

The Use of Vignettes in Evaluating Multilingual Questionnaires¹

M. Mandy Sha¹, Yuling Pan, Ph.D.²

¹RTI International, 230 West Monroe St., Suite 2100, Chicago, IL 60606-4901

²Language and Measurement Research Group, Statistical Research Division, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract

Vignettes are hypothetical situations that can be used to review questionnaire items, whereas it would be cost prohibitive to recruit participants that can represent all the situations in the questionnaire. The vignette method has been shown effective in evaluating questionnaires in English and Spanish, but little research is available on its application to other languages. This paper attempts to fill in the knowledge gap by examining the use of vignettes to evaluate questionnaire translation in Chinese, Korean, Vietnamese, and Russian. This study is based on a cognitive testing project undertaken at the U.S. Census Bureau to pretest the translation of the 2010 Census form in multiple languages through 92 cognitive interviews. We analyzed findings from the age reporting instructions and the relationship terms across the four non-English languages and explored future research directions.

Key Words: Translation, cognitive interviewing, vignettes, language and culture, multilingual questionnaires, cross-cultural survey methods

1. Introduction

Cognitive interviews are often used to solicit information on issues and implications for questionnaire construction and evaluation. As part of the cognitive interviewing methodology, vignettes can be designed to review questionnaire items. Vignettes are hypothetical situations that are presented to cognitive interview participants. They can be a story, a scenario, or a description, and are in general short (Willis 2005).

Past studies have used vignettes to evaluate aspects of English language questionnaires. They found that vignettes yield useful information, especially when context and wording effects are measured. For example, vignettes could be used to investigate participants' survey-relevant judgments, such as how they interpret information provided to them (Bates & DeMaio, 1989; Gerber 1996), what the hidden assumptions are about certain concepts (Gerber 1994), and whether participants' interpretations match the survey definitions, as well as comparing alternative questionnaires (Martin & Polivka, 1995; Martin 2004).

¹ Disclaimer: This paper is released to inform interested parties of research and to encourage discussion of work in progress. Any views expressed on (statistical, methodological, technical, or operational) issues are those of the authors and not necessarily those of the U.S. Census Bureau.

Vignettes have also been used in evaluating questionnaires in Spanish. Recent studies suggested that vignettes can identify participant comprehension issues with Spanish translations (Caspar and Goerman, 2007; Goerman & Clifton, 2009). In addition, vignettes present an advantageous approach when it may be too demanding or cost prohibitive to have all participants review every question item or recruit participants with corresponding characteristics for testing skip patterns. However, written vignettes need to be adapted to include drawings and verbal descriptions in order to accommodate Spanish-speaking participants of lower educational levels (Goerman & Clifton, 2009).

In spite of these research efforts on the use of vignettes, we are not aware of prior literature on the application of vignettes to non-English and non-Spanish languages, or their effectiveness in evaluating multilingual questionnaires. In this study, we attempted to fill in this knowledge gap. We examined whether vignettes are effective in evaluating two questionnaire items translated in Chinese, Korean, Vietnamese, and Russian. These items are the age reporting and household relationship questions that may be used on the 2010 census form.

Effectiveness refers to whether vignettes can generate the right information to satisfy the research objectives of the evaluation. The objectives were to identify the appropriateness and accuracy of word choice and the grammar and naturalness in sentence construction. It also aimed to reveal issues related to sociocultural context that affects participant comprehension. Ultimately, translation issues are identified and corrected. To examine the effectiveness of the vignettes, we investigated: (1) whether participants were able to complete the vignette task by reviewing the hypothetical vignette situations and forming a reaction or response to the question; (2) findings from the followup probes that were designed to see how a participant reached an answer and whether that answer is intended by the survey; and (3) how the vignette task and its followup probes helped with identifying comprehension issues with the translation and any alternative translation terms that participants might have found easier to understand.

2. Methodology

For multilingual questionnaires that involve translations, cognitive interviewing can be used as a pretesting method in detecting translation issues and improve translation quality (Goerman et al., 2007; Pan et al., 2006; Pan et al., forthcoming). This method is recommended in Census Bureau Guideline for Translation (Pan & de la Puente, 2005; U.S. Census Bureau, 2004) and supported by recent studies that pretest translations of survey instruments (Carter et al. 2009; Pan et al. 2009). We used data from a cognitive testing project undertaken at the U.S. Census Bureau that was designed to pretest the Chinese, Korean, Vietnamese, and Russian translations of the 2010 mailout/mailback self-administered census form.

For each of the four target languages on this cognitive testing project, a team of language expert was assembled. Because a translated instrument is most likely used by participants with limited English proficiency or who are immigrants, cognitive interview participants were recruited among this population. Ninety-two interviews were conducted over two rounds of cognitive interviews and with native Chinese, Korean, Vietnamese, and Russian speakers who spoke little or no English. As a comparison, 16 cognitive interviews in English were conducted. Three sites were selected for the interviews: the greater Washington, DC area, Illinois, and North Carolina. These sites were selected

based on the geographic proximity of the interviewers and they also have sufficient concentrations of the populations needed to represent the four target language groups. Participant recruitment focused on including a diverse range of educational attainment, age and year came to the U.S., and country or region of birth.

Each cognitive interview was administered using a protocol, which documented the administration details, informed consent procedures, and a guide of cognitive interview probes. The participants were first asked to fill out the census form by themselves as if they received it at home via mail. They were then asked to read certain texts aloud to gauge their naturalness. Additional language-specific probes were developed to test participants' comprehension and reaction. In the second round of cognitive interviewing, we revised the interview protocol guide in order to compare the current translations on the census form with the alternative wordings that were developed by each team of language experts based on round 1 findings. In order to elicit participants' feedback and reaction to the changes that we recommended, show cards were used to compare different versions of the translations.

The census form tested is the “short form” questionnaire that contains questions on age, gender, name, Hispanic origin, race, tenure, and household enumeration. The language experts identified specific key terms that could pose comprehension issues. In particular, the question on age includes a “baby instruction” that says “please report babies as age 0 when the child is less than 1 year old.” In the household enumeration section, a list of 14 terms describes the relationship between a household member and “Person 1” (first adult listed on the form, usually the owner or renter). The baby age instruction and the terms used to describe different relationships might not be covered when discussing about the participants' own household. For example, participants may not carefully review a relationship term if it does not apply to their own household composition. It is challenging to verify this. Additionally, it is not cost effective or pragmatic to recruit multiple sets of monolingual participants and rare subgroups in the four non-English language groups (Chinese, Korean, Russian, and Vietnamese) who can represent all question topics. The multilingual and cross-cultural context affecting the understanding of different relationship terms was an additional challenge.

In order to test both age reporting instruction and relationship terms, we constructed hypothetical vignettes of household composition. Based on prior experience evaluating questionnaires with Spanish speakers (Caspar & Goerman 2007, Goerman et al. 2007), we learned that the vignettes should not contain too much text or details for participants to attend to. The literacy level of participants also affected the time needed for processing the hypothetical vignettes. Therefore, we focused on having the participants select the appropriate relationship and react to the baby instruction.

The interview data is documented in the individual interview summary reports that were written by the interviewers who conducted the cognitive interviews. These reports detailed participants' answers to the vignettes and if appropriate, interviewer's interpretation of those answers. As much as possible, they also included actual quotes, dialogue, and spontaneous probing. In addition, we also incorporated findings from RTI interviewer debriefings to inform the actual administration of the vignette task across languages.

3. Vignettes

Vignettes were constructed for the “baby instruction” of the age question, and for five terms that appear in the list of answer categories of the relationship question. They were administered in Chinese, Korean, Vietnamese, and Russian after the participants completed the translation of a census form in that language.

3.1 Vignette for Baby Instruction

The question about age repeats throughout the Census form for all named household members. As shown in Figure 1, an instruction appears in italic and immediately follows the question text: *Please report babies as age 0 when the child is less than 1 year old.* This “baby instruction” aims to prompt participants to put “0” as the age for a child who has not reached one year old.

7. What is Person 1’s age and what is Person 1’s date of birth?			
<i>Please report babies as age 0 when the child is less than 1 year old.</i>			
<i>Print numbers in boxes</i>			
Age on February 1, 2008	Month	Day	Year of birth
_____	_____	_____	_____

Figure 1: Baby Instruction at the Age Question

We wanted to study comprehension of the baby instruction. However, not all participants have a baby in the household and may not notice the instruction. Therefore, we used a vignette situation and a followup probe, as follows:

The instruction in Question 7 says: “Please report babies as age 0 when the child is less than 1 year old.” If a person has a 4-month old baby girl, what age should she write here?

3.2 Vignette for the Relationship Question

The question about relationship between household members and Person 1 repeats throughout the Census form for all named household members. A total of 14 relationship terms are provided as answer categories, including biological, adopted, and in-law relationships, as well as nonrelatives who may share the same household.

2. How is [this person] related to Person 1? Mark x ONE box.	
<input type="checkbox"/> Husband or wife	<input type="checkbox"/> Parent-in-law
<input type="checkbox"/> Biological son or daughter	<input type="checkbox"/> Son-in-law or daughter-in-law
<input type="checkbox"/> Adopted son or daughter	<input type="checkbox"/> Other relative
<input type="checkbox"/> Stepson or stepdaughter	<input type="checkbox"/> Roomer or boarder
<input type="checkbox"/> Brother or sister	<input type="checkbox"/> Housemate or roommate
<input type="checkbox"/> Father or mother	<input type="checkbox"/> Unmarried partner
<input type="checkbox"/> Grandchild	<input type="checkbox"/> Other nonrelative

Figure 2: Relationship Question and Answer Categories

As part of the protocol guide development, the Chinese, Korean, Vietnamese, and Russian language experts contemplated the impact of each answer category in context. In general, they agreed that biological and in-law relationships are commonly referred to for their target ethnic/cultural population and the translation for them is not ambiguous. These relationship terms include husband or wife, son or daughter, brother or sister, father or mother, grandchild, parent-in-law, and son-in-law or daughter-in-law. These individuals are also more likely to be household members. Therefore, participants with such household composition could be more frequently recruited and represented in the study.

However, adopted, step, and nonkin relationships may not be as common among the Chinese, Korean, Vietnamese, and Russian populations. They include: adopted son or daughter, stepson or stepdaughter, roomer or boarder, housemate or roommate, and unmarried partner. This posed several challenges: (1) It is not always feasible for the researcher to verify that the selected relationship answer category is as intended, and (2) recruiting for a large enough group of participants for each of these specific households may not be realistic.

Therefore, we created four vignettes to investigate comprehension of these terms.

The vignette for the relationship question started with providing the participants with an example vignette so that they can practice with performing this cognitive task. Then, they were directed to read the first vignette. Each of the four vignettes described a fictitious household. The names of the residents were first developed in English and then substituted by culturally-appropriate names in Chinese, Korean, Vietnamese, and Russian. After studying the description of a household in a vignette situation, participants had to identify how one resident was related to another by choosing from the same list of 14 available answer categories in the relationship question. Followup probes were asked to understand how participants reached their answers and what alternative translation might be more appropriate. In the second round, participants were asked to compare alternative terms listed on a show card.

3.2.1 Vignette 1: Relationship Term “Housemate or Roommate”

The first vignette is designed to test the term “housemate or roommate”, which is intended for those who share a dwelling and expenses. It says:

Mary lives in an apartment with her best friend Nancy. They share all their housing expenses equally. How is Mary related to Nancy?

3.2.2 Vignette 2: Relationship Term “Unmarried Partner”

The second vignette is designed to test the term “unmarried partner”, which is intended for those who live together in a marriage-like relationship but are not legally married. In the first round, the vignette described a couple living together for 8 years. Because 8 years is a longer period, the description was modified in the second round to 2 years in order to avoid the perception of a common-law marriage. This was the only vignette that changed between rounds. The vignettes say:

(Round 1) David and Kimberly have lived together as a couple for 8 years but have never been legally married. How is David related to Kimberly?

(Round 2) David and Kimberly have lived together as a couple for 2 years but have never been legally married. How is David related to Kimberly?

3.2.3 Vignette 3: Relationship Term “Roomer or Boarder”

The third vignette is designed to test the term “roomer or boarder”, which is intended to describe roomers or boarders who pay for occupying a room or receiving meals. It says:

Ana rents a room from Mrs. Doe. Ana also pays for two meals a day that Mrs. Doe cooks. How is Ana related to Mrs. Doe?

3.2.4 Vignette 4: Relationship Term “Stepson or Stepdaughter” and “Adopted Son or Daughter”

The final vignette is designed to test two terms: “stepson or stepdaughter” and “adopted son or daughter”. Both terms describe non-biological children. The former is a child of the spouse from a previous marriage and the latter describes a child who is adopted through legal proceedings. The vignettes² have two parts and two followup probes were administered after each part. They are:

Angelina and Brad are married and have three children. The oldest child is Joe. Brad’s ex-wife Jennifer had given birth to Joe. How is Joe related to Angelina?

The next child is Max. Angelina adopted Max five years ago. How is Max related to Angelina?

3.2.5 Followup Probes

For each vignette, interviewers administered up to four followup probes, as shown below. Because round 2 interviews incorporated findings about alternative translation from the first round, a round 1 probe was revised to ask participants to compare two or more translation to describe the relationship term.

- (Vignettes 1, 2, 3, and 4) You chose _____. How did you choose that answer?
- (Vignettes 1, 2, 3, and 4) What do you think they mean by [relationship term]?
- (Vignette 1, 3) Does [term 1] mean the same thing as [term 2] or are they different things to you?
- (Vignette 1, 2, 3)
 - (Round 1) Is there another term you would use to describe this type of relationship?
 - (Round 2) Which of the following terms would you use to describe this type of relationship?

4. Results and Discussion

Participants interviewed in Chinese, Korean, Vietnamese, and Russian languages were asked to complete the vignette task by considering the vignette situation and then providing an answer. The followup probe led participants to elaborate how they reached their answer and considered alternative translations.

Table 1: *Problems with Vignettes by Language Group* shows the type of problems encountered by participants at each vignette for the baby instruction and the relationship

² The names of the residents were substituted by culturally-appropriate names in Chinese, Korean, Vietnamese, and Russian.

question and by language. There are three types of problems: (1) completing the vignette task; (2) choosing the intended answer; and (3) identifying comprehension issues and cultural implications through followup probes.

<i>Vignettes</i>	<i>Chinese</i>	<i>Korean</i>	<i>Vietnamese</i>	<i>Russian</i>
<i>Baby Instruction</i>	2	None	1, 2	2
<i>Housemate or Roommate</i>	2	2	1, 2	None
<i>Unmarried Partner</i>	None	None	None but discomfort	1, 2
<i>Roomer or Boarder</i>	2	None	2	None
<i>Stepson / Adopted Son</i>	None	1, 2	1, 2	2

As shown in Table 1, most of the participants did not encounter major problems when completing the vignette task (problem #1), except those interviewed in Vietnamese. The vignette task requires participants to consider the hypothetical vignette situation and then form a reaction or a response. Failure to do so also led to the problem of not choosing an answer intended by the survey (problem #2). Among the three types of problems, not choosing the intended answer was the most frequently encountered problem and occurred at least half the time among all participants. No problems were detected in identifying comprehension issues and cultural implications through followup probes (problem #3).

Based on the interview data, we determined that the vignette method seems effective across the target language groups. As participants tackled the problems in Table 1, translation issues were revealed, as well as the related sociocultural context that has an effect on participant comprehension. Followup probes were crucial to understanding how participants arrived at their answers and whether translation can be corrected to aid comprehension. They were particularly important when participants were unable to complete the vignette task, or when they selected the wrong relationship code or did not provide an intended answer according to the baby instruction.

We observed several factors that affect the effectiveness of the vignette method. First, participants exhibited difficulties with the vignettes when the vignette descriptions did not conform to the cultural and linguistic expectations of the participants. For example, several Russian participants could not consider the term “unmarried partner” at all because the vignette used in the first round described an 8-year living arrangement between an unmarried couple. Those participants would describe that the couple was related to each other as husband and wife, because they perceived that 8 years of living together has qualified any couple to be in a common-law marriage. In the second round, we had to revise the vignette from 8 years to 2 years.

Second, vignettes were not as effective when they included too much details or texts for participants to attend to. Although the vignette description for “stepson or stepdaughter” was short, it involved four people: a husband, a wife, an ex-wife, and the husband’s biological son with the ex-wife. It aimed to ask about how a child was related to one of the parent, however, the description started with both parents as if they were the focal point of the description. Cognitive interview participants had to sort out the relationship between each character before they could choose an answer about how the son related to the wife. As a result, a number of participants were unable to complete the vignette task and required help from the interviewer. Some other participants changed their answers,

chose multiple relationship answer categories, or chose an answer based on how one parent is related to another. The difficulties exhibited by one of the Vietnamese participants were noted in interviewer notes, as follows:

The participant took quite some time to give an answer. At first, his response was that no answer category was listed for the relationship. He then said it should be “vợ bé” (second wife) but soon changed his mind to “con nuôi” (adopted son or daughter). After reading the vignette again, he considered “con riêng” (stepchild or stepdaughter) but still was not quite sure. Then after a while he said it could be “mẹ ghê” (stepmother). Because stepmother was not among one of the listed categories, he finally decided that no answer categories would fit.

Additionally, there was some evidence that the order of vignettes might affect participants’ reaction to the question item. For instance, some participants in the Vietnamese language interviews were interpreting the vignette for “roomer and boarder” based on the discussion about unmarried partner in the immediately preceding vignette. They noted that the vignette did not describe a marriage between the residents; therefore, one was not related to the other as “unmarried partner” but “nonrelative”.

Moreover, providing a stronger context of the item tested was necessary in order to solicit relevant responses from a vignette. In the baby instruction vignette, we asked how someone who has a 4-month-old baby would record the baby’s age. This seemed to have led many participants to simply indicate what they themselves would like to record for the baby’s age, but not what they would do according to the instructions. In order to solicit participants’ response in the context of the baby instruction, interviewers had to use a series of spontaneous probes to redirect the participants to the instruction. It would have been better to provide the instruction to be tested last to provide a stronger context.

Furthermore, vignettes were effective in eliciting sociocultural implications that influence respondent comprehension of translated questionnaires. When completing the vignette task for the baby instruction, some participants did not provide the intended answer, which was “0”. Through followup probing, we learned that they did know what the instruction asked but chose to ignore it. They simply did not agree with the baby instruction because it did not conform to their perception of common practice and cultural expectations. For example, several participants would write the age of a 4-month old baby in months because that is how they record age. Not providing the intended answer was most pronounced among Vietnamese participants. Even though they seemed to comprehend that the instruction was to write “0”, half of the participants chose to answer in months or “1” for a baby. It was because they felt that age could not possibly be zero and decided that they would record the baby’s age either in months or “1”.

As illustrated in the conversation below, a participant in the Chinese language interview would write the baby’s age in months or “1”. Through multiple iterations of followup and spontaneous probing, the participant demonstrated that she comprehended the baby instruction as intended, but would choose not to follow it. First, she felt that a baby’s age cannot be zero. She would record the baby’s age in months and also cited it as a common practice among her friends of other races. She further explained that babies are often considered 1-year-old according to Chinese customs.

Interviewer: [Administered Vignette]
 Participant: 4 months.
 Interviewer: 4 months?
 Participant: Age 0 makes no sense. It means baby is still in the womb, not 4-month-old.
 Interviewer: According to this instruction, if a person has a 4-month-old baby girl, what age should she write here?
 Participant: We should count months, not age. But in [Chinese-speaking country/region], we'd say this baby is 1 year old.
 Interviewer: What do you think they mean by "Please report babies as age 0 when the child is less than 1 year old"?
 Participant: But I have a few [other races] friends; they all say a baby is "3 months", "9 months" [Participant said these months in English]
 Interviewer: What does that instruction mean to you?
 Participant: Is it that they are not treating babies as a person that should be count in the census?
 Interviewer: In your own words, how would you say this instruction?
 Participant: It is saying that if a baby has not yet reached 1-year-old, then write 0 year.

There is an Asian age counting tradition that adds to a person's age by one at the advent of the lunar New Year. In other words, a person's age would always be "plus 1" (1 year older than the actual age since birth). Thus, babies are considered 1 year old at birth. Among participants in Asian language interviews, those with Korean descent discussed the traditional "plus 1" age counting convention more frequently than Chinese and Vietnamese participants, possibly because it is practiced in modern Korea³. We found that despite the "plus 1" age counting convention, the presence of the baby instruction was helpful in guiding Korean participants to provide the intended answer. For example, a Korean participant indicated that this instruction prompted him to answer "0" for a baby's age. Without it, he felt that immigrants from Korea would write a baby's age as 1-year-old on the census form even when the baby is just a few months old. He even thought that this baby instruction was placed especially for immigrants who might be more familiar with a different age counting system. Although this instruction is not intended only for immigrants but for all Census form takers, it helped this participant to provide an answer "0", because he was willing to observe it.

Finally, vignettes may be effective in evaluating multilingual questionnaires because of several advantages that they offer. They led participants to consider a term or a phrase in a hypothetical, less personal situation. For example, most Asian participants were able to choose the intended answer of "unmarried partner", despite acknowledging the negative connotation it carries. Another benefit of applying the vignette technique is allowing more participants to give feedback on terms and concepts that could otherwise be done only by a subset of hard-to-recruit participants who represent a specific question topic, such as in the case of the multiple relationship terms. The vignettes for household relationships were quite useful when the unintended relationship code was selected. For example, because the concept of "housemate or roommate" is not germane in the Chinese, Korean, and Vietnamese languages, the translation needed to be descriptive and indicative about the co-renting and cost-sharing aspects. Otherwise, participants showed

³ Age is reported as "plus 1" unless the words 만나이 (Man-nai) was stated explicitly.

that they would choose “roomer or boarder”, “unmarried partner”, or “other nonrelative” to describe a housemate or roommate. For “roomer or boarder”, we learned that the translation generally covered the renting aspect only but lacked the concept of room and board. Several Chinese participants chose “other relative” for the vignette household. We also learned that while adoption involves a legal process in the U.S., it could be an informal arrangement outside of the U.S. Specifying the legal aspect in the translation for “adopted son or daughter” was particularly important in Vietnamese and Russian languages.

5. Conclusion

This study showed that vignettes could be practicably used to evaluate multilingual questionnaires. They are effective in examining comprehension issues related to translation and the sociocultural context surrounding it. The majority of the 92 participants interviewed in Chinese, Korean, Vietnamese, and Russian could complete the vignette task by considering the hypothetical situation and then providing a response. Followup probes were essential in gathering substantive findings about how a participant reached an answer and what translation would be better understood. We also benefited from engaging language experts who had identified key comprehension issues that might surface with the target language groups.

Additionally, we learned that participants experienced problems with the vignettes when they include too much text or details, or when the descriptions do not conform to cultural and linguistic norms. The order and context in which the vignettes are constructed also affected their effectiveness.

The vignette technique has several limitations. First, as described earlier, the interview data came from the interview summary reports, rather than interview transcripts. Thus, it was not always possible to conduct a deeper analysis such as comparing cultural cues across languages or spontaneous probes that elicit the appropriate answers. Further research based on the transcripts will benefit this research. Second, because the primary purpose of the project was to detect potential translation issues, the vignette situations and followup probes were geared toward eliciting only comprehension issues regarding key terms and phrases. They also were used with two questions on the U.S. Census form, which are not long or complex compared to survey questions in general. We are not sure if the vignette and followup probing techniques we employed in our cognitive interviews will be applicable in testing a long and complicated questionnaire. Furthermore, because the vignette task was used to compensate for recruiting of rare household compositions in our target language groups, it was not designed as an experiment. Future research should investigate the feasibility of using vignettes across languages to examine additional issues related to participants’ survey-relevant decisions and in an experimental design.

References

- Bates, N., & DeMaio, T. (1989). Using cognitive research methods to improve the design of the decennial census form. Proceedings of the U.S. Bureau of the Census Annual Research Conference.
- Carter, G., McAvinchey, G., Reed, L.M., Rodriguez, S.M., Schoua-Glusberg, A., and Sha, M. (2009). Cognitive testing of Spanish language translation of selected questions in

the American Housing Survey (AHS) final report. Washington, D.C.: U.S. Census Bureau.

Caspar, R., & Goerman, P. (2007). The use of vignettes in questionnaire evaluation: An application from the 2010 Bilingual Census form. Paper presented at 62nd annual conference of the American Association for Public Opinion Research, Anaheim, CA.

Gerber E. (1994). The Language of Residence: Participant Understandings and Census Rules. Final Report of the Cognitive Study of Living Situations. U.S. Bureau of the Census, Center for Survey Methods Research.

Gerber, E., Wellens, T. and Keeley, C. (1996) Who lives here?: The Use of Vignettes in Household Roster Research. Proceedings of the Section on Survey Research Methods, Alexandria, VA: American Statistical Association.

Goerman, P., Caspar, R.A., Sha, M., McAvinchey, G., and Quiroz, R.S. (2007). Census bilingual questionnaire research final report for round 2. Study Series (SM #2007-27). Washington, D.C.: U.S. Census Bureau.

Goerman, P. & Clifton, M. (2009). Vignettes in Cross-Cultural Cognitive Testing: Adaptation for Spanish-Speaking Participants of Lower Educational Levels
Paper presented at 64th annual conference of the American Association for Public Opinion Research, Hollywood, FL.

Martin, E. (2004). "Vignettes and Participant Debriefing for Questionnaire Design and Evaluation." In Presser S., et al (eds.), *Methods for Testing and Evaluating Survey Questionnaires*. New York: John Wiley & Sons.

Martin, E., & Polivka, A. E. (1995). Diagnostics for redesigning survey questionnaires: Measuring work in the Current Population Survey. *Public Opinion Quarterly*, 59, 547-567.

Pan, Y. and de la Puente, M. (2005). "Census Bureau Guideline for the Translation of Data Collection Instruments and Supporting Materials: Documentation on How the Guideline was Developed." In Statistical Research Division's Research Report Series (Survey Methodology #2005-06). Washington, DC: U.S. Census Bureau.

Pan, Y., Hinsdale, M., Schoua-Glusberg, A., and Park, H. (2006). Cognitive Testing of Translations of ACS CAPI Materials in Multiple Languages. Study Series (SM #2006-9). Washington, DC: U.S. Census Bureau.

Pan, Y., Sha, M, Park, H., Schoua-Glusberg, A. (2009). 2010 Census Language Program: Pretesting of Census 2010 Questionnaire in Five Languages. Statistical Research Division Research Report Series (Survey Methodology #2009-01). U.S. Census Bureau.

Pan, Y., Landreth, A., Schoua-Glusberg, A., Hinsdale, M., and Park, H. (Forthcoming) "Cognitive Interviewing in Non-English Languages: A Cross-cultural Perspective." In Hardness, J., et al (eds.) *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. Berlin: Wiley Press.

U.S. Census Bureau (2004). Census Bureau guideline: Language translation of data collection instruments and supporting materials. Washington, D.C.: U.S. Census Bureau.

Willis, G. (2005). Cognitive interviewing: A tool for improving questionnaire design. Sage: Thousand Oaks, CA.