

Field Methods

<http://fmx.sagepub.com>

Methods for Translating an English-Language Survey Questionnaire on Tobacco Use into Mandarin, Cantonese, Korean, and Vietnamese

Barbara H. Forsyth, Martha Stapleton Kudela, Kerry Levin, Deirdre Lawrence and Gordon B. Willis

Field Methods 2007; 19; 264
DOI: 10.1177/1525822X07302105

The online version of this article can be found at:
<http://fmx.sagepub.com/cgi/content/abstract/19/3/264>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *Field Methods* can be found at:

Email Alerts: <http://fmx.sagepub.com/cgi/alerts>

Subscriptions: <http://fmx.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 8 articles hosted on the SAGE Journals Online and HighWire Press platforms):
<http://fmx.sagepub.com/cgi/content/refs/19/3/264>

Methods for Translating an English-Language Survey Questionnaire on Tobacco Use into Mandarin, Cantonese, Korean, and Vietnamese

BARBARA H. FORSYTH
MARTHA STAPLETON KUDELA
KERRY LEVIN
Westat, Rockville, Maryland
DEIRDRE LAWRENCE
GORDON B. WILLIS
National Cancer Institute

This article reports research on procedures for translating a survey questionnaire on tobacco use from English into Mandarin Chinese, Cantonese Chinese, Korean, and Vietnamese. The goal is to offer practical guidelines for researchers involved in translating questionnaires. The authors operationalize a five-step process for translation and evaluation based on the frameworks presented in Harkness, Van de Vijver, and Mohler (2003) and the U.S. Census Bureau (2004). Based on qualitative observations, the five-step process produced effective questionnaire translations. The iterative nature of the process and the team-based approach the process encourages were particularly important to the success. Based on documented experiences, the authors identify lessons learned and make recommendations to other researchers who need to translate questionnaires.

Keywords: *survey translation; translation evaluation; TRAPD translation framework; survey methods; tobacco use survey*

Reliance on standardized questions is one feature that distinguishes survey methods from other field methods (Turner and Martin 1984; Groves 1989; Schaeffer and Maynard 2001). Survey methodologists generally agree that standardized question meaning is essential to sound survey measurement. The task of ensuring a standard meaning is complex when surveys are administered

Field Methods, Vol. 19, No. 3, August 2007 264–283

DOI: 10.1177/1525822X07302105

© 2007 Sage Publications

in a single language. When surveys are administered in multiple languages, the complexity is magnified because meanings must be standard not only within languages but also across them.

This article describes a process for developing and evaluating questionnaire translations for surveys administered in multiple languages. We focus on work for the National Cancer Institute, translating the English-language version of the Tobacco Use Survey component of the Current Population Survey (TUS-CPS) into Mandarin, Cantonese, Korean, and Vietnamese. We used a team-based process that involved five steps. This article describes that process and lessons learned. We hope that researchers can use the descriptions provided here as a practical guide for translating other survey instruments used in cross-cultural or multilingual research.

The TUS-CPS is a good candidate for translation for three reasons. First, there is concern that current English-language surveys of tobacco use underestimate tobacco use in some Asian American subgroups (e.g., Jenkins et al. 1995; Wewers et al. 1995; Ma et al. 2002; Lew and Tanjasiri 2003). Asian-language translations could make national survey results more representative and more accurate. Second, the TUS-CPS is widely used. It has been administered as part of the Census Bureau's Current Population Survey periodically since 1992, and it is a key source of national- and state-level data on smoking and other tobacco use in the U.S. household population. Third, the TUS-CPS questionnaire contains standard tobacco use prevalence items that are well established. English-language estimates from the TUS-CPS are consistent with the results of other large-scale tobacco surveys and therefore display construct validity. For example, estimates of current smoking in the United States in 2003 were 18.4% and 21.6% for the TUS-CPS and the National Health Interview Survey, respectively (Centers for Disease Control and Prevention 2005; National Cancer Institute 2006). In addition, research supports the notion that self-reported tobacco use is truthful (Patrick et al. 1994; Riesch et al. 2006).

OVERVIEW OF TRANSLATION PRACTICES

Researchers working on translation methods generally agree that an effective survey questionnaire translation is one in which the questionnaire translated into the target language "asks the same questions and offers the

We are grateful to Anne Brown Rodgers and three anonymous reviewers for their comments on earlier drafts of this article. This project was approved by the Westat Institutional Review Board, and all interview participants provided informed consent before beginning interviews.

same response options” as those provided in the source questionnaire (Harkness and Schoua-Glusberg 1998:92). Achieving this goal can be complicated for two reasons. First, words are not the only conveyors of meaning in language. Grammatical structures, syntax, and tone can also convey meaning, as can the communication context (Weidmer 1994; Pan 2003; Harkness, Pennell, and Schoua-Glusberg 2004). Second, complex mappings between words in different languages are routine. For example, words in the source and target languages that refer to similar concepts may have different connotations, different nuances of meaning, or different emotional effects (Harkness and Schoua-Glusberg 1998; Harkness, Pennell, and Schoua-Glusberg 2004).

Survey researchers have advocated diverse practices for producing effective translations (e.g., Brislin 1970; McKay et al. 1996; Harkness and Schoua-Glusberg 1998; Harkness, Van de Vijver, and Mohler 2003; U.S. Census Bureau 2004). Harkness, Van de Vijver, and Mohler (2003) provide detailed descriptions of these practices. Current practices build on the multi-step translation and evaluation approach articulated by Brislin (1970), which includes using back-translation to evaluate survey translations, bilingual respondents to test source and translated target-language questionnaires, and monolingual respondents to pretest target-language translations. More recent developments substitute alternative evaluation methods for back-translation, use bilingual reviewers and testers in more diverse ways, and use a broader set of pretest methods. Current practices also differ from previous methods in the numbers of translators involved, how translation work is distributed across individual translators, how review and revision activities are built into the translation process, and the roles bilingual staff (including interviewers) and/or members of the survey population play in translation and evaluation processes (Harkness, Van de Vijver, and Mohler 2003).

A growing consensus advocates team-based approaches to translation and evaluation. Compared with direct translation by one person, team approaches generate more translation options and provide sounder and less idiosyncratic translation review and evaluation (Harkness and Schoua-Glusberg 1998; European Social Survey 2002; Harkness, Van de Vijver, and Mohler 2003; U.S. Census Bureau 2004). Based on this general methodological preference, we chose to develop an iterative, team-based process. We began with the five-stage TRAPD framework described by Harkness, Van de Vijver, and Mohler (2003). The TRAPD framework identifies five general stages for developing and evaluating a questionnaire translation: translation (T), review (R), adjudication (A), pretesting (P), and documentation (D). Each of these stages can be implemented in a variety of ways, so the TRAPD framework

is very flexible. The five-step process we developed reflects one way of interpreting and operationalizing the TRAPD framework.

TRANSLATION AND EVALUATION PROCESSES

This section describes the process we used to translate the tobacco use questionnaire into Mandarin, Cantonese, Korean, and Vietnamese (see Table 1). We describe the research team, the procedures we used to translate the English-language TUS-CPS questionnaire, and lessons learned at each step. The lessons learned describe both what we found useful and what we would do differently the next time. They evolved from a variety of sources, including observed translator performance, comments and feedback from independent reviewers, and unexpected events.

The TUS translation and evaluation team included public health researchers at the National Cancer Institute familiar with tobacco research, methodologists at Westat familiar with questionnaire design, and experienced translators.

Step 1: Translation

The translation staff consisted of three independent professional translators. Each translator spoke one of the target languages as their strongest language. One translator was multilingual in English, Mandarin, and Cantonese; one was bilingual in English and Korean; and one was bilingual in English and Vietnamese. All three translators had at least 10 years of professional translation experience. Their experience translating survey questionnaires varied.

A translation coordinator worked with research staff and the translators to define the task. The project aimed to develop translations that “ask the same questions” (e.g., Harkness, Van de Vijver, and Mohler 2003). Based on these instructions, the three translators worked independently to produce target-language translations. The translation coordinator supervised their work and was available as needed to answer questions and provide guidance. However, results from subsequent review and cognitive interviewing activities suggested that these instructions were insufficient, often producing literal, word-for-word translations that were awkward, confusing, or cumbersome.

The translation step yielded three draft questionnaires: a single target questionnaire for the Mandarin and Cantonese dialects of Chinese, a target questionnaire in Korean, and a target questionnaire in Vietnamese. We opted to use a single Chinese-language questionnaire based on advice from the Mandarin and Cantonese translator who believed a single written questionnaire adequately conveyed the survey questions in both dialects. In

TABLE I
Five-Step Process for Translating the English-Language Tobacco Use Survey Component of the Current Population Survey Questionnaire

<i>Step</i>	<i>Activities</i>	<i>Personnel</i>
1. Translation	Create draft target-language questionnaires by translating source questionnaire into target languages. Document translation challenges and decisions on how to deal with them.	Professional translators
2. Review	Review translation materials, including source questionnaire, draft versions of target-language questionnaires, and translator documentation. Identify potential revisions.	SLCs having extensive language skills, knowledge of cultural groups, and familiarity with survey objectives
3. Initial adjudication	Review and decide on suggested revisions. Collaborate with bilingual reviewers and generally monolingual English-speaking research staff to develop pretest-ready target-language questionnaires.	Bilingual adjudicators (university consultants); project staff; SLCs
4. Cognitive interview pretest	Identify and train cognitive interviewers. Recruit interview participants. Complete cognitive interviews, analyze results, and develop recommendations.	SLCs, bilingual interviewers trained in cognitive interviewing, and English-speaking project staff
5. Final review and adjudication	Review recommendations with SLCs and English-speaking research staff. Select final revisions for target-language questionnaires.	Project director, advised by SLCs and English-speaking project staff

NOTE: SLCs = survey language consultants.

subsequent steps, both the Chinese survey language consultant (SLC) and the Chinese-language adjudicator agreed with this decision.

In addition to the three target-language questionnaires, translators provided documentation that described specific translation challenges they encountered and decisions they made to deal with these challenges. These translation products informed the succeeding review step.

Because of the way the current research was funded, the translation step was independent of the review, adjudication, and pretesting steps. We would prefer to instead follow advice from translation researchers and set up collaborative working relations among translators, reviewers, and adjudicators early in the translation step (Harkness and Schoua-Glusberg 1998; Harkness, Van de Vijver, and Mohler 2003; U.S. Census Bureau 2004). Fortunately, we were able to implement more collaborative approaches after the initial translation step was finished.

Lessons Learned: Translation (Step 1)

- Involve review, adjudication, and pretest staff early in translation step, while setting up translation goals.
- Include reviews of early translations as part of translation step to clarify translation task specifications as needed before translation proceeds very far.
- Give detailed instructions to translators, including background on the reasons for and structure of the survey interview conversation.

Step 2: Review

After the translation step, we looked for opportunities for collaboration and feedback across the subsequent review, adjudication, and pretesting steps. Notably, we created the position of SLC for each of the three languages. The SLCs fulfilled two broad functions: They reviewed the initial target-language questionnaire translations to identify translation options, and they supervised cognitive interview pretest activities in their assigned target languages. Using the same staff for review and pretest roles was a way to ensure that review results informed pretesting designs and that pretesting results informed ongoing review activities.

Because SLCs were asked to fill both reviewing and pretesting roles, we looked for a range of skills when selecting them. We were interested in SLCs who were able to communicate fluently and skillfully both in the target language(s) and in English. Other desirable qualities we sought were the ability to offer insight into the cultures associated with their target languages in the United States, access to community resources for recruiting pretest interviewers and respondents, familiarity with survey methods, awareness of research on tobacco use and tobacco-related disease, and exposure to or experience with research requiring translation activities. None of the SLCs had all of these characteristics, but each reviewer had a mix of some of them. The one required attribute was the language fluency.

We hired four SLCs. Three were engaged early in the project. The fourth was hired to replace the original Vietnamese-language SLC after we learned of errors in initial rounds of reviewing and pretesting for the Vietnamese-language

translation. All four SLCs spoke the target language as their strongest (or first) language.

We set up detailed and relatively formal review and adjudication processes, involving a template that SLCs used to structure their reviews and to document results. SLCs recorded the following information on the template:

- the item number for a translation identified as problematic,
- a brief description of the type of problem,
- the reasons the translated item was problematic,
- a suggestion for a possible solution, and
- the reason the suggested solution would improve the item translation.

Because the SLCs had different levels of exposure to and experience with survey methods, we developed a variety of training materials, activities, and support resources to provide the survey methods background they needed to complete appropriate and consistent reviews for all three target-language questionnaires. An initial 2-hour training session gave SLCs information about the TUS-CPS purposes and measurement goals, reviewed the objectives for the translation review tasks, and instructed SLCs on using the review template to accomplish their reviews. The session also gave an overview of cognitive pretest interview methods to prepare SLCs for later tasks.

After the training, we set up biweekly group meetings to discuss review progress, emerging questions, potential impacts of cultural issues on translation effectiveness, and important interim results. In addition, we had informal meetings and conversations with SLCs to discuss detailed questions, language-specific issues, and unexpected problems. The informal exchanges occurred often and required considerable time from study research staff and from the individual SLCs. We believe these informal exchanges were particularly effective for helping SLCs understand how the survey items should function and how translation decisions can affect response accuracy.

Together with the original translations and translator notes, the templates that SLCs used to document the results of their reviews were key inputs to the initial adjudication phase. SLCs found that the translations were more problematic than we originally anticipated. Their comments identified complex, awkward, or unusual grammar and syntax and confusing word choices, which possibly reflected overly literal translations that violated conversational expectations.

Lessons Learned: Review (Step 2)

- Previous experience with survey methods is helpful but not necessary to ensure useful input from reviewers.

- SLCs who lack previous survey methods experience benefit from ongoing interactions with research staff and each other about item intent, item wording, and alternative translation options.
- Engage SLC reviewers early, during the translation step, to reduce the need for large-scale revisions after translation activities are completed and to benefit from direct interaction between translators and reviewers.

Step 3: Initial Adjudication

The role of the adjudicator is to make decisions about recommendations for improving the translations. Effective adjudication requires knowledgeable and versatile adjudicators (e.g., Harkness, Van de Vijver, and Mohler 2003; Harkness, Pennell, and Schoua-Glusberg 2004; U.S. Census Bureau 2004). Wherever SLCs identified potential problems with a target-language translation, the adjudicators' task was to review the problem and the suggested solution and to make a decision whether to retain the original translation, make the revision suggested by the SLC, or make an alternative revision.

Working through university research centers specializing in tobacco-related research, we found a lead adjudicator who had subject-matter expertise, translation experience, and a strong survey methods background. The lead adjudicator spoke Chinese as her first language and was skilled in the Mandarin and Cantonese dialects. She also communicated well in English. Through her academic appointment, she had access to tobacco researchers who spoke Vietnamese or Korean as their strongest languages. These three researchers formed the adjudication team. Because the adjudicators routinely worked together in their academic setting, they had already established strategies for communicating and collaborating with each other. Because of their expertise, we relied on the adjudicators to set appropriate translation decision criteria and accepted all their decisions during the initial adjudication phase. We folded the decisions into the original translations to produce a pretest questionnaire version for each target-language questionnaire.

Lessons Learned: Initial Adjudication (Step 3)

- University research centers are good resources for knowledgeable and versatile adjudicators.
- The structure of the adjudication process was effective. All adjudicators' decisions fit within the task guidelines, and reviewers and translators respected the adjudicator role.

Step 4: Cognitive Interview Pretest

Cognitive interviews are one method researchers and survey methodologists use to evaluate the quality and effectiveness of survey translations

(e.g., Weidmer 1994; Harkness and Schoua-Glusberg 1998; de la Puente, Pan, and Rose 2003; Harkness, Van de Vijver, and Mohler 2003; Harkness, Pennell, and Schoua-Glusberg 2004; U.S. Census Bureau 2004). These structured, open-ended interviews are designed to gather detailed information about the cognitive thought processes respondents use to understand and answer survey questionnaire items (Forsyth and Lessler 1991; Beatty 2004; Willis 2004, 2005). The goal is to identify and remove potential causes of response error, including items that use unfamiliar or inappropriate terminology, items that respondents interpret in unexpected ways, or items that ask for information that respondents have difficulty remembering. Cognitive interviews are an effective complement to psychometric methods typically used to quantify survey measurement characteristics (Willis 2005).

When cognitive interviews are used to test a questionnaire translation, results can be used to identify additional deficiencies related to the translation itself, such as target-language terms that respondents interpret differently than intended, target-language terms that are unfamiliar, or target-language terms that have culture-specific meanings. We chose to pretest the survey translations using cognitive interviews because we were interested in both sets of issues: cognitive difficulties respondents may have in understanding and answering the intended questions and translation issues that may alter question meaning.

We planned to conduct nine cognitive interviews to test the Chinese translation, nine cognitive interviews to test the Korean translation, and nine cognitive interviews to test the Vietnamese translation. As described below, we later conducted an additional round of five cognitive interviews to test a second revision of the Vietnamese-language translation.

Although we decided to test a single Chinese-language questionnaire, more recent cognitive testing research suggests minor but potentially important differences between the two dialects, at least in some survey domains (Levin et al. 2005). In retrospect, we would recommend including a larger number of Chinese-language respondents to explicitly test the recommendation to use a single Chinese-language questionnaire. Based on the relatively small number of interviews we conducted, we found no data to suggest that separate Mandarin and Cantonese translations would improve survey responses.

Lessons Learned: Cognitive Interview Pretest (Step 4)

- When dialects may be of interest, include enough interviews to test translation effectiveness separately for each relevant dialect.

Cognitive interview research normally requires the five steps in Table 2. The following paragraphs describe how we accomplished each of these steps.

TABLE 2
Five Steps for Conducting Cognitive Pretest Interviews

-
-
1. Develop cognitive interview scripts.
 2. Identify and train cognitive interviewers.
 3. Recruit eligible respondents.
 4. Conduct and summarize cognitive interviews.
 5. Identify key findings and make recommendations.
-

Develop cognitive interview scripts. We developed an English-language cognitive interview script that consisted of the questionnaire items with cognitive probes inserted after selected questions. Probe questions asked respondents to describe how they interpreted survey items, how they understood particular question and response wordings, whether they had difficulty understanding how to use response sets, whether the response sets seemed incomplete, the factors they considered as they selected their answers, and whether any questions asked for information they either did not have or could not remember. Once the English-language script was complete, SLCs translated the cognitive probes into the target languages and added them to the target-language questionnaires to create three target-language cognitive interview scripts.

Identify cognitive interviewers. Each SLC hired two cognitive interviewers to conduct cognitive pretest interviews in the target languages. Both of the Vietnamese-language SLCs had access to bilingual staff and/or coworkers whom they hired. The Korean- and Chinese-language SLCs developed strategies for locating and hiring bilingual interviewers that involved a mix of community networking, outreach to community and research resources, Web postings, and use of advertising media.

We were able to pay cognitive interviewers relatively high rates for interviewers, but the pay was lower than professional researchers would expect. Fortunately, we found that qualified candidates were attracted to the cognitive interviewing work because of its social and scientific value and because of the potential for additional research opportunities in the future.

Lessons Learned: Identify Cognitive Interviewers (Step 4)

- Document interviewer hiring activities in detail. Staff who have trouble finding suitable candidates will benefit from others' successes.
- Community networking and other outreach activities were far more effective for identifying and recruiting cognitive interviewers than were Web postings or other advertising activities.

Train cognitive interviewers. We conducted a 6-hour session to train interviewers to administer cognitive interviews in their target languages. The SLCs also attended the session as trainees to make certain that they had the knowledge and skills required to effectively monitor and supervise cognitive interviewing activities. Because most of the cognitive interviewers had no previous interviewing experience, the training session gave overviews of standard interview practices and conventions and the structure and content of the TUS-CPS questionnaire. In addition, the training covered cognitive interview goals and techniques, reviewed the cognitive interview probes and their purposes, and included an English-language demonstration of a cognitive interview. Most of the training was in English, but trainees spent roughly 2 hours using role-playing methods to conduct practice cognitive interviews in their target languages.

Lessons Learned: Train Cognitive Interviewers (Step 4)

- In training, review both the survey questionnaire and the cognitive interview script. Separate reviews will help inexperienced cognitive interviewers recognize the different functions of survey questions and cognitive interview probes.
- Provide ample time for monitored practice and feedback. Conduct monitored practice interviews both in English and in the target language.
- Provide guidelines and additional practice about adapting cognitive interview probes as necessary to help interviewers avoid repetition and follow up unexpected responses.

Recruit eligible respondents. SLCs filled important roles marketing the study by screening potential volunteers for eligibility, convincing them to participate, and setting cognitive interview appointments. Eligible respondents met two screening criteria: They currently smoked or had stopped smoking within the past 5 years, and they reported limited ability to speak or understand English. From volunteers identified as eligible, SLCs recruited and scheduled interviews with respondents who had a mix of education, age, years in the United States, gender, and socioeconomic status (represented by occupation).

The SLCs were reluctant to use personal contacts to recruit respondents because of the sensitive nature of smoking habits within some Asian American communities. Therefore, they contacted potential respondents using flyers, newspaper advertisements, and tables at community locations and events. They also made extensive use of word-of-mouth recruiting activities by working through community networks. These activities were time consuming because contacts and potential respondents asked for considerable detail about the general study purposes, the specific tasks respondents would be asked to complete, and the SLCs' role in the study.

We anticipated the cognitive interviews would last about an hour, and we planned to pay respondents \$35 to thank them for their time. SLCs expressed some concerns about the payments. On one hand, they thought some volunteer respondents would hesitate to accept incentive payments because of worries about work permit status and questions about reporting the payment for tax purposes. On the other hand, they believed respondents without those fears would find the incentive payment too small. We were able to increase the incentive to \$40 for a second set of interviews testing the Vietnamese-language questionnaire. SLCs suggested supplementing the incentive payments with small gifts such as fruit or cookies, particularly when interviews were held in respondents' homes. We allotted \$15 per respondent for this purpose.

Based on their knowledge of the communities they worked in and their respondent recruiting conversations, SLCs were concerned about potential respondent reactions to study materials and procedures. They felt that some respondents would refuse to participate in interviews because of fears that involvement in the survey could draw attention to their immigration status. SLCs also predicted that respondents would resist being tape-recorded and would be nervous about signing the confidentiality form and the study receipt. Research staff worked with SLCs to develop recruiting approaches that would address issues like these if respondents raised them. Fortunately, volunteer respondents actually expressed few concerns about participating in the cognitive interview pretest. We observed no sensitivities related to immigration status, tape-recording, accepting financial compensation, or signing study confidentiality and incentive receipt forms.

Lessons Learned: Recruit Respondents (Step 4)

- Document respondent recruiting activities. Details will be useful for evaluating alternative approaches and suggesting better approaches to other investigators who have trouble finding eligible volunteer respondents.
- Word-of-mouth contacts through community networks were effective for finding and recruiting eligible volunteer respondents. Contacts through professional networks were particularly productive.
- SLCs should be prepared to provide extensive explanation about the study including study goals, what it requires of respondents, and the SLCs' role in the study.
- Pay attention to SLCs' concerns about respondent participation and monitor respondent reactions as necessary. For this study, respondents had few concerns about participating in the pretest interviews.

Conduct and summarize cognitive interviews. We conducted two rounds of cognitive interviews. In the first round, six interviewers conducted twenty-seven interviews. Nine interviews tested the Chinese-language translations.

Five of the Chinese-language interviews used the Mandarin dialect, and four used the Cantonese dialect. Nine interviews tested the Korean-language translation, and nine interviews tested the Vietnamese-language translation. As reported below, results from the first round of cognitive interviews revealed important problems with the Vietnamese-language translation. After the Vietnamese-language translation was rereviewed and revised, two new cognitive interviewers conducted a second round of five interviews to test the revised Vietnamese-language translation.

In most cases, the cognitive interviewers conducted the interviews in respondents' homes. SLCs observed all cognitive interviews and took detailed notes. After obtaining respondent consent, interviewers audiotaped the cognitive interviews for later review and quality control. SLCs reviewed their notes and the audiotape for each interview and wrote a summary that described the respondent's reactions to target questionnaire items, highlighting key findings and results.

SLCs were responsible for monitoring cognitive interview quality based on their observations, notes, and audiotape reviews. We scheduled debriefing meetings with individual SLCs every second or third interview to give them an opportunity to report any procedural issues or concerns about interviewer performance. SLCs used these meetings as opportunities to gather suggestions, advice, and operational support. Research staff used these meetings as opportunities to monitor quality and identify interim results.

In early meetings, SLCs reported that interviewers were having difficulty administering the cognitive interview probes and that respondents seemed impatient because of redundancy in interviewers' questions. Further discussion revealed three issues in cognitive interview administration. First, interviewers were administering the probes word-for-word, even when respondents had already provided the relevant information. Second, interviewers were hesitant to deviate from scripted probes and use their own words to get at identified pretest issues. Third, interviewers were consistently missing opportunities to use their own probes to follow up unanticipated problems. Based on this early feedback, we conducted retraining sessions separately with each SLC. The SLCs, in turn, retrained their cognitive interviewers. This retraining activity seemed effective. After retraining, the cognitive interviews were shorter, we heard fewer reports of respondents' being annoyed by redundancy, and SLCs' interview summaries were more focused and more useful for identifying reasons for respondent confusion or error.

Interim results reported by the Vietnamese-language SLC led us to suspect the quality of the cognitive interviews and eventually the quality of the translation itself. To address our concerns, we hired a second Vietnamese-language SLC. His initial task was to review tape recordings from cognitive

interviews testing the Vietnamese-language translation. He discovered a variety of errors. For example, interviewers read the “true or false” items without pauses, giving respondents no opportunity to answer. Also, cognitive probes and interviewer instructions were translated incorrectly, tested survey items were administered improperly, and interviewers mistakenly read interviewer instructions and skip patterns to respondents.

Based on these findings, we decided we could not rely on results from the first round of cognitive interviews testing the Vietnamese-language translation. The new Vietnamese-language SLC followed our established review and adjudication process to revise the Vietnamese-language translation. Also, he hired and trained one new cognitive interviewer with whom he conducted a second round of five interviews testing the revised Vietnamese-language translation.

Lessons Learned: Conduct and Summarize Cognitive Interviews (Step 4)

- Anticipate retraining, particularly if cognitive interviewers are inexperienced.
- Add active quality control activities during initial cognitive interviews to support interview observation and SLC debriefing activities.
- Identify retraining needs based on SLCs’ summaries, audiotape reviews, and interviewer notes from early interviews. Look for adequate detail, minimal redundancy within interviews, and evidence of proper use of scripted and spontaneous probes.
- Carry out the following active quality control activities because they increase the likelihood that early interviews will be as useful as later interviews.
 - Divide the first round of cognitive interviews into small sets of one or two interviews.
 - Allow time between sets to observe and/or review interviews and retrain interviewers as needed.
 - When possible, observe early interviews in target language(s) using simultaneous interpretation to convey interview content to observers who are not fluent in the target language(s).
 - Engage a range of staff in interview observation to ensure review and comments from several viewpoints.
- Consider using SLCs to conduct cognitive interviews rather than hiring a separate cognitive interviewing staff to capitalize on knowledge and perspective SLCs gain during earlier review and adjudication.

Identify key findings and make recommendations. Detailed descriptions of specific cognitive interview results are provided elsewhere (Willis et al. 2005). Our focus here is on identifying effective translation processes. SLCs’ interview summaries were the primary source for identifying key findings and recommendations. We asked SLCs to write summaries that “described everything that happened” during an interview. We coached them to avoid editing out details that seemed unimportant to them, we asked them

to watch for both verbal and nonverbal reactions, and we encouraged them to include information about interviewer behaviors and observer reactions as well as information about respondent behaviors and reactions.

We used qualitative analytic methods to review these inclusive summaries, identify general themes, and document detailed variations in themes across individual interviews. After we completed analyses within each target language, we reviewed the full set of results to explore whether any general themes or variations were evident across the target languages.

As we analyzed the interview summaries, we often consulted with individual SLCs to verify that we were interpreting interview results correctly and gather advice about recommendations and priorities. Based on the qualitative analyses, we drafted a preliminary report that detailed major findings and included lists of recommendations for each target-language translation. SLCs reviewed the draft report and recommendations for their own target languages, and we used their feedback to finalize a set of recommendations for the final review and adjudication.

Lessons Learned: Identify Key Findings and Develop Recommendations (Step 4)

- The general, nondirective instruction to “describe everything that happened,” supported with frequent consultation with SLCs, is a useful approach for producing detailed and useful interview summaries.

Step 5: Final Review and Adjudication

Finally, we convened a meeting to conduct the final review and adjudication. At this meeting, the National Cancer Institute project director determined which recommendations to accept and which revisions to make to the three target-language questionnaires. The final review meeting was attended by all research staff working on the translation project and the three SLCs who were still involved with the project.

We did not attempt to define specific roles for research staff and the SLCs during the final review and adjudication step. Therefore, it was interesting to note that research staff and SLCs seemed to take on different functions. Research staff tended to focus on identifying and classifying problems respondents had in understanding and answering survey questions. SLCs used examples and context to illustrate the types of problems observed and to clarify the effects these problems had on survey responses. The project director found that both types of information were necessary to make good decisions about revising the three target-language questionnaires.

It is instructive to notice differences between the initial and final adjudication steps in terms of the number of changes considered and the kinds of

changes considered. At the initial step, the adjudicators made extensive revisions to the three target-language questionnaires based on many suggestions from the SLC reviewers. Furthermore, most of the changes made during the initial adjudication step were language specific, focused on improving the individual target-language translations to make the questionnaires more conversational, less wordy, and easier to understand.

In contrast, at the final step, the project director made relatively few additional revisions to the target-language questionnaires. Some of these revisions were made to improve specific language aspects of individual translations. However, other revisions were made to all three target-language questionnaires to clarify wording that seemed ambiguous across the languages.

We believe that differences in the number and kinds of revisions made during the initial and final adjudication steps indicate the general success of the five-step process for translation and evaluation in Table 1. In particular, we found that the review and initial adjudication steps identified most of the shortcomings in the draft translations. These shortcomings were effectively addressed by changes made in the initial adjudication step. Thus, the cognitive testing and final adjudication steps could focus on more universal issues such as clarity of question purpose and response set completeness.

Lessons Learned: Final Review and Adjudication (Step 5)

- Aside from the decision-making role filled by the project director, leave other roles unspecified. This gives research staff and SLCs freedom to select the topics about which they feel qualified to speak.
- When early steps of the translation process effectively address translation errors and shortcomings, then later steps can focus on more general questionnaire design issues that may influence responses regardless of the language used to administer interviews.

CONCLUSIONS

This article described a five-step translation and evaluation process to produce translated versions of a tobacco use questionnaire. Based on careful documentation and appropriate qualitative evaluation, we believe the translation process described here—involving multiple steps, activities, and types of staff—was successful and necessary, even though the process was somewhat involved. Results from review and cognitive testing activities indicate that all three Asian target-language questionnaires accurately reflect the content and the measurement goals of the English-language source questionnaire.

The procedures used at each step supported and reinforced the team-based collaboration we intended. For example, translators provided information to reviewers about the translation challenges they faced and the decisions they made, and reviewers used this information as one resource for their initial review. Reviewers worked collaboratively with project research staff and with each other to identify recommendations for revision and document their reasons. Adjudicators made decisions as prescribed, either accepting an initial translation, accepting a reviewer's recommendation, or developing their own solutions to a reviewer-identified problem. Reviewers and translators respected the adjudicator role and accepted adjudicators' decisions. Cognitive interviewers and interview observers built on early results to identify additional refinements that seemed necessary based on suggestions from monolingual volunteer respondents. In the final adjudication step, reviewers and project research staff worked together to support the project director's final decision making.

We anticipate the translation process described here will apply to non-Asian languages as well, based in part on earlier success using similar though less comprehensive procedures to produce a Spanish-language version of the tobacco use questionnaire (Westat 2002).

Building on experiences documented here, we make four preliminary recommendations for future translation efforts:

- Use SLCs throughout the translation and evaluation process.

The SLC role provides staff continuity so knowledge gained in the early review step informs subsequent cognitive testing activities and recommendations. SLCs serving as reviewers and cognitive interview supervisors can be selected to have a mix of necessary skills. SLCs without previous survey research experience should be given early opportunities to gain working knowledge of survey goals and methods.

- Integrate initial translation activities with subsequent review and testing.

In this study, practical arrangements limited integration. For future translations, we anticipate using an enhanced model that draws translators more closely into iterative sets of translation and evaluation activities.

- Use SLCs as cognitive interviewers.

By the time we conducted cognitive interviewer training, SLCs already had pragmatic experience with survey methods. This experience gave them

a firm foundation for conducting and analyzing cognitive interviews. When it is impractical to use SLCs as cognitive interviewers, cognitive interviewer training should be expanded.

- Include proactive quality control measures, particularly during cognitive interviewing. This step is important even when interviewers are highly skilled.

Schedule initial interviews to allow considerable time between interviews for observation, review, debriefing, and retraining. In future projects, we will use this procedure to observe at least two interviews per cognitive interviewer. To ensure consistent interview quality, we will continue to observe and retrain between interviews until individual interviewers are certified as ready to continue interviewing with less intensive oversight.

At this point, our assessment of the TUS-CPS translations is based on qualitative observations. Our next action is to evaluate the five-step translation and evaluation process quantitatively, using behavior coding methods (Fowler and Cannell 1996) to quantify data quality. Our goal is to compare behavior coding results with the review and cognitive interview results obtained here to determine whether revisions made based on review and cognitive testing activities appear to enhance survey data quality.

REFERENCES

- Beatty, P. 2004. The dynamics of cognitive interviewing. In *Methods for testing and evaluating survey questionnaires*, ed. S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, and E. Singer, 45–66. New York: John Wiley.
- Brislin, R. W. 1970. Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology* 1 (3): 185–216.
- Centers for Disease Control and Prevention. 2005. Cigarette smoking among adults—United States, 2003. <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5420a3.htm> (accessed January 30, 2007).
- de la Puente, M., Y. Pan, and D. Rose. 2003. An overview of a proposed Census Bureau guideline for the translation of data collection instruments and supporting materials. Paper presented at the meeting of the Federal Committee on Statistical Methodology, Arlington, VA, November 17–19.
- European Social Survey. 2002. An outline of ESS translation strategies and procedures. http://naticent02.uuhost.uk.uu.net/ess_docs/translation.pdf (accessed September 30, 2005).
- Forsyth, B. H., and J. T. Lessler. 1991. Cognitive laboratory methods: A taxonomy. In *Measurement errors in surveys*, ed. P. P. Biemer, R. M. Groves, L. E. Lyber, N. A. Mathiowetz, and S. Sudman, 393–418. New York: John Wiley.
- Fowler, F. J., and C. F. Cannell. 1996. Using behavioral coding to identify cognitive problems with survey questions. In *Answering questions: Methodology for determining cognitive and communicative processes in survey research*, ed. N. Schwarz and S. Sudman, 15–36. San Francisco: Jossey-Bass.

- Groves, R. M. 1989. *Survey errors and costs*. New York: John Wiley.
- Harkness, J. A., B.-E. Pennell, and A. Schoua-Glusberg. 2004. Survey questionnaire translation and assessment. In *Methods for testing and evaluating survey questionnaires*, ed. S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, and E. Singer, 453–74. New York: John Wiley.
- Harkness, J. A., and A. Schoua-Glusberg. 1998. Questionnaires in translation. In *Cross-cultural survey equivalence*, ed. J. A. Harkness, 87–126 (ZUMA-Nachrichten Spezial Number 3). Mannheim, Germany: ZUMA.
- Harkness, J. A., F. J. R. Van de Vijver, and P. Mohler. 2003. *Cross-cultural survey methods*. Hoboken, NJ: John Wiley.
- Jenkins, C. N. H., S. J. McPhee, N.-T. Ha, T. V. Nam, and A. Chen. 1995. Cigarette smoking among Vietnamese immigrants in California. *American Journal of Health Promotion* 9 (4): 254–56.
- Levin, K., A. Norberg, F. Thompson, G. B. Willis, and A. Grauman. 2005. CHIS 2005 adult questionnaire: Additional modules, nutrition, body image, and dieting behavior modules C & D, Cantonese Chinese interview summary. Unpublished technical report prepared by Westat for the National Cancer Institute.
- Lew, R., and S. P. Tanjasiri. 2003. Slowing the epidemic of tobacco use among Asian Americans and Pacific Islanders. *American Journal of Public Health* 93 (5): 764–68.
- Ma, G. X., S. Shive, Y. Tan, and J. I. Toubbeh. 2002. Prevalence and predictors of tobacco use among Asian Americans in the Delaware Valley region. *American Journal of Public Health* 92 (June): 1013–20.
- McKay, R. B., M. J. Breslow, R. J. Sangster, S. M. Gabbard, R. W. Reynolds, J. M. Nakamoto, and J. Tarnai. 1996. Translating survey questionnaires: Lessons learned. *New Directions for Evaluation* 70 (Summer): 93–105.
- National Cancer Institute. 2006. Tobacco Use Supplement, Current Population Survey: Table 1. 2003 Tobacco Use special Cessation Supplement to the Current Population Survey (TUSCS-CPS). <http://riskfactor.cancer.gov/studies/tus-cps/results/data03/table1.html> (accessed January 30, 2007).
- Pan, Y. 2003. The role of sociolinguistics in the development and conduct of federal surveys. Paper presented at the meeting of the Federal Committee on Statistical Methodology, Arlington, VA, November 17–19.
- Patrick, D. L., A. Cheadle, D. C. Thompson, P. Diehr, T. Koepsell, and S. Kinne. 1994. The validity of self-reported smoking: A review and meta-analysis. *American Journal of Public Health* 84 (7): 1086–93.
- Riesch, S. K., L. S. Anderson, N. Angresano, J. Canty-Mitchell, D. L. Johnson, and K. Krainuwat. 2006. Evaluating content validity and reliability of the Children's Health Risk Behavior Scale. *Public Health Nursing* 23 (4): 366–72.
- Schaeffer, N. C., and D. W. Maynard. 2001. Standardization and interaction in the survey interview. In *Handbook of interview research*, ed. J. F. Gubrium and J. A. Holstein, 577–602. Thousand Oaks, CA: Sage.
- Turner, C. F., and E. Martin. 1984. *Surveying subjective phenomena*, vol. 2. New York: Russell Sage.
- U.S. Census Bureau. 2004. *Census Bureau guideline: Language translation of data collection instruments and supporting materials*. Suitland, MD: U.S. Census Bureau.
- Weidmer, B. 1994. Issues and guidelines for translation in cross-cultural research. Paper presented at the meeting of the American Association for Public Opinion Research, Danvers, MA, May 12–15.
- Westat. 2002. Summary report: DHQ Spanish translation. Unpublished technical report prepared by Westat for the National Cancer Institute.

- Wewers, M. E., R. K. Dhatt, M. L. Moeschberger, R. M. Guthrie, P. Kuun, and M. S. Chen. 1995. Misclassification of smoking status among Southeast Asian adult immigrants. *American Journal of Respiratory and Critical Care Medicine* 152 (6): 1917–21.
- Willis, G. B. 2004. Cognitive interviewing revisited: A useful technique in theory? In *Methods for testing and evaluating survey questionnaires*, ed. S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, and E. Singer, 23–44. New York: John Wiley.
- . 2005. *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.
- Willis, G. B., D. Lawrence, M. Kudela, K. Levin, and K. Miller. 2005. The use of cognitive interviewing to study cultural variation in survey response. Paper presented at the Questionnaire Evaluation Standards (QUEST) Workshop, Santa Fe, NM, August 8–12.

BARBARA H. FORSYTH is a senior study director at Westat. She is a cognitive psychologist with research interests in survey questionnaire design and cognitive testing methods for improving questionnaires. Recent publications and presentations have focused on questionnaire development and testing for surveys conducted by the National Cancer Institute, the National Center of Education Statistics, the U.S. Census Bureau, and the Internal Revenue Service.

MARTHA STAPLETON KUDELA is a senior study director at Westat with special interests in survey design and development. Recent project work includes cognitive interview and focus group testing for the National Cancer Institute, the National Highway and Traffic Safety Administration, the Internal Revenue Service, the National Institute for Dental and Craniofacial Research, and the Army Research Institute.

KERRY LEVIN, an associate director at Westat, is a social psychologist with experience conducting large business and household surveys related to customer service and evaluation for agencies such as the National Institute of Standards and Technology, the Internal Revenue Service, and the U.S. Patent and Trademark Office. She is also involved in cross-cultural qualitative research applying techniques including cognitive testing, retrospective debriefings, and behavior coding to evaluate survey items in languages such as Mandarin and Cantonese Chinese, Spanish, and Korean.

DEIRDRE LAWRENCE is an epidemiologist in the Risk Factor Monitoring and Methods Branch in the Division of Cancer Control and Population Sciences, National Cancer Institute. Her research interests include using surveillance, statistical, and epidemiologic research methods to monitor trends in tobacco and analyzing differences in predictors and patterns of cancer-related risk factors. She is currently conducting work to gain an in-depth understanding of how Japan is working to reduce tobacco use and other lifestyle risk factors.

GORDON B. WILLIS is a cognitive psychologist in the Division of Cancer Control and Population Sciences at the National Cancer Institute. He now works mainly in the area of the development and evaluation of surveys on cancer risk factors, and serves as chair of the National Cancer Institute Institutional Review Board for human subjects protection in epidemiological research. His research interests increasingly involve cross-cultural issues in self-report surveys and studies.