

Cognitive Testing of English and Spanish Versions of Health Survey Items

Roger Levine (American Institutes for Research)
Raquel González (American Institutes for Research)
Beverly Weidmer (RAND)
Patricia Gallagher (Center for Survey Research,
University of Massachusetts, Boston)

Background

A survey of patients' hospital experiences is being developed as a CAHPS[®] survey. Since survey items will be administered in English and Spanish, it is important to develop items that will function equivalently in these different languages. As part of the item development process, English versions of survey items went through several rounds of cognitive testing. An English version of the instrument was prepared and field tested.

This version was translated into Spanish and cognitively tested. The protocol for cognitive interviewing in Spanish was based on the English protocol. This enabled comparisons of the types of item problems that were found in the English and Spanish cognitive testing sessions. These items also provided insights into the judgment formation processes used by English and Spanish language respondents in responding to items about their health care. In previous CAHPS[®] surveys, it was noted that even though Spanish language respondents had worse reports of actual health experiences, their ratings of their overall health care, their personal doctor, specialists, and their health plan was higher than comparable ratings provided by white, English speaking respondents (Weech-Maldonado et al., 2003).

There is a growing body of research that seems to indicate that Latinos may have a tendency to provide more socially desirable responses than their non-Latino counterparts. While there isn't a definitive study that's been able to provide an answer as to why this is the case and while there is a certain level of disagreement among the research community as to why this can happen, many researchers attribute this finding to certain Latino cultural values. The Latino culture has been characterized by several researchers (Marin & Triandis, 1985 and Hofstede, 1980) as an allocentric culture, that is, a culture that emphasizes the needs, objectives and attitudes of an in-group (as opposed to individualistic cultures which emphasize the personal needs, objectives, attitudes and values of the individual). The importance of the in-group in Latino culture promotes behaviors that lead to smooth and pleasant relationships, emphasize positive behaviors and de-emphasize negative behaviors, and avoid interpersonal conflict. The Latino culture has also been characterized as demanding politeness, respect and discouraging criticism, confrontation and assertiveness (values associated with *simpatía*). Both of these cultural values (*allocentrism* and *simpatía*) have been associated with greater socially desirable responses

by Latino respondents.

Another hypothesis to explain this tendency to provide socially desirable responses among Latinos is that respondents with lower socioeconomic status are also more likely to provide more socially desirable responses. That is, respondents with low social power or prestige may be more deferential when responding than those with more social power and prestige. There are studies that support this hypothesis (Ross and Mirowsky, 1984) and that also found that socioeconomic status and degree of acculturation are highly correlated among Latinos. Finally, some studies (Marin et. al., 1991; Hui & Triandis 1989) have found that Latinos prefer the extreme response choices in greater proportions than non-Hispanic whites, which may also be explained by the need to avoid confrontation and criticism.

Methodology

After preparation of an initial "cognitive testing" version of a hospital CAHPS[®] instrument (called the HCAHPS instrument), a draft cognitive interviewing protocol was prepared and administered by each of the grantee teams. This protocol provided a listing of scripted probes that could be employed to provide insights into each respondent's cognitive processes as he or she answered the pilot items. It also included a series of general questions about the items, to allow the respondent to provide additional feedback about the items and to help assess the comprehensiveness of the instrument. A think aloud training exercise, with practice questions and a scripted response for the interviewers to use in modeling appropriate thinking aloud behaviors, was also included. A similar protocol was prepared for the second round of cognitive testing and administered by the grantee teams.

Slightly different procedures were used by each of the CAHPS[®] II grantee teams (the American Institutes for Research (AIR), Harvard, and RAND). These procedures reflected the belief that the instrument would probably be administered by both mail and telephone. Accordingly, it was reasonable to test the items as both self-administered items and as interviewer-administered items. The different procedures also reflected methodological preferences of each of the grantee teams.

To learn how the items functioned as self-administered items, AIR researchers administered the survey as a self-administered instrument. After training, respondents were asked to think out loud as they completed each item; scripted and unscripted probes were employed, as necessary. Harvard researchers used interviewers to administer the survey, conducting some of the interviews as in-person interviews and others as telephone interviews. Prior to the start of testing, respondents were asked several questions about their hospitalization experiences. These introductory items were used to prepare the respondent for answering a series of items about their hospitalization. In addition, the information

obtained was used to validate responses, triggering unscripted probing to determine the etiology of perceived inconsistencies. RAND researchers administered the survey as an in-person interviewer-administered instrument. All of the teams used the protocol as a guide, administering scripted probes as deemed appropriate and developing and employing other probes as needed. Each grantee team used the same procedures for both rounds of English language cognitive interviewing and for the Spanish language cognitive interviewing.

We conducted a total of 18 interviews in the initial English language round of testing (December 2002 - January 2003) and conducted 13 interviews in the second English language round of testing (February 2003). After the second round of English language cognitive testing, the instrument underwent minor revisions and was translated into Spanish. The Spanish language cognitive interviewing protocol was nearly identical with the English language protocol. Certain probes were modified or eliminated, reflecting anticipated item revisions. Other probes were added, to inform about the efficacy of alternative wordings of items (in Spanish).

Respondent characteristics. In all rounds of cognitive testing, respondents had to have been hospitalized for at least 24 hours, within the previous five months. In the first round of testing, 4 interviews were conducted in Palo Alto, CA; 3 were conducted in the Raleigh-Durham, NC area; 5 were conducted in the Los Angeles, CA area, and 6 were conducted in the Boston, MA area. Seven of the respondents were male; 11, female. Respondent ages ranged from 24 - 86 (mean = 51.1). Thirteen respondents were white; 4, Black; and 1, Hispanic.

In the second round of cognitive testing, 6 interviews were conducted in Palo Alto, CA; 3 were conducted in the Los Angeles, CA area, and 4 were conducted in the Boston, MA area. Five of the respondents were male; five, female. Respondent ages ranged from 22 - 91 (mean = 53.2). All of the respondents were white.

For the Spanish language cognitive testing, additional criteria were imposed: Respondents had to be of Hispanic/Latino descent, had to speak Spanish and, if given a choice about being interviewed in English or Spanish, they had to choose to be interviewed in Spanish. Four interviews were conducted in Palo Alto, CA; 4 were conducted in the Los Angeles, CA area, and 4 were conducted in the Boston, MA area. Five of the respondents were male; 7, female. Respondent ages ranged from 25 - 54. Four respondents were from Mexico; 1 from Central America; 5 from South America (2 Argentina, 1 Peru, 1 Venezuela, 1 Colombia); 1 from the Dominican Republic; and 1 was a native of the U.S. with a Dominican Republic background.

Results and Discussion

Similar problems. Most of the problems that were identified were either found in both the Spanish and

English versions or were idiosyncratic to the respondent (and almost certainly independent of language or culture). These problems are discussed below.

Underlying cognitive investigations of data collection instruments is a model of the questionnaire response process (Tourangeau, 1984; Sudman, Bradburn, and Schwarz, 1996). This model decomposes the process of responding to an item into four stages: comprehension, retrieval, judgment formation, and response production. The types of problems observed are classified according to this schema.

PROBLEMS FOUND WITH BOTH ENGLISH AND SPANISH LANGUAGE RESPONDENTS

Common comprehension problems. The first stage of the survey item response process is comprehension of the question. If respondents do not understand the question's real intent, the item will almost surely fail. There were several issues that interfered with item comprehension that were common to both English and Spanish language respondents. These are summarized below.

(1) Medical surveys often ask about health-related procedures and issues for which the use of technical terms is almost unavoidable. At least one Spanish language respondent and at least one English language respondent had difficulties with certain words or phrases. For example, "urinal/orinal (*bedpan*)" was interpreted by at least three English language respondents as referring to a porcelain urinal in a men's room. At least two Spanish language respondents were unfamiliar with orinal. Two others referred to orinal as "the duck," another referred to it as the "gallon." There seems to be a tendency for Spanish speakers to use euphemisms when referring to things they prefer not to discuss, leading to the use of regional slang.

(2) The term "supplements," referring to vitamins, minerals, or proteins, was not well understood by English language respondents. Similarly, "suplementos o complementos" was not well understood by all Spanish language respondents. Some English and Spanish language respondents thought this referred to over the counter medicines and/or herbal remedies.

(3) "Trauma unit/urgencias" was another phrase that was not universally understood by English language and Spanish language respondents. Respondents, regardless of language, did not distinguish this from the Emergency Department/Sala de Emergencia or simply did not know what the term meant.

(4) In order to determine if there were unreasonable delays during the admissions process, it is important that respondents have a common understanding of when the admissions process begins and ends. Problems were noted among both Spanish and English language respondents with the following items:

Think about when you were admitted to the hospital for this stay. Were there any unreasonable delays during the admission process?

Recuerde cuando le ingresaron al hospital esta vez, ¿se tardaron más de lo razonable durante el proceso de admitirlo al hospital?

Regardless of language, there was some confusion about when the admission process began for those who came in through the Emergency Room (ER). Similarly, for voluntary admissions, regardless of language, some respondents included “pre-admission” behaviors while others did not.

(5) In order to determine whether patients would have the help they needed after they left the hospital, the following item was developed:

Before you left the hospital, did you talk with someone about whether you would have the help you needed when you were discharged?

Antes de salir del hospital, ¿habló alguien con usted sobre si tendría la ayuda que necesitaba cuando le dieran de alta?

This item was intended to be inclusive of all kinds of help -- including both professional help and help from friends and family members. An explicit probe was developed to determine the types of help the respondent was thinking about. Most English language and Spanish language respondents thought the item was asking about both professional help and help from family and friends. However, some English language respondents thought it was asking only about professional help and others thought it was asking only about non-professional help. Similarly, at least one Spanish language respondents thought it was asking only about professional help and at least one thought it was asking only about non-professional help.

(6) “Living will/testamento en vida” was not understood by at least half of the English and half of the Spanish language respondents. As a result of the English language cognitive testing, a definition of living will was prepared and included in the Spanish cognitive testing. Even with a definition, at least five Spanish language participants did not know what the term meant.

(7) “Double negative” items present comprehension problems for both English language and Spanish language respondents. One item asked:

How often did were these tests and procedures done without causing you too much pain? (Always/ Usually/ Sometimes/ Never)

¿Con qué frecuencia le hacían esos procedimientos o exámenes médicos sin causarle demasiado dolor? (Nunca/ A veces/ Normalmente/ Siempre)

A response of “They were never done without causing me too much pain” contains two negative terms. These

types of items are difficult to cognitively process and resulted in several confused English language respondents¹. Spanish language respondents also had similar problems. At least one did not correctly use this set of response options. He responded “Never” to mean that they had never caused him pain. (For a different “double negative” question, translation into Spanish was able to avoid a double negative. This is discussed subsequently.)

(8) Occasionally, an item’s literal meaning and its intended meaning (that is to say, its subtext) are not identical. The following items were intended to provide an indicator of how often hospital staff would volunteer an explanation of a medicine’s purpose before they gave a patient a new medicine:

Before giving you any new medicine, how often did doctors, nurses, or other hospital staff tell you what the medicine was for?

Antes de darle cualquier medicamento nuevo, ¿con qué frecuencia le dijeron los doctores, enfermeras u otro personal del hospital para qué era ese medicamento?

One English language respondent only got one new drug but said, “Usually.” This was because he had to ask to find out -- they didn’t tell him spontaneously. Similarly, one Spanish language respondent responded, “Normalmente” because he had to ask the staff what the new drugs were for. They both recognized the item’s subtext and did not respond to the item’s literal meaning. However, neither responded with respect to the desired interpretation of the question, indicating an item problem.

Common comprehension/retrieval issues. Working memory capacity is limited to about five bits of information. When questions containing more than five bits of information are presented orally, respondents are forced to ignore certain parts of the item. (These items, when presented in written form, do not require the respondent to keep all of the information in working memory.)

(1) For this reason, the following item did not work well as an interviewer administered item:

Before you left the hospital, did you get information in writing about what activities you could and could not do?

Antes de salir del hospital, ¿le dieron información por escrito sobre las actividades que podía o no podía

¹ This item was originally phrased positively, asking about how often unnecessary pain was associated with the tests and procedures. Unfortunately, a positive (“always”) response is indicative of poor care. Since there was an objective to have “always” serve as a positive indicator (to facilitate the reporting of results, the item was rewritten as the version that was tested.

hacer?

At least two English language respondents attended to the word “activities” and interpreted activities to include taking medications and diet activities. One Spanish language respondent attended to the phrase “written instructions” and reported about written instructions to continue with antibiotic treatment at home. The length of the item seems to be responsible for both English and Spanish language respondents including medicine taking information -- a behavior that the item did not intend to capture.

(2) For similar reasons, the following item did not work well as an interviewer administered item:

Before you left the hospital, did you get information in writing about what symptoms or health problems to look out for after you were discharged?

Antes de salir del hospital, ¿le dieron información por escrito sobre los síntomas o problemas de salud a los que debía poner atención cuando le dieran de alta?

When administered by an interviewer, English language respondents focused on one part -- usually “information” and ignored “writing.” So, if the information was provided verbally, they would respond affirmatively. Conversely, at least three Spanish language respondents (including one patient who was administered this item in written form), would focus on “written information” and report about written information about their medicines or information about their condition in general.

Common judgment formation/response production issues. In order to solicit an evaluation of overall care, CAHPS® asks respondents to evaluate care received on a 0 - 10 rating scale, where “0 is the worst possible care and 10 is the best possible care.” An alternative rating scale (Perfect/Excellent/Very good/ Good/Fair/Poor; Perfecto/Excelente/Muy Bueno/Bueno/ Regular/Malo) was tested as an alternative. Several of the English respondents and at least three of the Spanish respondents refused to give ratings of Perfect/ Perfecto. They explained, “there’s no such thing as Perfect,” “Perfecto is an ideal,” “Perfecto does not exist,” and “Perfecto is too much to ask.”

One of the CAHPS® survey design principles is that respondents must be knowledgeable informants. If a respondent lacks the experiential basis for forming a judgment, but provides a response, the validity of the response should be questioned. Both English language and Spanish language respondents provided responses to the following items with an experiential basis:

During this hospital stay, did your family and visitors receive the help they needed when they called or visited the hospital?

Mientras estuvo en el hospital esta vez, ¿con qué frecuencia recibían sus familiares y amigos la ayuda que

necesitaban cuando llamaban o iban al hospital?

At least four of the English language respondents and at least two of the Spanish language respondents didn’t know whether or not their family or visitors received the help they needed but assumed that the absence of complaints was indicative of the fact that they had no problems. Accordingly, they responded “Always/Siempre.”

PROBLEMS FOUND WITH ONLY ENGLISH LANGUAGE RESPONDENTS OR WITH ONLY SPANISH LANGUAGE RESPONDENTS

Idiosyncratic responses. Regardless of language or culture, certain situations that create item problems invariably arise. Although the following issues arose for either an English language respondent or a Spanish language respondent, we feel that these issues are idiosyncratic and not related to language or culture.

(1) One English language respondent had difficulty with the following item:

How often did hospital staff respond quickly when you asked for pain medicine?

She answered, “Never,” and explained that she didn’t ask for pain medicines but asked for the staff to “reduce her pain.” She wanted holistic treatments in addition to pharmaceutical options. Since the staff didn’t do this, she responded negatively.

(2) One Spanish language respondent had a great deal of difficulty with the following item:

Before you left the hospital, were you told to take any medicine at home that you had not taken before this hospital stay?

Antes de salir del hospital, ¿le dijeron que tomara algún medicamento en casa que no había tomado antes de que estuviera en el hospital esta vez?

One respondent had a great deal of trouble answering this question. He was given a heart medication in the hospital that was the same as the one he had taken before being admitted, but it was under a new name. When he filled this prescription after discharge, he took both the “new” medicine and the old medicine and ended up with a mild overdose. There is probably no way we could have crafted this question that would remedy this particular problem without losing the meaning for a majority of respondents.

(3) Certain idiosyncratic responses have cultural origins. In private clinics, one’s expectations of care are much higher than in a regular hospital. One Spanish language respondent who went to the hospital for the birth of her child had exceedingly high expectations for care. She felt that the doctors didn’t come around frequently enough, even though the doctor was present for the birth of her child, would come around every 6 hours, and spent a lot of time with her. She wanted the doctor to come around

more frequently during childbirth. Similarly, she didn't feel the nurses spent enough time with her because they weren't always in her room. (They came by every 10 - 15 minutes.) She said that to get a rating of 9 or 10, the nurse would have to be her private nurse -- and she felt an "8" rating was very good.

(4) Certain items were interpreted differently by respondents who were admitted through the Emergency Department. Such problems were more typical of English language respondents, probably reflecting the fact that only one Spanish language respondent in our sample was admitted through the Emergency Department.

Comprehension: Linguistic issues. Linguistic issues can create, exacerbate, minimize, or eliminate problems with survey items. For several different items, linguistic issues helped ameliorate or eliminate certain problems that were observed with the English versions of these items.

(1) In English, the word "you" can be singular or plural. In Spanish, there are different words for "you:" "usted," "tu," and "ustedes." Usted is the formal way of saying you. One would use this form if one was speaking to an elder. Tu is the informal form, used when speaking with a friend. In the survey, the formal form was always used. However, there are different singular and plural versions of "you" in Spanish. "Usted" is singular and "Ustedes" is plural. (Tu is always singular). So, when someone says "usted," it is interpreted as singular. As a result, the following question:

During your hospital stay, how often did the nurses treat you with courtesy and respect?

led one English language respondent to include experiences of family members in forming a judgment. Not infrequently, English language respondents would include their child's experiences when forming judgments for analogous items. For example, when asked,

During your hospital stay, how often did the doctors explain things about your illness or treatment in a way you could understand?

one respondent focused on her daughter's illness and treatment.

These situations did not arise when the Spanish versions of these items were administered:

Mientras estuvo en el hospital esta vez, ¿con qué frecuencia le trataban las enfermeras con cortesía y respeto?

Mientras estuvo en el hospital esta vez, ¿con qué frecuencia le explicaban las cosas los doctores en una forma que usted pudiera entender?

Even though "usted" did not appear in the first item, the verb form makes "usted" implicit.

(2) Spanish translations enabled the avoidance of another problem: Confusion associated with a "double negative" item. One item asked:

How often did nurses answer your call button without a long wait? (Always/ Usually/ Sometimes/ Never)

Después de usar el botón para llamar a la enfermera, ¿con qué frecuencia le atendían tan pronto como usted quería? (Nunca/ A veces/ Normalmente/ Siempre)

A response of "They never answered the call button without a long wait" contains two negative terms. These types of items are difficult to cognitively process and resulted in several confused English language respondents². This issue did not arise with the Spanish translation, as the double negative was avoided.

(3) At least three English language respondents to the above item indicated that "answering the call button" could mean a voice over an intercom system as well as a live person who comes to your room. This was not the intent of the item. This problem did not arise with Spanish language respondents. The use of "atendian," which means "attend to you," seems to have reduced or eliminated this problem.

Comprehension: Cross-cultural issues.

(1) There were a series of items that began with the word "Before/ Antes":

a. Before giving you any new medicine, how often did doctors, nurses, or other hospital staff tell you the name of the medicine?

b. ... tell you what the medicine was for?

c. ... ask you if you were taking any other medicines or supplements?

d. ... ask if you were allergic to any medicines?

e. ... describe possible side effects of the medicine in a way you could understand?

a. Antes de darle cualquier medicamento nuevo, ¿con qué frecuencia le dijeron los doctores, enfermeras u otro personal del hospital el nombre del medicamento?

b. ... ¿para qué era ese medicamento?

c. ... ¿si estaba tomando otros medicamentos, complementos o suplementos?

d. ... ¿si era alérgico(a) a algún medicamento?

e. ... ¿los efectos secundarios del medicamento en una forma que usted pudiera entender?

² Several alternative wordings that avoided the double negative were considered and rejected prior to cognitive testing. For example, instead of "without a long wait," "as soon as you needed" was considered but rejected since respondents couldn't assess need. And, "quickly" was rejected because its idiosyncratic meanings.

The intent of these items was to determine if these behaviors occurred in the time period immediately prior to administration of the medication. When these items were administered to English language respondents, “before” was interpreted differently for items c and d by many of the respondents. The respondents would retrieve experiences associated with being asked about other medicines and medicine allergies. These experiences occurred in physicians’ offices prior to surgery, as part of the pre-admittance paperwork, and as part of regular admissions paperwork, as well as during the period immediately prior to administration of the medication. They would typically include these experiences.

We did not detect such differences with Spanish language respondents. However, we are loathe to conclude that the implicit time frame evoked by “antes” was the same for all Spanish language respondents. The supplementary probes that were associated with item c focused on definitions of the term “suplementos” and the supplementary probes associated with item d focused on how often this occurred and whether the same people asked more than once. If explicit probes had been developed and administered, we suspect this issue would also have arisen for several of the Spanish language respondents.

(2) Several items asked about behaviors associated with preparations for care after the patient was discharged from the hospital. None of the English language respondents had difficulties understanding “discharged.” However, at least one Spanish language respondent did not know what “alta” meant.

(3) Certain demographic items were problematic for Spanish language respondents because of cultural differences. An item about education asked:

What is the highest grade or level of school that you have completed?

¿Cuál es el grado o nivel escolar más alto que ha completado?

- “ 8 años de escuela o menos*
- “ Algo de escuela secundaria, preparatoria, bachillerato, o high school, pero sin obtener el diploma*
- “ Graduado(a) de la escuela secundaria-- Diploma de escuela secundaria, preparatoria, bachillerato, high school o su equivalente (por ejemplo: GED)*
- “ Algunos cursos universitarios o un título universitario de un programa de 2 años (por ejemplo: AA, AS)*
- “ Título universitario de 4 años (por ejemplo: BA, AB, BS)*
- “ Estudios de postgrado o estudios superiores al título universitario de 4 años*

Since this was a standard demographic item, it was not administered in English. However, it was noted that at least two Spanish language respondents did not know what GED meant; one thought a masters degree was higher than a Ph.D.; one had never heard of

“preparatoria,” and another did not understand what “postgraduate studies” meant. Another Spanish language respondent, who was a mechanical engineering graduate of a California State University System school didn’t recognize “BA, AB, BS.” Familiarity with American educational terms seems to be a cultural issue. In addition, schools systems vary throughout Latin America, as do the names for comparable education levels. For example, secondary school might be referred to as escuela secundaria, escuela preparatoria, or escuela bachillerato. (In one country, a bachillerato meant high school, while in another country it meant a bachelor’s degree.)

(4) With one item, different problems were found in the English and Spanish versions. In the following question:

During the hospital stay, how often did doctors, nurses, or other hospital staff introduce themselves when they first came to care for you?

“other hospital staff” created confusion for some English language respondents. One respondent included cleaning staff as “other hospital staff,” which was not the intent of the item. (The item was intended to capture introductions from the care staff, who are expected to introduce themselves.) The failure to find the inclusion of care staff by the Spanish language respondents may be artifactual. In the Spanish cognitive interviews, additional probes focused on gaining insights about respondents’ understanding of the term “se le presentaban.” The probes were not as focused on “other hospital staff” or “first came to care for you” as in the English protocol.

In Spanish, different issues arose for the same item, which was translated as:

Mientras estuvo en el hospital esta vez, ¿con qué frecuencia se le presentaban los doctores, enfermeras u otro personal del hospital cuando le atendían por primera vez?

The phrase “se le presentaban” caused confusion with at least four respondents. Several respondents’ initial understanding of “se le presentaban” was that the item was asking about behaviors that occurred when the doctor or nurses were present. Although “se le presentaban” can be translated as “introduce themselves,” “presentaban” is similar to “presente” (present). This might explain the confusion.

Response formation differences: Cultural. There was evidence that some Spanish language respondents were using different criteria than most English-language respondents when responding to items.

(1) There is a tendency for all respondents, regardless of language, to evaluate their providers positively. For example, in response to:

During your hospital stay, how often did the doctors explain things about your illness or treatment in a way you could understand?

one English language patient answered, “Always,” in spite of the fact that there were 2 - 3 times when the doctors’ explanations were unclear.

(2) Even though most people have a tendency to evaluate their physicians’ performance in a positive light, this tendency is exacerbated by Spanish-language respondent’s greater tolerance for poor service. The following item, asking about doctors’ treatment of patients, seems to have been answered differently by Spanish language respondents:

During your hospital stay, how often did the doctors treat you with courtesy and respect?

Mientras estuvo en el hospital esta vez, ¿con qué frecuencia le trataban los doctores con cortesía y respeto?

Two Spanish-language respondents reported negative experiences with physicians. One patient spoke with 2 doctors, once each time, while in the hospital for 2 nights. The first doctor he spoke with treated him poorly, refusing to release the patient when the patient did not want to stay overnight. He said the doctor started to scream at him and he started to scream back. The patient spoke with another doctor, who released him. A second patient also reported a bad experience. However, both of them responded to the item with “Normalmente (usually).” It was surprising that these negative experiences resulted in a response as positive as *normalmente*; it is unlikely that many native, English speaking Americans would have responded as positively. It appears these respondents had a higher tolerance for poor service than the vast majority of English language respondents.

(3) Similarly, when these two Spanish-language respondents answered:

Using any number from 0 to 10 where 0 is the worst possible care and 10 is the best possible care, what number would you give the care you got from all the doctors who treated you during your hospital stay?

the first rated his care as an 8 and the second rated his doctor care as a 9. Based on their experiences with at least 30 English-language respondents, the interviewers did not feel that any of their English-language respondents would have given these doctors as high a rating.

One could hypothesize that the Latino respondents may think that providing a low provider rating could lead to interpersonal conflict either with the interviewer or subsequently with the provider if the responses were shared with him/her. (Trust in the confidentiality of the data may be another issue that can lead to providing more socially desirable responses.) Latino respondents

may be providing higher ratings in order to promote a smooth social relationship with the interviewer and to avoid interpersonal conflict with the provider.

(4) The Spanish language respondents that participated in the study appeared to have different expectations about health care, with respect to their involvement in making treatment decisions. In response to the following items:

During this hospital stay, how often did doctors, nurses or other hospital staff involve you in decisions about your treatment as much as you wanted?

Mientras estuvo en el hospital esta vez, ¿con qué frecuencia, le involucraban los doctores, enfermeras u otro personal del hospital, tanto como quería, al tomar decisiones sobre su tratamiento?

at least four of the Spanish language respondents insisted that the doctors know what is best for the patient and that the doctors should be the ones making these decisions. One said that he didn’t think he was qualified to give his opinion, since he didn’t know much about medicine. However, these beliefs might have been exacerbated by the translation of “involved” into Spanish. *Involucrado* has negative connotations, along the lines of being involved in an accident or involved in a crime.

(5) Further evidence of different expectations about involvement in decision making was noted in the following item, which was administered only to Spanish language respondents:

If you could change one thing about your care during this hospital stay, what would it be?

Si pudiera cambiar una cosa sobre el cuidado que recibió mientras estuvo en el hospital esta vez, ¿qué cambiaría?

At least one Spanish language respondent had trouble understanding this hypothetical item. He responded that he couldn’t change anything -- it was all out of his hands. This respondent did not feel empowered. It is not clear whether this is a cultural difference, a socioeconomic phenomenon, or idiosyncratic.

(6) Difficulties associated with the use of hypothetical questions were much more common among Spanish language respondents. Several of the probes that were used in our protocol were projective probes (i.e., hypothetical questions). English language respondents did not have much difficulty responding to these probes. However, interviewers from AIR, CSR, and RAND all noted that these probes were much more difficult to administer to the Spanish language respondents.

Conclusions

The application of cognitive interviewing techniques for testing questionnaire items in Spanish is a feasible and effective way to detect item problems. Most of the cognitive techniques (concurrent think-alouds, scripted comprehension probes, unscripted probes) worked well,

enabling detection of item problems. However, it was observed that projective probes did not perform well with the Spanish-language respondents.

Results from the Spanish cognitive interviewing were compared with results from the English cognitive interviewing. It was noted that:

(1) Many problems and issues were independent of language. The same medical care related terms were often sources of comprehension problems that were independent of language. Regardless of language, respondents had problems with verbal administration of items that exceeded their working memory capacity. They also had problems with “double negative” items. In responding to an item asking about problems associated with calling or visiting the hospital that friends or family might have had, both English and Spanish language respondents interpreted the absence of complaints as being indicative of the absence of problems.

(2) Certain problems arose in the English cognitive interviewing that did not arise in the Spanish cognitive interviewing. In response to questions about their hospital care, English language child birth patients would include the care and treatment their newborn when forming judgments. (The items were intended to refer to only the care the patient received.) This problem was avoided in Spanish, due to the fact that there are singular and plural forms of the word for “you.” In other cases, it was possible to eliminate problems by using words that more clearly expressed the construct of concern. For example, when asked about “answering the call button,” English language patients would include a verbal response over an intercom as an example of answering the call button. This was not the intent of the item. However, the use of “attendian,” which means “attend to you,” seems to have avoided this problem.

(3) Other issues unique to the Spanish language instrument were identified. A demographic item about the patient’s level of education did not work well for many Spanish language respondents. This reflects differences in the structure and nomenclature of educational systems in different Spanish-speaking countries. (These problems arose even though the items were translated into Spanish, with examples and attempts to compensate for such differences.)

Spanish language respondents had difficulties with the Federally mandated race question, which followed a question about whether or not the respondent was Hispanic or Latino. This difficulty was noted by the Bureau of the Census when similar items were administered in the 2000 Census (U.S. Dept. of Commerce, 2001).

Difficulties with a hypothetical question were noted for several Spanish language respondents. An item asking if the respondent could change one thing about their care during the hospital stay evoked a response that it was all

out of his hands. Similar difficulties were noted in the use of projective (hypothetical) probes.

(4) Differences were noted with respect to judgment formation and response synthesis strategies employed by Spanish and English language respondents. Some Spanish language respondents appeared to have different expectations about health care than English language respondents. They expected to have less involvement in making treatment decisions. We believe that several Spanish language respondents rated the care they received from physicians more positively than English speaking respondents would have rated comparable care.

References

Hofstede, G. (1980). *Culture's consequences*. Beverly Hills, CA: Sage.

Hui, C. H., & Triandis, H.C. (1989). *Effects of culture and response format on extreme response style*. *Journal of Cross-Cultural Psychology*, 20, 296-309.

Marin, G. & Triandis, H.C. (1985). Allocentrism as an important characteristic of the behavior of Latin Americans and Hispanics. In R. Diaz-Guerrero (Ed.), *Cross-cultural and national studies in social psychology* (pp. 85-104). Amsterdam: Elsevier Science Publishers.

Marin, G., Gamba, R.J., & Marin, B.V. (1991). *Acquiescence and extreme response sets among Hispanics: The role of acculturation and education*. *Journal of Cross-Cultural Psychology*.

Ross, C. E., & Mirowsky, J. (1984). *Socially-desirable response and acquiescence in a cross-cultural survey of mental health*. *Journal of Health and Social Behavior*, 25, 189-197.

Sudman, S., Bradburn, N., & Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass Publishers

Tourangeau, R. (1984). Cognitive science and survey methods: A Cognitive perspective. In: T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 73-101). Washington, DC: National Academy Press.

U.S. Department of Commerce, U.S. Census Bureau. (2001). *The Hispanic Population, Census 2000 Brief, May 2001*. C2KBR/01-3.

Weech-Maldonado, R., Morales, L, Elliott, M., Spritzer, K., Marshall, G., & Hays, R. (2003). *Race/Ethnicity, Language, and Patients' Assessments of Care in Medicaid Managed Care*. Health Services Research, 38(3), 789-808.