is shown in italics.	

Figure 3: Spanish Version of Two Concern Scales for Questions on Privacy and Security of Electronic Health Records

3. Data and Methods

Participants were recruited from the Chicago metropolitan area by a third party vendor for a one-hour interview. The demographic profile of participants for this qualitative study tried to match socio-demographic characteristics of the general population, with an “oversample” of African Americans and Hispanics/Latinos. A total of 25 interviews were completed between July 8 and August 1, 2013; 15 of them were conducted in English and in 10 Spanish. Figure 4 displays the demographic characteristics of the cognitive interview participants.

<i>TOTAL</i>	<i>Number of participants</i>	<i>% of total interviews</i>
Gender:		
Male	12	48%
Female	13	52%
Race/Ethnicity:		
AA	9	36%
Hispanic/Latino	10	40%
White/Caucasian	6	24%
Asian	0	0%
Age ranges:		
Age 18-25	2	8%
Age 26-35	3	12%
Age 36-45	6	24%
Age 46-55	5	20%
Age 56-70	4	16%
Age 71+	4	16%
Missing	1	4%

Figure 4: Cognitive Script Used to Evaluate Changes in Usage of a 4- and 5-point Concern Scale

Cognitive interviews were conducted at NORC offices in Chicago by survey methodologists and project staff. The set of 10 Spanish interviews were conducted by native Spanish-speaking NORC survey methodologist. Conventional consent forms were administered in person to participants, who received a \$40 incentive for their participation. Given that the actual survey design corresponds to a telephone survey, the draft survey questionnaire was administered first by an interviewer over the telephone from a separate conference room and, upon completion, a cognitive interview was conducted. On average, completing the draft questionnaire took 20:43 (min:sec) —the average administration time for 15 English interviews was 19:56 (min:sec), and the average administration time for 10 Spanish interviews was 21:54 (min:sec).

After the draft survey was administered on the phone (i.e., after the first 20 min), the interviewer returned to the room to conduct a face-to-face cognitive interview on a set of specific items in the remainder of the hour. A delayed retrospective probing approach was followed for cognitive testing; namely, after all survey questions were asked, the cognitive interview was conducted. Typical semi-standardized cognitive scripts were used to investigate issues on comprehension, memory erosion, judgment and response.

As part of the cognitive retrospective protocol, we investigated issues related to interpretation and use of response scales for five survey items. These five items asked on concerns about privacy and security of electronic health records, and they used the previously described four-point concern scale (Figure 2). During the cognitive testing, a five-point scale was administered. Specifically, cognitive interviewers were instructed to administer the script displayed in Figure 5 during cognitive testing.

I am going to repeat some of the questions from the survey, but this time I will give you a slightly different set of answer choices. These choices are the same except in between “very” and “somewhat” the option “fairly concerned” has been added.

Figure 5: Cognitive Script Used to Evaluate Changes in Usage of a 4- and 5-point Concern Scale

Cognitive interviewers were instructed to present the five privacy and security questions on paper with respondents’ original answers that they provided on the 4-point scale. Interviewers were instructed to have the respondent review each of the answers to determine whether the respondent would have changed responses based on the new five-point scale (very/fairly/somewhat/not very/not at all concerned).

To further understand respondents’ comprehension and use of the 4- and 5-point scales, we explored a method to evaluate survey response scales. This method relies on paper and pencil instruments during cognitive probing. Respondents were presented a line (approximately 6 inches long) with labeled ends. The leftmost end of the continuum displays the label “Not concerned at all” whereas the rightmost end is labeled as “Very concerned.” For the 4-point scale, participants were asked to place two categories on the continuum: “Not very concerned” and “Somewhat concerned”. Additionally, for the 5-point scale, respondents were asked to indicate where they would place three response options over the end-labeled line: “Fairly concerned,” “Somewhat concerned,” and “Not very concerned.” The administration of the end-labeled lines was rotated across participants. Figure 6 shows the script used for this cognitive testing and also presents an illustration of the end-label continuum used.

<p><i>Here is a line. If “very concerned” is on this end, and “not concerned at all” is on this end, where along the line would you put “somewhat concerned” and “not very concerned”? Put X’s on the line to show me where these options would go.</i></p>	<p><i>Now here is another line. This time, please show me where “fairly concerned”, “somewhat concerned”, and “not very concerned” would be.</i></p>
<div style="text-align: center; margin-bottom: 10px;"> <hr style="width: 80%; margin: 0 auto;"/> </div> <div style="display: flex; justify-content: space-between; font-size: small;"> Not concerned at all Very concerned </div> <div style="margin-top: 20px;"> <ul style="list-style-type: none"> × Somewhat concerned × Not very concerned </div>	<div style="text-align: center; margin-bottom: 10px;"> <hr style="width: 80%; margin: 0 auto;"/> </div> <div style="display: flex; justify-content: space-between; font-size: small;"> Not concerned at all Very concerned </div> <div style="margin-top: 20px;"> <ul style="list-style-type: none"> × Fairly concerned × Somewhat concerned × Not very concerned </div>

Figure 6: Cognitive Script and Illustration of End-labeled Lines Used to Evaluate Comprehension and Order of Categories in a 4- and 5-point Concern Scale

4. Results

Out of the 25 respondents presented with both scales (4- and 5-point concern scale) during cognitive probing, 3 were excluded from analysis because they were *not very* or *not at all concerned* on survey items, 4 who changed answers for reasons unrelated to the scale, and 1 respondent with missing data, leaving 17 cases for analysis. 8 of 17 (47%) respondents selected *fairly* for at least one of the five concern questions when this option was presented (as described in Figure 5). 9 of 17 (52%) respondents chose not to change their answers at all. 5 of 9 (55%) Spanish speakers reported that “*mucho*” (*very*) and “*bastante*” (*fairly*) meant the same thing (as described in Figure 2).

Participants conducted a task in which they were asked to graphically show the relationship between response categories on the 4- and 5-point concern scale (they conducted the exercise separately for each scale). They marked on the line where they thought the middle two or three categories fell.

Figure 7 shows results from the five-point scale exercise placing “fairly concerned” and “somewhat concerned” on the end-labeled continuum. These are aggregated results—it is not possible to see where a single respondent put *fairly* and *somewhat* but the overall placement of these terms on the line is shown.

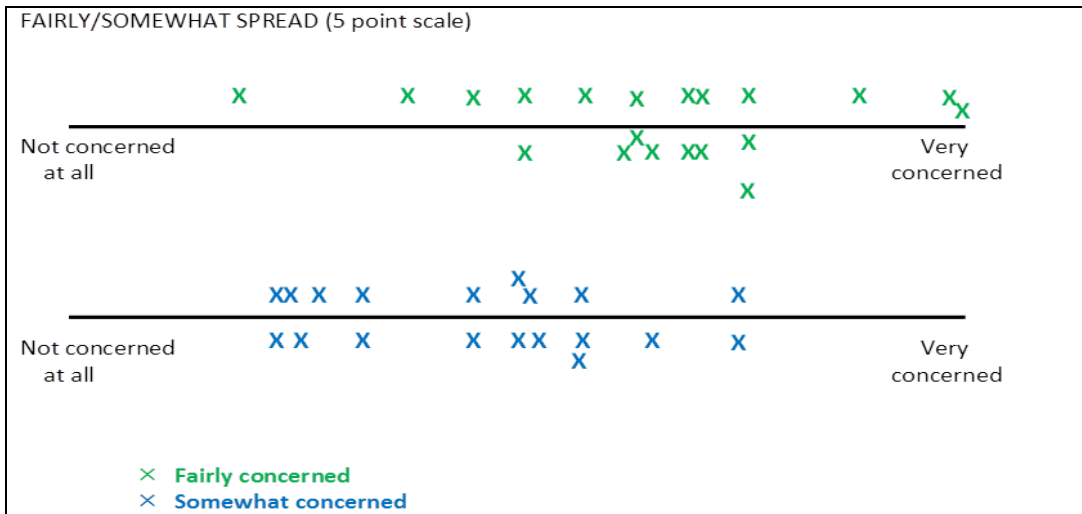


Figure 7: Results From a Task in Which Respondents Were Asked to Graphically Show the Relationship between the Response Categories on the Concern Scale (Results for 5-point concern scale)

Some data were missing because one participant left early, responses were illegible or because of unmarked lines (i.e., hard to tell which was fairly/somewhat), or interviewer error (i.e., forgot to administer this task). Nonetheless, with the existing data we observe that *fairly concerned* is spread a little more than *somewhat concerned*, but both show great variation across respondents.

Four of 19 (21%) participants who did the exercise thought that *somewhat* was a stronger term than *fairly*. Fourteen of 19 participants (74%) placed *fairly* and *somewhat* close together (within same quarter on the line). Only five of 19 (26%) placed *fairly* and *somewhat* farther apart.

In the context of this project, we recommended using the 4-point scale instead of the 5-point scale. Most respondents did not choose *fairly* when it was offered, and there was great overlap across respondents in how intense the terms *fairly* and *somewhat* were perceived to be. Most respondents did not perceive a great deal of distance between *fairly* and *somewhat*. Further, given that the Spanish translation for *fairly* (we used “bastante”) was challenging, a better translation of the 5-point scale would require changing the labels for all the points on the scale; we recommended that the 4-point scale would be more likely to increase reliability.

References

- Fowler, F. J., 1995. *Improving Survey Questions: Design and Evaluation*. Thousand Oaks: Sage.
- Fowler, F. J., Cosenza & Carol, 2008. Writing effective questions. In: J. H. D. D. Edith D. de Leeuw, ed. *International Handbook of survey methodology*. New York: Lawrence Erlbaum Associates.
- Schwarz, N., 2007. Cognitive Aspects of Survey Methodology. *Applied Cognitive Psychology*, Issue 21, pp. 277-287.
- Tourangeau, R., 2003. Cognitive aspects of survey measurement and mismeasurement. *International Journal of Public Opinion Research*, 15(1), pp. 3-7.
- Tourangeau, R., Rips, L. J. & Rasinski, K., 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.
- Willis, G. B., 2005. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks: Sage.
- Willis, G. B., 2008. Cognitive Aspects of Survey Methodology (CASM). In: *Encyclopedia of Survey Research Methods*. s.l.:Sage, pp. 104-107.