Adapting and Improving Methods to Manage Cognitive Pretesting of Multilingual Survey Instruments[1]

M. Mandy Sha[1], Yuling Pan[2], Barbara Lazirko[2]

[1]RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
[2]U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746

Abstract

Current methods to manage cognitive pretesting of multilingual survey materials and instruments are based on well documented lessons learned. Yet, they did not follow a systematic approach because translation pretesting was still at its early development stage. This paper examines the application of the methods in the current literature to a large-scale translation pretesting project, and recommends improvements and areas for future research. The U.S. Census Bureau pretested the Chinese and Korean translation of the American Community Survey (ACS) Language Assistance Guide (LAG), which included a variety of housing and population questions, such as mortgage, health, education, and ancestry. A total of 258 respondents who were 129 Chinese and 129 Korean speakers with limited English language proficiency participated in cognitive interviews. The large number of respondents was selected from completed screener records after contacting more than 1,000 candidates in three states. By following a systematic process and guided by sociolinguistic research framework, we were successful in adapting the methodology in the current literature, particularly with using cognitive interviews to identify translation problems at the pragmatic level that cannot be revealed in traditional translation review process. We also found implementing the following steps to be advantageous: conducting a systematic translation review process prior to cognitive testing, tailoring respondent recruitment methods to recruitment criteria, and implementing an interview lifecycle process to control and monitor the progress of interviewing.

Key Words: cognitive interviewing, pretesting, ACS, translation

1. Introduction

Managing cognitive pretesting involves many operational considerations. Willis (2005) referred to subject recruitment as an activity that "determines the feasibility of cognitive interviewing" and described how to determine respondent selection parameters and common recruitment methods. He also presented twelve logistical issues, including staffing, number of interviews, length of time, and various planning activities. When these considerations were applied to multilingual cognitive pretesting, Forsyth, Kudela, Levin, Lawrence, and Willis (2007) discussed five areas as part of the lessons learned after conducting a total of 32 cognitive interviews for Chinese, Korean, and Vietnamese translations of a survey questionnaire of tobacco use. Those areas were: interview protocol development, staffing "survey language consultants" and training "bilingual interviewers", respondent recruitment, interviewing, and reporting.

[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.

Pan, Landreth, Hinsdale, Park, and Schoua-Glusberg (2008) also used cognitive interviews to evaluate Spanish, Chinese, Korean, and Russian translations of survey prenotification materials for the American Community Survey (ACS). A total of 112 cognitive interviews were conducted across these four target languages and English. They recommended using a panel of language experts and survey methodologists, using culturally appropriate interview protocols and translating them, and providing a comprehensive interviewer training. More importantly, they identified a set of issues that are not normally informed by the traditional translation-review process. Pan, Sha, Park, and Schoua-Glusberg (2009) successfully applied the same methodology to pretest survey instruments. They conducted a total of 92 cognitive interviews in Chinese, Korean, Russian, and Vietnamese to evaluate translations of the 2010 Census form. To provide a comparative perspective, sixteen English language interviews were also conducted.

In a case study of pretesting of the U.S. Census Bureau bilingual Spanish/English questionnaire for the 2010 Census, Goerman and Caspar (2010) provided a list of issues that echoed the methods noted in prior literature about managing multilingual cognitive pretesting. Based on 85 Spanish language cognitive interviews and 25 English interviews, they discussed possible solutions about staffing qualified interviewers and training, developing suitable protocols, identifying monolingual respondents, and including appropriate details in the interview summaries. Sha, McAvinchey, Rodriguez, and Carter (2010) came to similar conclusions after conducting 96 cognitive interviews in Spanish and 32 in English. They reported that having formal interviewer trainings increase interviewer buy-in and compliance with the research protocol and procedures, and that respondent recruitment should follow a well-tested recruitment process while adapting to changing needs.

In summary, current methods to manage cognitive testing of multilingual translation can be described in five steps:

(1) Using and training language experts to conduct cognitive interviews
(2) Recruiting monolingual respondents of the target language, sometimes coupled with English interviews to anchor findings
(3) Developing protocol in English and target language
(4) Conducting interviews (in target language) for at least two rounds and writing interview summaries (in English with specific quotes in target language)
(5) Analyzing and reporting results

While prior studies have identified these main steps, they were unable to follow a systematic approach to manage those steps because translation pretesting was still at its early development stage. The study design and analysis were not well guided by an established theoretical framework and sometimes were conducted in an ad hoc manner. In addition, most of the recommendations were based on lessons learned from a "trial and error" format. For example, efficient methods to recruit monolingual respondents were not determined based on recruitment data but anecdotal information. What is more, past efforts were usually limited in the number of interviews conducted for each target language and the source language materials concentrated in single or more topics.

Our efforts to apply a systematic approach on a large-scale study offer a unique opportunity to further research in managing multilingual cognitive testing. This translation pretesting project was large-scale by several standards. First, a team of highly-

qualified interviewers completed 259 cognitive interviews in Chinese and Korean. Second, the survey instrument under testing included a large number of questions that covered a variety of topics. Third, we tailored efficient respondent recruitment methods and screened over 1,000 potential respondents. Next, we reported and analyzed results using a coding scheme that was grounded in the theoretical framework of sociolinguistics. In addition, the testing was a multi-year effort and project progress was closely monitored using a project management approach. By following a systematic process, we were able to fill in the knowledge gap by adapting and improving the current methods to manage multilingual cognitive testing.

2. ACS Language Assistance Guide (LAG) Translation Pretesting

Over a two-year period, RTI International (RTI) assisted the U.S. Census Bureau to pretest the Chinese and Korean translations of the American Community Survey (ACS) in the form of a self-administered language assistance guide (LAG). The LAGs had been translated by another contractor prior to cognitive testing. As shown in Table 2-1, a total of 1,084 potential respondents were screened from which 129 Chinese and 129 Korean speakers were selected to participate in cognitive interviews for a total of 258 cognitive interviews. A core team of eight language experts experienced with multilingual pretesting recruited the respondents, designed interview protocol guides in the target language, conducted cognitive interviews, and prepared interview summary reports.

Table 2-1. Results of Screening, Monolingual Eligibility, and Interview Selection

Language	*Screened*	*Monolingual Eligibility*	*Selected and Interviewed*
Chinese	404	351	129[2]
Korean	680	423	139
Total	1,084	774	258

The ACS asks about 48 population and 21 housing questions, including housing tenure, health and health insurance, ancestry, Industry/Occupation, education, income, mortgage, and a variety of other topics. The LAG provides translation for the English version of the ACS form that respondents use to complete the survey. The intent of the LAG is for actual respondents to read the questions, instructions, and answer categories in their language from the LAG and then record the answer on the English version of the form and mail the English version into the Census Bureau. For the purpose of the cognitive interview, pretest respondents wrote their answers directly on the LAG.

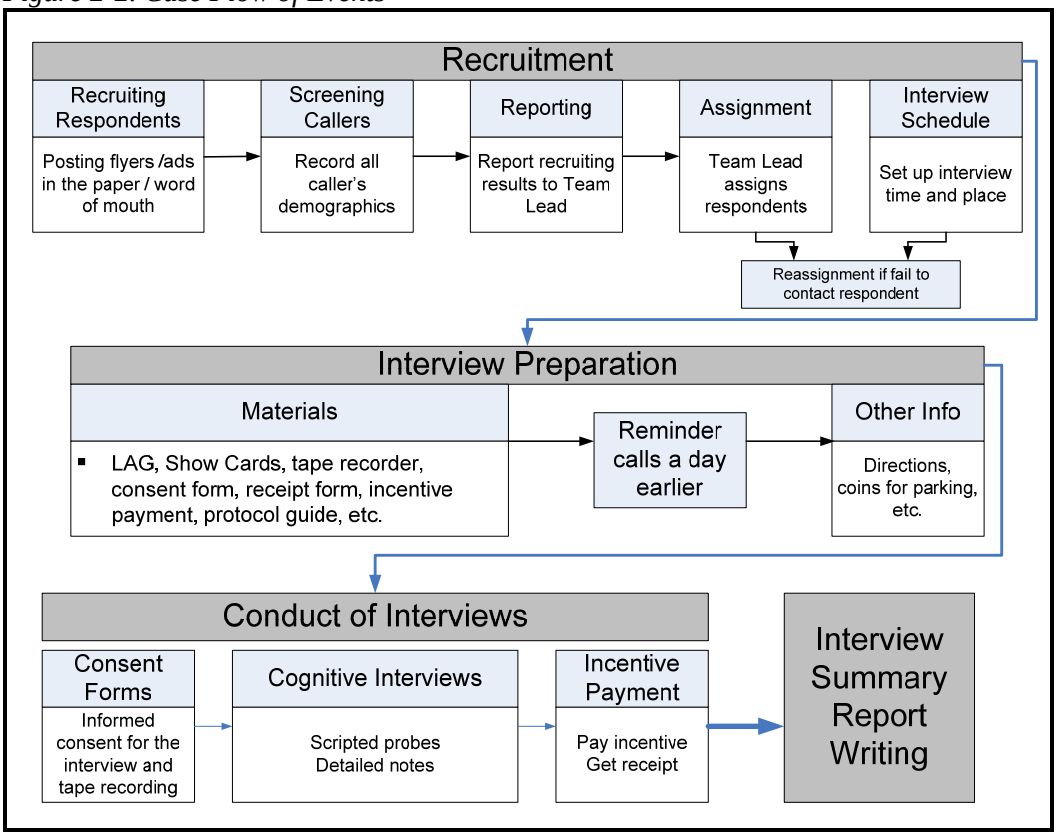
Because the entire LAG was too long for examination in one cognitive interview, a phased approach was designed. ACS questions were divided into three phases, and two rounds of testing took place during each phase. The first round of interviews focused on identifying problematic issues. In the second round, we tested the recommended changes based on findings from the first round. In each phase, two versions of interview protocols were developed to each cover a subset of the ACS questions. Because the ACS was a fixed instrument (the English source version did not allow for changes at the time), the translation pretesting protocol used mainly meaning-oriented probes to evaluate the translation.

[2] Ninety-one interviews were completed in Mandarin and 38 were in Cantonese.

During each cognitive interview, informed consent for participation was obtained before the interview started. If the respondents agreed to be audio recorded, their consent for recording was also obtained in writing. At the completion of the interview, respondents participated in a debriefing session with their interviewer to elicit additional feedback. After the debriefing, a $40 incentive payment was given to the respondents. The average time to complete an interview was 60 to 90 minutes.

This project was large-scale because of the high number of interviews, the large team of highly-qualified interviewers, and the complexity of the survey instrument being tested. It was also a multi-year effort. The RTI project manager took a project management approach to manage this study. As shown in Figure 2-2, an interview lifecycle consisted of respondent recruitment, interview preparation, conduct of cognitive interviews, and reporting of results. This was closely monitored.

Figure 2-2. Case Flow of Events



3. Adapting and Improving Current Methods

The following section describes our key efforts on this large-scale project to adapt and improve current methods to manage multilingual cognitive pretesting. The information is organized by the five main steps recommended in the current literature. It is important to note that we made continuous improvements on the methods based on regular debriefings among the language experts, results from the interviews, and the testing priorities communicated by the Census Bureau.

3.1 Maximizing benefits from prior studies in staffing and training interviewers
Because of our experience on prior multilingual studies, we already had a panel of three Chinese and three Korean language experts in place. Not only were they experienced with cognitive interviewing in the target languages, they were also bicultural and bilingual, having worked or studied in both their native and American cultures. However, in order to complete this many interviews, we knew that we had to train more interviewers. By the end of the data collection period, we used four interviewers for each of the languages to complete 20-24 interviews for a round in a 4-week[3].

The Chinese and Korean language expert panels were each led by a language expert who was a bilingual survey methodologist familiar with questionnaire design and multilingual cognitive testing. This was a rare combination of skills, and we were able to avoid some of the constraints described in previous studies, such as having to conduct interviewer trainings in English to accommodate research team members who do not speak the target languages. The leads also provided one-on-one coaching to the inexperienced interviewers we added.

There were a few months when the Chinese language lead oversaw the Korean language expert panel because its lead was unavailable. Even though the Chinese language lead did not speak Korean, she was sensitive to Korean language issues because of its sociolinguistic proximity to the Chinese language and culture. Based on the interview summaries (written in English with quotes in Korean), She was also able to make accurate assessment on the translation quality and clarify for the interviewers issues that resulted from misunderstanding of the intent of the questions. This was made possible because of a team approach – confirmations from the Korean language panel, appropriate details in the interview summary reports, this lead's background in questionnaire design, and ACS subject area expertise at the Census Bureau.

For interviewer training, we held a two-day in-person training for the entire team at the start of the project. At this training, most of the content was delivered in the common language, which was English. Because the vast majority of our interviewers were experienced, we were able to devote most of the time on detailed question-by-question walkthrough of the ACS and discussing the intent of the probes in the protocol guide. While we paired interviewers for practice interview, as recommended in the literature, after conducting six trainings, we found the "round-robin" style offered many benefits. In this setting, all trainees participate and in turn act as an interviewer with the team lead acting as the respondent. This style keeps the team communications intact, allows experience-sharing between trainees, and misunderstandings do not go unnoticed by the team lead and can be resolved immediately.

In subsequent rounds, trainings were one-day and conducted separately by language team to allow more synergy. These trainings were conducted using Chinese and Korean languages, which facilitated discussing language and culture-specific issues. After the trainings, the leads then convened to resolve issues that were raised during those trainings. All of these trainings were in-person or via videoconference that allowed for face to face interactions. The training for the first round focused on becoming familiar

[3] The Chinese language panel had a core team of four interviewers who completed 80% of the interviews. There were three additional language experts (two experienced and one novice cognitive interviewers) who temporarily replaced the core members at different periods during the project because of unexpected personal obligations.

with the ACS questions being tested in its English source language and its translation, as well as the probes. Practice interviews were conducted. Because the second round of interviewing was used mainly to resolve the problems identified previously, we felt that instead of practice interviews, the training should focus on reviewing possible changes to the translation and to the protocol to understand their impact. We also planned for remedial trainings after evaluating the first summary report written by the interviewers.

Our experience has been that the desired level of details in the interview summary reports is something that is difficult to convey. Thus, the trainings provided very specific instructions on writing interview summary reports. We provided both "good" and "bad" examples of reports to demonstrate the level of details needed. We also implemented precoded answer categories for the interviewers to check off, such as whether respondents answered the question or followed skip instructions as intended. This helped to create more context for the analysts who review these summaries and also helped to organize the large amount of qualitative data collected.

3.2 Identifying and recruiting monolingual respondents using systematic recruitment methods

The large number of interviews required a continuous respondent recruitment effort. The recruitment period spanned from December 2009 to July 2011, with recruitment and screening activities peaking before and during the 4-week active interviewing period for each of the two rounds in three phases. Respondents were recruited from three locations: Illinois, North Carolina, and the District of Columbia/Virginia/ Maryland metropolitan areas. These three sites were determined based on the geographic proximity of the language experts. But they also have sufficient concentrations of Chinese- and Korean-speaking populations to represent the target language groups, and this site selection approach helped to control overall costs.

The recruitment target determined by the Census Bureau was a combination of individuals whose characteristics fall into the desired categories that represent a great cross-section of monolingual Chinese and Korean speaking population in the United States. It included recruiting Chinese- and Korean-speaking respondents of various education levels (i.e., less than high school, high school graduates, and college graduates), of both genders, and from a variety of age groups. To be eligible for recruitment, the majority of the respondents needed to have lived in United States for at least a year. For Chinese speakers, the recruitment plan also included populations from diverse origins: those who were born in mainland China, Taiwan, or United States or other countries and those who spoke Mandarin and Cantonese dialects.

During the study, additional requirements about respondent recruitment were communicated by the Census Bureau to test specific questions. These requirements considered the sociocultural background of the population being recruited. For example, the ACS contains three questions about military service. Because English speaking and writing are required to serve in the U.S. military, we did not actively attempt to recruit monolingual respondents who were on active duty. In consultation with the Census Bureau, we instead recruited family members of current or past U.S. military service persons. This way, these questions could be tested with those who were more likely to be familiar with specific military terminology, since they were the intended users of the ACS LAG as well.

In addition, we used a combination of recruiting methods that had been effective in the past with similar target populations: leveraging community organizations, posting flyers at a variety of public places frequented by potential respondents, and issuing advertisements in ethnic newspapers and media outlets. About halfway through the project, we conducted a preliminary investigation of the efficiency of the recruitment methods. Liu, Park, and Sha (2012) learned that advertisement in local ethnic, in-language newspapers demonstrated a prominent advantage in reaching potential respondents in a short time. Word-of-mouth was a helpful strategy for recruiting hard-to-reach respondents meeting specific study requirements. These data-driven findings guided our recruitment strategies for the rest of the study. At the end of the recruitment, we were successful in fulfilling the recruitment criteria set forth by the Census Bureau and we screened over 1,000 potential respondents.

3.3 Tailoring interview protocol guide to translation pretesting

Based on lessons learned from Phase 1, we administered concurrent probing in round 1 and retrospective probing in round 2 of the subsequent phases. In the first round, we divided the LAG into a set of questions with one theme and had respondents answer one set of questions at a time followed by concurrent probing. This administration approach allowed better control of the length and flow of the interview, and helped identify issues the respondents had when they were actually answering the questions compared to the retrospective probing approach. Throughout the interview, the interviewer applied spontaneous probes to follow up on any observed problems, issues, or questions brought up by the respondent. In round 2, we used the retrospective probing approach—once the respondent finished filling out the section of the LAG being tested, the interviewer consulted the interview protocol, administered the scripted and spontaneous probes, and discussed the specific statements, phrases, or terms with the respondent question by question. Because the second round was used mainly to test recommended changes, having respondents finished filling out the section of the LAG being tested in one sitting was desirable.

Before the cognitive interview began, the protocol guide instructed the interviewers to provide a practice session to the respondents. We adapted the practice question about counting windows described in Willis (2005) and Goerman (2006) and had respondents answer “How many windows are there in the house or apartment where you live?” However, we quickly realized that this was not an ideal practice question for this type of translation pretesting. The “window” practice question was originally designed to induce the respondent to think aloud, while most of the cognitive tasks in our interview involved meaning-oriented probes designed to evaluate whether respondents comprehended the translation. Other types of probes that could improve questionnaire design were not used as frequently mainly because the source questionnaire could not be changed, a common constraint faced by many studies. In subsequent rounds, we used variations of the window question for the practice session. Although these attempts were not designed as experiments, we determined that the practice session must be kept simple, natural, and tailored to translation pretesting. For example, using the “window” practice question was problematic for Korean language interviews because “window” was understood phonetically to mean Microsoft Windows. This led to unnecessary confusion during the practice session and achieved the opposite of what a practice was intended.

The latest version of practice question included an easy-to-spot translation issue. In general, respondents more readily pointed out the obvious error. This allowed the interviewer to naturally transition to administering some of the meaning-oriented probes

and engaged the respondents in a simulation of the larger interview to come. However, this latest version of the practice question still had some limitations. We are in the process of analyzing the data from the interview summary reports and plan to report our findings.

3.4 Using consistent method to report results

The LAGs had been translated into Chinese and Korean by another contractor. Prior to cognitive testing, the panel of Chinese and Korean language experts conducted a detailed Expert review of the translation. After the review, they provided comments, corrections, and suggestions on the translations in addition to justification for the suggested corrections. Because of the advance translation expert review step, few translation errors were identified during pretesting. This allowed the testing to focus on uncovering linguistic and sociocultural issues that may only be observed through testing with monolingual respondents.

We compared two methods for conducting translation reviews: (1) a translation appraisal system (TAS) developed by Sha, Park, and Pan (2010)[4] and (2) the traditional expert review. Half of the panel was assigned to one of the methods and the other half was assigned to the other. The TAS group examined each translated sentence by considering preidentified specific categories of translation issues in a step-wise fashion and decided whether a translated item exhibited features that were likely to cause problems at each step. Those in the Expert Review Group did not follow such a system. Sha, Park, and Pan (2010) found that when aided with TAS, less experienced language experts could do as good of a job (or better in some aspects of the process) as the more experienced language experts of identifying translation issues during the review. Major benefits of the TAS were that it was a systematic process that prompted reviewers to consider issues at each step and problem. It also provided documentation and problem tally, and was relatively low cost.

Using a coding scheme guided by sociolinguistic approaches to language and culture (Pan and Fond 2010, 2011), we were able to evaluate and clearly communicate the results by classifying translation issues in terms of Linguistic Rules, Cultural Norms, and Social Practices. This allowed us to compare across the Chinese and Korean language interview results, especially when they did not appear to be comparable. We provide examples below from actual ACS questions and the type of issues we encountered.

<u>Linguistic rules</u>

- Complex sentence structure that did not reflect Chinese and Korean linguistic practice
- Instances where a single concept in English needed to be expressed in two words or phrases in Chinese or Korean. For example, for the housing question asking about the number of rooms, the word “room” in English can indicate many rooms in a residence, such as a living room, a bedroom, or a dining room. However, depending on the usage, each room has its own name in Chinese: 厅 (ting), 堂 (tang), 房 (fang) or 室 (shi).

[4] The TAS provides a systematic appraisal of translation and helps spot potential problems in the wording (lexical error) and structure (syntactical error) of the questions that may lead to difficulties in respondent comprehension, miscommunication, or other failings.

- Korean translation used difficult Hancha-rooted words or phonetic expression that cannot be readily understood

Cultural norms

- Ways of communication. For example, Chinese way of responding is to have the response follow the verb used in the question.
- Reasoning and cultural conventions, such as confusion and negative reactions toward the "when did you last get married" question because multiple marriages are traditionally not desirable in the Asian cultures.

Social practice

- No equivalent concept, such as the educational system
- Lack of knowledge of financial terms and U.S. government programs
- Inexperience with survey conventions, such as questionnaire layout and routing instructions, concept of an LAG, and "Surveyspeak" (Harkness 2003), e.g. repeating "house, apartment, and mobile home" in every question.

3. Discussion and Future Research

This large-scale translation pretesting effort successfully implemented the current methods in the literature, but we had to make adaptations to increase their "fit for use". We also identified several areas for improvement.

In terms of interviewers, we recommend using experienced cognitive interviewers who are language experts in the target languages. Recent interest in multilingual pretesting has produced a (small) cadre of qualified candidates. However, language experts experienced with cognitive testing who also have expertise in questionnaire design is still a rare combination. Such individuals, as used in our study, can lead the panel of experts and avoid some of the constraints experienced by previous studies where the lead researcher's native language was English. In addition, our experience suggests that it is possible to use a lead researcher who is *not* proficient in the target language – knowing the intent of the questions being tested and having some language sensitivity, can help to mend the gap with the help of language experts. When training interviewers, we suggest using the round-robin technique for practice interviews to allow immediate feedback and coaching. But we feel that they can be optional in the training for the second round of interviewing. Instead, we recommend using the time to discuss the changes since the first round. We also recommend providing examples of both good and poor interview summary reports to help the interviewers learn the level of details desired for the study.

By analyzing the efficiency of common recruitment methods, our data-driven approach to using recruitment methods guided us to tailor our efforts to recruiting different kinds of respondents. For example, we used advertisements in local, ethnic in-language newspaper when we needed to reach a larger pool of potential respondents quickly. Because of the large number of interviews, the recruiting and screening efforts were extensive and afforded us the opportunity to study the efficiency of recruitment methods for reaching monolingual respondents. To continue this line of research, future large-scale multilingual cognitive testing studies should give collecting and managing recruitment data a similar level of attention as the interview data.

When interview protocol is concerned, we preferred the concurrent probing technique in the first round that is intended to detect translation problems. It allowed respondents to report their observations while they were answering the survey. The level of interruption to the respondent's question-answer process seemed acceptable to us. For the second round where recommended changes to the translation were tested, we found retrospective probing better suited as it allowed respondents to finish filling out the LAG in one sitting before doing cognitive tasks. In addition, we had to modify the practice questions several times so that they are specific to the project and the types of probes in the interview protocols. We are not aware of prior research that demonstrates the utility of practice sessions for non-English cognitive interviews. This is clearly a topic for future research and we are in the process of completing our analysis.

In addition, when the translation and pretesting teams are different, we found it advantageous to implement a systematic translation review process, such as the TAS used on this study. By fixing language-specific errors prior to the cognitive interview, we were able to focus our attention on linguistic and sociocultural issues during the testing. The cognitive interviews identified a set of translation issues that may only be observed through testing with monolingual respondents. The coding scheme developed by Pan and Fond (2010, 2011) articulated what those issues were and the root causes. We recommend both processes for future large-scale or small-scale multilingual pretesting studies. Nevertheless, some respondent difficulty simply cannot be "fixed" within the parameters of the translation but must be addressed at the source language questionnaire level.

Finally, all research studies must balance the constraints of schedule, budget, and scope, even when aided with experts and expertise. For the ACS LAG study, the RTI project manager used a project management approach and implemented an interview lifecycle to control and monitor the progress of interviewing. Despite the large number of interviews and interviewers, multi-year schedule, and the complexity of the survey instrument being tested, we completed the study on time, within budget, and met all the requirements set forth by the Census Bureau. We recommend that future multilingual cognitive testing projects adopt a similar model.

Acknowledgements

The authors thank Yuling Pan who conceptualized the research framework for the ACS LAG study. We also thank the American Community Survey Office (ACSO) of the U.S. Census Bureau for advising this study: Barbara Lazirko, Todd Hugues, Herman Alvarado, Dameka Reese, and Debbie Klein. We acknowledge RTI project manager and lead researcher Mandy Sha. We also thank RTI survey methodologist Hyunjoo Park; this project would not have been successful without her. In addition, we acknowledge RTI's panel of language experts and cognitive interviewers: Lu Liu, Jiyoung Son, Qinghui Guo, Grace Chan, Suzie Kim, Younhee Harm, Michelle Yuan, Guangya Liu, Hyunjoo Park, Mandy Sha, and former Census Bureau analyst Virginia Wake.

References

Forsyth, B. H., Kudela, M. S., Levin, K., Lawrence, D., & G. B. Willis, G. B. (2007). Methods for Translating an English-Language Survey Questionnaire on Tobacco Use into Mandarin, Cantonese, Korean, and Vietnamese. *Field Methods 19*, 264-283.

Goerman, P. (2006). *Adapting cognitive interviews for use in pretesting Spanish language instruments*. Washington, DC: Statistical Research Division, U.S. Census Bureau.

Goerman, P.L. and Caspar, R.A. (2010). Managing the Cognitive Pretesting of Multilingual Survey Instruments: A Case Study of Pretesting of the U.S. Census Bureau Bilingual Spanish/English Questionnaire. In *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* (pp. 75–90). Hoboken, NJ: Wiley and Sons, Inc.

Harkness, J. A., Mohler, P. P., & van de Vijver, F. J. R. (2003). Questionnaire Translation. In J. A. Harkness, F. J. R. van de Vivjer, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 3-16). Hoboken, NJ: John Wiley & Sons.

Liu, L., Sha, M. & Park, H. (2012). *Efficiency of recruitment strategies to reach monolingual Asian Participants for Cognitive Interviews*. In JSM Proceedings, 67th Annual Conference of the American Association for Public Opinion Research (AAPOR), Alexandra, VA: American Statistical Association, pp. TBD.

Pan, Y., Landreth, A., Hinsdale, M., Park, H. & Schoua-Glusberg, A. (2008). *Methodology for Cognitive Testing of Translations in Multiple Languages.* Statistical Research Division Research Report Series (Survey Methodology #2008-02). U.S. Census Bureau. Available online at: http://www.census.gov/srd/papers/pdf/ssm2008-02.pdf.

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. Available at http://www.census.gov/srd/papers/pdf/rsm2009-01.pdf.

Pan, Y., & Fond, M. (2010). Developing Multilingual Questionnaires: A Sociolinguistic Perspective. Proceedings of the Joint Statistical Meetings, 2197-2210. Alexandria, VA: American Statistical Association.

Pan, Y., & Fond, M. (2011). Evaluating Multilingual Questionnaires: A Sociolinguistic Perspective. Research and Methodology Directorate, Center for Survey Measurement Study Series (Survey Methodology #2012-04). U.S. Census Bureau. Available online at http://www.census.gov/srd/papers/pdf/ssm2012-04.pdf.

Sha, M.**,** McAvinchey G., Rodriguez S., & G. Carter. (2010). *Respondent Recruitment, Interviewing and Training: Lessons Learned from a Spanish Language Cognitive Interviewing Project*. In JSM Proceedings, 65th Annual Conference of the American Association for Public Opinion Research (AAPOR), Alexandra, VA: American Statistical Association, pp. 6372-6381.

Sha, M., Park, H., & Pan, Y. (2010, May). *Developing a systematic process for translation expert review: The Translation Appraisal System (TAS-10)*. Paper presented at the 65th annual conference of the American Association for Public Opinion Research, Chicago, IL.

Sha, M., Park, H., & Pan, Y. (2012, March). *Translation Review and Cognitive Testing of American Community Survey (ACS) Language Assistance Guides in Multiple Languages*. Report prepared for the U. S. Census Bureau.

Willis, G. B. (2005). *Cognitive interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage.