Methods for Translating Survey Questionnaires[1]

Barbara H. Forsyth[2], Martha Stapleton Kudela[2], Deidre Lawrence[3], Kerry Levin[2], and Gordon B. Willis[3]
Westat, 1650 Research Blvd., Rockville, MD, U.S.A., 20850[2]
National Cancer Institute, 6130 Executive Blvd., Rockville, MD, U.S.A., 20852[3]

1. Introduction

This paper explores procedures for developing and evaluating questionnaire translations for surveys administered in multiple languages. We focus on recent work for the National Cancer Institute (NCI) translating an English-language questionnaire on tobacco use into Mandarin Chinese, Cantonese Chinese, Korean and Vietnamese. We used an iterative translation, evaluation and review process. This paper describes the iterative process and lessons learned. Our purposes are pragmatic. Can we identify useful practices for developing and testing questionnaire translations?

The Tobacco Use Survey (TUS) is administered in the United States as a component of the Current Population Survey (CPS). It asks questions about tobacco use patterns, smoking prevalence, workplace smoking policies, level of nicotine dependence, medical advice to quit smoking, quit attempts, cessation methods used, and changes in smoking norms and attitudes. Several features of the TUS questionnaire and the content area in general make this questionnaire a good candidate for translation into four Asian languages. Although the general Asian-American population tends to have a lower rate of tobacco use than the general U.S. population (Centers for Disease Control and Prevention, 1998), tobacco use rates vary considerably among Asian-American subgroups (Ma et al., 2004; Ma et al., 2002). Furthermore, local surveys show high tobacco use rates for some Asian-American subgroups (Lew & Tanjasiri, 2003; Ma et al., 2002). Because tobacco use is associated with serious health risks, it is important to conduct tobacco use surveys using methods that ensure complete and accurate data from Asian-Americans.

Survey researchers have advocated a range of practices for producing effective translations (e.g., Census Bureau, 2004; Harkness, Van de Vijver & Mohler, 2003; McKay, Breslow, Sangster, Gabbard, Reynolds, Nakamoto & Tarnai, 1996). Translation researchers seem to prefer team-based approaches when study

resources permit them. Team approaches generate more translation options and provide sounder and less idiosyncratic translation review and evaluation (Census Bureau, 2004; European Social Survey, 2002; Harkness et al., 2003). We selected an iterative, team-based approach based on the multi-stage translation frameworks described in Harkness et al. (2003) and a U.S. Census Bureau (2004) translation guideline.

2. Translation and Evaluation Processes

We developed a 5-step process for translating the tobacco use questionnaire into Mandarin, Cantonese, Korean and Vietnamese. The five steps are: translation; review; initial adjudication, cognitive interview pretesting and final review and adjudication. This paper focuses on the procedures we used to conduct each step.

2.1 Translation

The translation staff consisted of three independent professional translators. One was multilingual in English, Mandarin and Cantonese, one was bilingual in English and Korean and one was bilingual in English and Vietnamese.

The translation project aimed to develop translations that "ask the same questions" (e.g., Harkness et al., 2003). Based on these instructions, the three translators worked independently to produce target-language translations. A translation coordinator supervised their work and was available as needed to answer questions and provide guidance.

The translation step yielded 3 sets of translated items: a single translation for the Mandarin and Cantonese dialects of Chinese, a second translation in Korean, and a third translation in Vietnamese. Translators also provided documentation that described specific translation challenges they encountered and decisions they made to deal with these challenges. These translation products were input for the next review step.

Because of the way our research was funded, the translation step was independent from succeeding review, adjudication and pretesting steps. We would

---

prefer to set up collaborative working relations among translators, reviewers and adjudicators early in the translation step. Fortunately, we were able to implement a more collaborative approach after the initial translation step was finished.

Lessons Learned: Translation Step

- Involve review, adjudication and pretest staff early in translation step, while setting up translation goals.
- Include early review as part of translation step to identify and eliminate ambiguities in translation task specifications before translation moves too far ahead.
- Give unambiguous instructions to translators, including the reasons for and structure of the survey interview conversation.

2.2 Review

For the review, adjudication and pretesting phases, we created the position of Survey Language Consultant (SLC) for each of the target languages. The SLCs fulfilled two broad functions: they reviewed the initial questionnaire translations to identify translation options, and they supervised cognitive interview pretest activities. 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.

We hired a total of four SLCs. Three of them were engaged early in the project and we hired a fourth later to replace the original Vietnamese-language SLC when we recognized errors in initial rounds of reviewing and pretesting for the Vietnamese-language translation. One SLC was multilingual in English, Mandarin Chinese and Cantonese Chinese, one was bilingual in English and Korean, and two were bilingual in English and Vietnamese.

At the outset, we anticipated that the review and initial adjudication steps would be relatively brief and straightforward. Early on, it became clear the project would benefit from more formal review and adjudication processes. We developed a template that SLCs used to structure their reviews and to document results from their reviews. SLCs completed the template by identifying items in the target translations that seemed problematic, describing the reasons an item was problematic, suggesting a possible solution, and describing how each suggested revision would improve the target translation.

The SLCs had different levels of experience with survey methods. We developed training materials, on-the-job training activities, and other feedback and support resources to provide the survey methods background needed to ensure comparable levels of review across the three target-language questionnaires. We conducted a 2-hour training session to give SLCs information about the survey purposes and measurement goals, objectives for the translation review tasks, using the review template to accomplish their reviews, and cognitive pretest interview methods.

We set up routine biweekly group meetings to discuss review progress, questions, and interim results. In addition, we set up informal meetings with SLCs as needed to discuss more detailed questions, language-specific issues, and unexpected problems. The templates that SLCs used to document the results of their reviews together with the original translations and translator notes were inputs to the initial adjudication phase.

Lessons Learned: Review Step

- Previous experience with survey methods is useful but not necessary to ensure effective input from reviewers.
- SLCs who lack previous survey experience benefit from ongoing conversations with research staff and each other about item intent, wording and translation options.
- Engage reviewers early, during translation, to reduce needs for large-scale revisions after translation is completed and to benefit from direct interaction between translators and reviewers.

2.3 Initial Adjudication

Effective adjudication requires knowledgeable and versatile adjudicators (e.g., Census Bureau, 2004; Harkness et al, 2003; Harkness, Pennell & Schoua-Glusberg, 2004). Working through university research centers specializing in tobacco-related research, we found a lead adjudicator who had subject-matter expertise, translation experience, and familiarity with survey data collection methods. The lead adjudicator spoke Chinese as her first language, and was skilled in the Mandarin and Cantonese dialects. Through her academic appointment, she had access to other tobacco researchers who spoke Vietnamese or Korean as their strongest (or "first") languages. These researchers formed the adjudication team.

Wherever SLCs identified potential problems with a target-language translation, the adjudicators' tasks

were to review the problem and the suggested solution and to make a decision whether and how to revise the original translation. We adopted all decisions made by the adjudication team during the initial adjudication phase and folded them into the original translations to produce a pretest version for each target-language.

Lessons Learned: Initial Adjudication Step

- 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.

2.4 Cognitive Interview Pretest

Cognitive interviews are structured, open-ended interviews designed to gather detailed information about the cognitive thought processes respondents use to understand and answer survey questionnaire items (Beatty, 2004; Forsyth & Lessler, 1991; Willis, 2005). For example, results can identify items that use unfamiliar or inappropriate terminology, items that respondents interpret in unexpected ways, or items that ask for information respondents have difficulty remembering. When cognitive interviews are used to test a questionnaire translation, results can be used to identify additional translation deficiencies. For example, 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 developed an English-language cognitive interview script that consisted of the questionnaire items with cognitive probes inserted after selected questions. For example, probe questions asked respondents to describe how they understood particular question and response wordings, whether the response sets seemed incomplete, and whether any questions asked for information respondents have difficulty recalling. SLCs translated the cognitive interview script into the target languages. Each SLC hired two cognitive interviewers to conduct cognitive pretest interviews in the appropriate target language.

We conducted a 6-hour session to train interviewers to administer cognitive interviews in their target languages. The training session gave overviews of standard interview practices and conventions and the content of the tobacco use items. 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

- Include separate reviews of the questionnaire and the cognitive interview script to highlight 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.

- Give guidelines and additional practice adapting cognitive interview probes as necessary to avoid repetition and to follow up unexpected responses.

SLCs recruited respondents who had limited ability to speak or understand English and a mix of education, years in the United States, gender, and socioeconomic status (represented by occupation). Our budget was not large enough to also include respondents from a mix of regional dialects within each target language.

Lessons Learned: Recruit Respondents

- Document respondent recruiting activities in detail. Details will help staff evaluate alternative approaches and suggest better approaches to anyone who has 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.

- For word-of-mouth recruiting methods, SLCs should be prepared to provide extensive explanation about the study including study goals, what it requires of respondents, and the SLC's role in the study.

- Volunteer respondents had relatively few concerns about participating in the pretest interviews.

We anticipated the cognitive interviews would last about an hour and we paid respondents a small incentive to thank them for their time. SLCs suggested supplementing the incentives with small gifts such as fruit or cookies, particularly when interviews were held in respondents' homes. For example, the Vietnamese-language interviews coincided with the Moon Festival, and interviewers brought moon cakes,

an ethnic Vietnamese sweet, for respondents. Other interviewers brought fruit baskets or gift certificates for local ethnic food markets.

We conducted two rounds of cognitive interviews. In the first round, six interviewers conducted a total of 27 interviews. Nine interviews tested the Chinese-language translations. Roughly half of the Chinese-language interviews were in Mandarin and roughly half were in Cantonese. Nine interviews tested the Korean-language translation, and nine interviews tested the Vietnamese-language translation. Results from the first round of cognitive interviews revealed important problems with the Vietnamese-language translation. After the Vietnamese-language translation was re-reviewed and revised, two new cognitive interviewers conducted a second round of five interviews to test the revised Vietnamese-language translation. SLCs observed all cognitive interviews, took detailed notes, reviewed interview audiotapes, and wrote a summary for each interview.

SLCs monitored 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 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.

Based on interim results reported by the Vietnamese-language SLC, we suspected the quality of the cognitive interviews and of the translation itself. We hired a second Vietnamese-language SLC who reviewed tape recordings from the Vietnamese-language interviews. He discovered a variety of errors. 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. We decided we could not trust results from the first round of Vietnamese-language interviews. The new Vietnamese-language SLC followed our established review and adjudication process to revise the Vietnamese-language translation. Also, he hired and trained two new cognitive interviewers to conduct a second round of five interviews testing the revised Vietnamese-language translation.

We have reported the cognitive interview results elsewhere (Willis et al., 2005a; 2005b). The focus here is on identifying effective analytic 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." We coached them to avoid editing out details that seemed unimportant to them, and to include information about interviewer behaviors and observer reactions as well as information about respondent behaviors and reactions. The resulting, inclusive summaries gave analysts access to a full range of reactions, including verbal, nonverbal and emotional reactions from respondents and reactions from interviewers and observers.

Lessons Learned:
Conduct and Summarize Cognitive Interviews

- Introduce independent quality control activities early, during initial review and adjudication steps. Careful quality control activities are particularly important when relatively few research staff are skillful with the target languages.

- Anticipate retraining. Use SLCs' summaries and reviews to identify retraining needs. Look for adequate detail, minimal redundancy within interviews, and evidence of proper use of cognitive interview probes.

- Add active quality control activities during initial cognitive interviews to support interview observation and SLC debriefing activities.

  • The following active quality control activity seems promising to us because it increases the likelihood that early interviews will be as useful as later interviews.

  • Break the first round of cognitive interviews into small sets of 1 or 2 interviews.

  • Separate these sets, giving time to observe and/or review interviews and re-train interviewers as needed.

- When possible, observe initial interviews using simultaneous interpretation to ensure observers have access to all interview details.

- Engage a range of staff in interview observation and review to ensure input from several viewpoints.

- Consider using SLCs to conduct cognitive interviews rather than hiring a separate cognitive interviewing staff to further benefit from the knowledge and perspective SLCs gain during review and adjudication steps.

We used qualitative analytic methods to review the summaries and identify general themes. As part of the analysis, we often consulted with SLCs to verify that

we were interpreting interview results correctly and to gather advice about recommendations and priorities. SLCs reviewed draft reports and recommendations and provided additional feedback. We used their feedback to finalize a set of recommendations in preparation for the final review and adjudication.

Lessons Learned: Identify Key Findings and Develop Recommendations

- The general and nondirective instruction to "describe everything that happened" helped SLCs produce detailed and useful interview summaries.

2.5 Final Review and Adjudication

We convened a meeting to conduct the final review and adjudication. At this meeting, NCI's project director determined which recommendations to accept and what revisions to make to the three target-language questionnaires. We did not attempt to define specific roles for research staff and the SLCs during the final review and adjudication step. We found that research staff tended to focus on identifying and classifying problems respondents had understanding and answering survey questions. SLCs focused on giving examples and context to illustrate the types of problems observed and to clarify effects 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.

The initial and final adjudication steps differed considerably in terms of the number of changes considered and also in terms of the kinds of changes considered. At the initial adjudication step, adjudicators made extensive revisions to the target-language questionnaires, and most of the changes focused on improving the individual translations. In contrast, the project director made relatively few changes during the final adjudication step. Some of these changes improved the translations, but several changes were more general, addressing problems observed across all three translations.

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. The review and initial adjudication steps identified a lion's share 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

- 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.

3. Conclusions

In summary, we believe the 5-step translation and evaluation process implemented here produced effective target-language translations for the English-language tobacco use items. Our assessment is based on qualitative observations. Review and cognitive interview results suggest the target-language translations effectively represent the source questionnaire. Successive rounds of review and evaluation produced successively smaller revisions. The translation and evaluation methods we employed supported a collaborative research environment. Thus, the project benefited from the diverse kinds of expertise that individual team members brought to their tasks.

We are currently conducting research to evaluate the 5-step translation and evaluation process more objectively, using behavior coding methods (Cannell, Fowler & Marquis, 1968) to quantify data quality. Our goal is to compare behavior coding results with the review and cognitive interview results reported here to determine whether revisions made based on review and cognitive testing activities actually enhanced survey data quality.

References

Beatty, P. (2004), "The Dynamics of Cognitive Interviewing", In Presser, S., Rothgeb, J.M., Couper, M.P., Lessler, J.T., Martin, E., Martin, J. & Singer, E. (eds.) *Methods for Testing and Evaluating Survey Questionnaires*, New York: Wiley.

Cannell, C. F., Fowler, F. J., & Marquis, K. (1968). The influence of interviewer and respondent psychological and behavioral variables on the reporting in household interviews. *Vital and Health Statistics*, Series 2, No. 26. Washington, DC: U.S. Government Printing Office.

Census Bureau (U.S.) (2004). Census Bureau Guideline: Language translation of data collection instruments and supporting materials. Suitland, MD. Issue date April 5, 2004.

Centers for Disease Control and Prevention (U.S.) (1998). Tobacco use among U.S. Racial/Ethnic Minority Groups – African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, and Hispanics: A report of the Surgeon General. Atlanta, Georgia.

European Social Survey (2002). An outline of ESS translation strategies and procedures. Available at http://naticent02.uuhost.uk.uu.net/ess_docs/translation.pdf. Accessed September 30, 2005.

Forsyth, B.H. & Lessler, J.T. (1991), "Cognitive Laboratory Methods: A Taxonomy", In Biemer, P.P., Groves, R.M., Lyber, L.E., Mathiowetz, N.A. & Sudman, S. (eds.) *Measurement Errors in Surveys*, New York: Wiley.

Harkness, J.A., Pennell, B-E, & Schoua-Glusberg, A. (2004), "Survey Questionnaire Translation and Assessment". In Presser, S., Rothgeb, J.M., Couper, M.P., Lessler, J.T., Martin, E., Martin, J. & Singer, E. (eds.) *Methods for Testing and Evaluating Survey Questionnaires*, New York: Wiley.

Harkness, J.A., Van de Vijver, F.J.R., & Mohler, P. (2003), *Cross-Cultural Survey Methods*, Hoboken, N.J.: Wiley.

Lew, R. & Tanjasiri, S.P. (2003). Slowing the epidemic of tobacco use among Asian Americans and Pacific Islanders. *American Journal of Public Health*, *93*, 764- 768.

Ma, G. Tan, Y., Toubbeh, J.I., Su, X., Shive, S.E., & Lan, Y. (2004). Acculturation and smoking behavior in Asian-American populations. *Health Education Research*, *19*, 615-625.

Ma, G.X., Shive, S., Tan, Y. & Toubbeh, J. (2002). *American Journal of Public Health*, *92*, 1013-1020.

McKay, R.B., Breslow, M.J., Sangster, R.L., Gabbard, S.M., Reynolds, R.W., Nakamoto, J.M., & Tarnai, J. (1996). "Translating survey questionnaires: Lessons learned", *New Directions for Evaluation*, 70, pp. 93-105.

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

Willis, G.B., Lawrence, D., Kudela, M., Levin, K. & Miller, K. (2005a). "The use of cognitive interviewing to study cultural variation in survey response." Paper presented to the Questionnaire Evaluation Standards (QUEST) Workshop.

Willis, G.B., Lawrence, D., Thompson, F., Kudela, M., Levin, K. & Miller, K. (2005b). "The use of cognitive interviewing to evaluate translated survey questions: Lessons learned." Paper presented to the Federal Committee on Statistical Methodology. Washington, DC.