Survey Research Methods (2014)
Vol. 8, No. 3, pp. 181-194

ISSN 1864-3361
http://www.surveymethods.org

Evaluating Multilingual Questionnaires: A Sociolinguistic Perspective

Yuling Pan
Center for Survey Measurement
U.S. Census Bureau

Marissa Fond
Center for Survey Measurement
U.S. Census Bureau

This study aims to develop an assessment tool to evaluate multilingual questionnaires by categorizing the types of translation issues that can lead to measurement errors in cross-cultural surveys. Based on the results of two multilingual projects that cognitively pretested the 2010 U.S. Census questionnaire in five languages and the American Community Survey questionnaire in two languages, we developed a coding scheme guided by sociolinguistic approaches to language and culture to evaluate translated questionnaires by classifying translation issues. In this paper we discuss how the coding scheme is useful in the evaluation of multilingual questionnaires and how it could be integrated productively into the development of such questionnaires and the early rounds of translation. We also suggest feasible solutions to translation issues, to ensure translation quality and achieve not only semantic but functional equivalence across translations.

Keywords: survey translation; translation evaluation; functional equivalence; sociolinguistics; multilingual questionnaire; measurement errors

1 Introduction

With globalization facilitated by technology and the surge of immigration across linguistic and cultural boundaries, government statistical agencies and survey research organizations increasingly realize the need for multilingual questionnaires in social science research. Development of multilingual questionnaires inevitably requires translation of source language questionnaires into target languages at some point in the process. However, reliable translations cannot result simply from the production of translations that are grammatically correct in a target language (Pan & de la Puente, 2005). Rather, a quality translation incorporates the social, cultural, and linguistic elements of each target language to better match respondents' experiences and ensure data quality. This practice is viewed as one of the best practices in the survey translation scholarship community, and research on survey translation (e.g., Forsyth, Kudela, Levin, Lawrence, & Willis, 2007; Harkness, 2003; Pan & de la Puente, 2005) has extensively refined the procedural aspect of the translation-review process. In addition, there is a large and growing research program aimed at evaluating the effectiveness of the result and the practicalities and challenges of its implementation (Dept, Ferrari, & Wäyrynen, 2010; Levin et al., 2009, for a review).

However, one aspect of the cross-linguistic, cross-cultural survey development and evaluation process that has received less attention in the literature is the communication of findings from cognitive testing of translations to decision makers who might not be trained as translators or linguists, and who might not have a clear classification of translation issues and identification of solutions. In this paper we first demonstrate the challenges in evaluating translated survey questions through the method of cognitive pretesting, including the difficulties inherent in advocating for translations that are adapted rather than adopted from source questionnaires. We then argue for the need to develop a translation-review framework guided by sociolinguistic tenants (e.g., Schiffrin, 1988) that makes clear the rationale for flexible survey and survey material translations. The goals of this framework are twofold: first, to classify translation problems revealed by cognitive testing with monolingual survey respondents; and second, to frame these problems in a clear, uncomplicated way that allows translators and survey methodologists to advocate for translation flexibility when negotiating with survey designers and other interested parties who might not be familiar with the languages in question. The final goal is for more functionally equivalent translations to be deployed in the field, resulting in better data quality.

Based on two U.S. Census Bureau projects involving cognitive testing of multilingual questionnaires, we developed a coding scheme to classify translation issues[1] with roots in different linguistic conventions (words and grammar), cross-

Contact information: Yuling Pan, U.S. Census Bureau, 4600 Silver Hill Road, Washington, District of Columbia 20233 (panyuling0822@gmail.com)

[1] We use the term "issue" throughout this paper to describe aspects of translations that were not successful in conveying the intended meaning from the survey designers. This term encompasses typographical errors, linguistic messages and cultural differences,

cultural communication norms (appropriate expressions of a concept), and social practices (knowledge needed to process a concept or to answer a question). The first project involves 112 cognitive interviews conducted to evaluate the 2010 Census questionnaire in five languages. The second one involves 41 cognitive interviews conducted to evaluate the Chinese and Korean translations of the American Community Survey. In addition to classifying the types of translation issues through the coding scheme, we also explore a number of feasible solutions to address the types of translation issues encountered. Thus, this paper discusses how the coding scheme can be useful in evaluating multilingual questionnaires and in effecting change in order to disseminate the best translation possible within institutional constraints.

2 Background

Before we present our study, some discussion of various approaches to survey translation is necessary to highlight the need for systematic evaluation of multilingual questionnaires, and to illustrate why it can be difficult to advocate for flexibility in translation practices. Typically, in the United States federal government system, the first step in developing multilingual questionnaires is the translation of a questionnaire from a source language (English) into several target languages (e.g., Spanish, Chinese). This is not the case in all survey design endeavors; for example, the development process of the European Social Survey includes cross-linguistic and cross-cultural input from inception and incorporates drafts of translations before the source questionnaire is finalized (e.g., Dorer, 2011; Fitzgerald, Widdop, Gray, & Collins, 2011). Also, translatability assessments (Conway, Patrick, Gauchon, & Acquadro, 2010) attempt to identify potential obstacles to cross-cultural translations of questionnaires before the source version is finalized. However, the U. S. federal government statistical agencies rarely develop questionnaires with multiple languages in mind. In practice, survey designers often do not know, when they develop a questionnaire in English, into which language(s) the survey may eventually be translated. Moreover, certain concepts that are mandatory inclusions on many federal surveys (e.g., ethnicity and race reporting) are inherently difficult to translate.

Thus, we address a particular type of problem that arises within the U.S. Census Bureau, the largest statistical organization of the U.S. federal government, and other agencies: that of how to reconcile a translated survey with the source survey within a framework of functional equivalence. This problem has two parts: first, survey materials are developed in English and are typically closed for editing by the time the translation process is initiated; thus, while a translation can be adjusted during the translation review process and pretesting, the source may not be adjusted in tandem. Second, survey sponsors are deeply reluctant to diverge from a translation that appears, at least on a surface level, to mimic the English original; translations must not include information not represented in the source, or be presented in a different layout than the source. Overall, the burden of proof rests with those who advocate adaptation, rather than adoption (see below).

With these constraints in mind, the need for guidance on survey translation and evaluation has been acknowledged by many key statistical agencies, locally based organizations (many in the area of health care delivery and education), and professional researchers in the fields of cross-cultural studies and survey methodology. Over the past few decades, the scholarly community conducting cross-cultural studies has provided useful insights on different approaches to the translation of data collection instruments in multiple languages.[2] The cross-cultural survey literature describes a number of approaches used to produce questionnaires in multiple languages, and we will briefly describe two of the most influential here. The first approach is to *adopt* the questionnaire from the source language into the target language and the second is to *adapt* it.[3]

Adoption calls for the most direct translation of the questionnaire from the source language to the target language without incorporating the flexibility in the survey content that might be required to convey the intended meaning of the question in its entirety, in multiple languages. This approach is based on what is frequently referred to as the "Ask-the-Same-Question" model (see Harkness, 2003), and its goal is to ensure standardization of stimuli in both the source and target languages. This approach is also based on the assumption that a question that seems to be understood by respondents in the source language will also be equally comprehensible in the target language, and it ignores semantic and cultural differences that exist across languages.

The second approach, *adaptation*, also uses the source questionnaire as the base, but allows for components of the survey questions to be extensively modified, in ways independent of the inevitable changes that result from the translation from the source to the target language, in order to make the questionnaire interpretable and meaningful in the target language. The adaptation approach can help to en-

---

and respondent errors that relate to the usability of the survey instrument and/or format. We believe that issue is a more accurate descriptor than "error" because there is no ascription of blame, to the survey designers or the respondents, in this paper.

[2]Harkness (2003) provide an extensive review of cross-cultural survey methodology. Also, the work of Behling and Law (2000) serves as an example of how the translation of data collection instruments is explicitly addressed, and other guidance can be found in specific case studies of survey translation, for example, McKay et al. (1996), Potaka and Cochrane (2002), and Schoua-Glusberg (1992).

[3]We emphasize that in these two approaches outlined, the source questionnaire is developed before the target language questionnaire; they are not developed simultaneously.

sure that survey questions (in both the source and target languages) measure the same or similar constructs and avoid concept biases, and therefore achieve *functional equivalence*; this means that the instrument conveys the meanings of questions, instructions, and response options comparably across versions, even if they differ in length or other characteristics. Such questionnaires are more likely to provide reliable, complete, accurate, and culturally-appropriate information than instruments developed using other techniques (Behling & Law, 2000; McKay et al., 1996).

The adaptation approach, with its inherent flexibility, focuses on the functional equivalence of survey questions and has gained attention from survey researchers when it comes to the actual task of translation (see e.g., Harkness, 2003). Adaptation of translation can be performed using a variety of methods or techniques.[4] Recent literature in cross-cultural studies advocates the use of the "committee approach" (Forsyth et al., 2007; Harkness, 2003; McKay et al., 1996; Schoua-Glusberg, 1992) as a technique for adaptation. The committee approach is more comprehensive and collaborative because it relies on input from a team whose members (subject-matter experts, survey designers, and researchers) have skills that supplement those of a translator.

An additional step is called for by the TRAPD[5] approach (Harkness, 2003), a specific type of committee approach: the pretesting of the translated instrument. In this approach, several translators independently translate the instrument from the source language to the target language. Then the translators, the translation reviewers, and other members of the team discuss the translated versions of the instrument. A reconciled version of the translated instrument is produced, and this data collection instrument is then pretested. After the pretesting is complete, the adjudicator and other committee members convene again to finalize the instrument. The TRAPD committee approach is the foundation of several approaches to cross-cultural translation. Translation verification (e.g., Dept et al., 2010) addresses the need for quality assurance during the translation process, providing a meticulously detailed review procedure that documents problems at each phase of translation. The U.S. Census Bureau Translation Guidelines recommend this committee approach as well (Pan & de la Puente, 2005). Because the source questionnaires (in English) often cannot be modified by the time they are sent out for translation, the Guidelines require that every translation team assigned to produce final versions of Census Bureau translated questionnaires and supporting materials must involve five different groups of professionals to produce the best possible translation: translators, reviewers, subject matter experts, survey methodologists, and adjudicators. The process follows includes five steps: prepare, translate, pretest, revise, and document. The steps and the actors involved are summarized in the table 1.

Table 1
U.S. Census Bureau Translation Guidelines Team Process

Steps	Actors
1 Prepare	
The questionnaire is conceptualized and designed in the source language.	Survey methodologists, designers, survey sponsors or organizations[a]
2 Translate	
The source questionnaire is translated using the committee approach.	Translators
3 Pretest	
The translated questionnaire is cognitively tested with monolingual speakers of the target language.	Language experts
4 Revise	
The translated questionnaire is revised based on the cognitive testing results.	Translators, reviewers, adjudicators
5 Document	
Approved changes are recorded so that they can be referred to for future surveys.	Translators, language experts, survey methodologists

[a] For example, in the United States an organization sponsoring a survey carried out by the Census Bureau might be the Department of Housing and Urban Development.

As we alluded to earlier, research on survey translation has focused primarily on the translation process and techniques for the performance of translation, rather than the evaluation of the result, or the principled incorporation of the pretesting results into revisions of the translations (steps 3 to 4 in table 1) in consultation with survey designers who are responsible for final decisions regarding content. While most agree that achieving functional equivalence is the goal of survey translation, achieving this goal can be deceptively complicated; and it can be difficult for translators and language experts to communicate their findings and recommendations to the survey sponsors who are not part of the process. Detailed evaluation systems have been developed for translators to use in their work; Levin et al. (2009), in a thorough review of work on cross-cultural cognitive interviewing methods for pretesting surveys, note that the range of issues identified in multi-

[4] Some of the most commonly used techniques are "simple direct translation," "modified direct translation," "back translation," and "committee approach." See Harkness (2003), Behling and Law (2000), and Pan and de la Puente (2005) for details.

[5] TRAPD stands for Translation, Review, Adjudication, Pretesting and Documentation.

lingual pretesting is extensive (p. 14). They summarize six classification systems that researchers have used to explain the problems that were uncovered in translation pretesting, and while the systems differ in the descriptions or operationalizations of each category, they all include references to linguistic or "translation" issues, problems that are "culturally" based, and issues with survey navigation generally (e.g., Carrasco, 2003; Goerman & Caspar, 2007; Harkness, 2003; Schoua-Glusberg, 2006; Willis et al., 2008; Willis & Zahnd, 2007).

While these research efforts have fomented a systematic and complete examination of translation issues, there is still a need to communicate these issues effectively to decision makers (i.e., survey designers and sponsors). Perhaps more pressing is the need to design a coding scheme that simply and clearly translates evaluations of translations into compelling arguments that survey sponsors can take into account as they make final decisions about approvals of translated documents.

Keeping in mind the need to communicate translation issues uncovered in cognitive interview pretesting across different groups (survey developers/designers, sponsors, translators, etc.), we attempt to address the following research questions in this study: (a) How can we determine the acceptability of a translation of an English-language source questionnaire? (b) When translation issues are identified in the expert review or cognitive pretesting process, how can we best articulate and describe the nature of the problems? (c) Can patterns of inadequacies in translated materials be identified so that they can be anticipated and addressed efficiently? (d) How can language experts articulate the scope of the translation problems encountered and effectively argue for flexibility in translation to survey sponsors? (e) What are the feasible and effective solutions to the problems identified?

3 Identification of Functional Equivalence

Sound, effective survey translation requires that a translation function at micro- and macrolinguistic levels. The microlinguistic level concerns the word choice and sentence structure in a translation, meaning that the translation should consist of accurate, appropriate wording and use the correct terminology to convey the meaning of the source text. Moreover, the translation should follow the grammatical structure of the target language, and sound natural to a native speaker of that language. The macrolinguistic level is oftentimes referred to as the pragmatic level of translation (Pan, Landreth, Hinsdale, Park, & Schoua-Glusberg, 2007), which is the frames of reference, or schemata, that respondents rely upon to interpret a translated item. This involves the social and cultural contexts as well as background knowledge or experience and communication norms common among speakers of the target language.

In order to ensure that translated questionnaires are appropriate at the pragmatic or "functional" level as well as the microlinguistic level of lexicon and syntax, we need to promote one basic principle in sociolinguistics: that a language is inseparable from the culture and society in which it is used (Gumperz, 1999; Holmes, 1992; Tannen, 2005). Language use inevitably reflects, and perpetuates, the values and social practices of a given culture. To tackle problems in translation, we need to attend to not only linguistic rules that govern the sentence structure or word order of a specific language, but also the cultural norms of expressing certain concepts and the social practices encoded in linguistic expressions. Thus, there are three components that we need to consider in our analysis: *Linguistic Rules*, *Cultural Norms*, and *Social Practices*:

- Linguistic Rules refer to language-specific rules, such as the grammar, the word order, or internal sentence structure of a language or code. These issues are identified at the word or clause level.
- Cultural Norms refer to the ways of doing certain things in a given culture, such as communication style, the discourse sequences for presenting information, and culture-specific ways of showing politeness. This category allows us to examine language in use, and how a communicative event unfolds (e.g., the question-answer rhythm of a survey).
- Social Practices, in contrast to Cultural Norms, refer to daily or institutionalized practices in a society, including social institutions, educational systems, or personal experiences as influenced by culture and society. For example, survey interviewing may be a common practice in American culture, but it could be a foreign concept to people who are from China or Vietnam (Pan, 2008; Pan, Sha, Schoua-Glusberg, & Park, 2009).

As language encodes the cultural values and salient social practices of a particular cultural group, and language use is always a reflection of cultural norms and social knowledge (Gumperz, 2001), the three components of Linguistic Rules, Cultural Norms, and Social Practices can serve as the guiding principles for us to evaluate the quality of a translation and advocate for translations that are functionally equivalent to the source.

Our approach distinguishes between cultural and social factors in language use (see e.g., Schiffrin, 1988) in evaluating multilingual questionnaires, because these two aspects of language require different resolutions from translation teams and survey designers, as we will describe below.

3.1 Development of the Coding Scheme: Method

The coding scheme was based on findings from two cognitive testing research projects: 1) the 2010 U.S. Census questionnaire in five languages (English, Chinese, Korean, Vietnamese, and Russian) and 2) the American Community Survey in two languages (Chinese and Korean). The first

project involved the 2010 U.S. Census questionnaire, a self-administered survey with 10 basic demographic questions. First, the entire questionnaire was translated from English into the four non-English target languages. That is, the multilingual versions were not developed simultaneously with the English; they were developed after the English version was finalized. Then, a total of 112 cognitive interviews were conducted with monolingual speakers of English and the four non-English language groups. English was included in the cognitive testing as a baseline for comparison. Cognitive interviewing, referenced above, is a semi-structured interview method used to "study the manner in which targeted audiences understand, mentally process, and respond to the materials" provided by survey researchers (Willis, 2005, p. 3). Based on the summary of findings from this first project, we developed the coding scheme to code the translation problems observed.

The methodology in the cognitive testing procedures involved the following steps that ensured a systematic evaluation of the translated questionnaires: 1) a team of three language experts (native speakers with experience in cross-cultural survey methodology and cognitive testing) was assembled for each language, and these team members went through a two-day training program in cognitive interviewing and project-specific requirements; 2) using a committee approach, the language teams translated the cognitive interview protocol, which was developed by Census Bureau survey methodologists, from English into the target languages; 3) the language teams conducted cognitive interviews with respondents representing a range of demographic characteristics who spoke each target language; 4) after the first round of interviews, the language teams met to identify problematic translations, suggest alternative translations, and articulate their justifications for their suggestions, based on the findings of the interviews. The suggested changes were tested in the second round of interviews; 5) finally, after all of the cognitive interviews across the languages were completed and summarized, the language teams met once more to reassess the translation of the census questionnaire, and suggested final recommendations for alternative translations, based on the outcomes of the interviews (Pan et al., 2007).

After the cognitive testing was completed, we analyzed the results documented by the language experts. In the cognitive interview summaries produced for each interview, each issue uncovered was noted and explained by the language experts, who then offered recommendations for possible solutions based on their linguistic and cultural knowledge and expertise. Working with the explanations of the issues and the suggested solutions, it became apparent that while most translated census questions and instructions could be considered grammatically correct and included words that, in isolation, were familiar to respondents, the respondents demonstrated difficulties understanding the aggregate, overall intended meaning of the questions and providing answers that felt satisfactory to them. We then classified each difficulty according to whether the problem pertained most significantly to issues in Linguistic Rules, Cultural Norms, or Social Practices.

In addition, we noticed that there were some problems that arose from production errors like typographical errors or accidentally-omitted words. We also noticed that some issues were caused by respondents not paying attention to the task and making mistakes in answering the questions, which the respondents themselves later noted or corrected. To address these issues, we added the following two categories: Production Errors and Respondent Errors. We introduce these as separate from the categories of Linguistic Rules, Cultural Norms, and Social Practices because they are qualitatively different; they represent two types of errors or "mistakes" that are easily corrected with a second review. For example, a misspelled word that language experts consider a clear typographical error, and not a possible alternative spelling, would be grouped under Production Errors. A respondent who answered a question incorrectly because she omitted a word while reading the question because she was momentarily distracted would have committed a Respondent Error. In this situation, if the respondent is asked to read the question again, they do so without any errors. We found it necessary to record these categories of errors, rather than dismiss them outright as easily remedied, because in the event that we found more Production or Respondent Errors in a particular section of the survey, within a single translation or across multiple languages, these issues would be worthwhile to address in subsequent usability testing.

3.2 The Coding Scheme

The completed coding scheme includes the three main categories listed above that reflect the components of Linguistic Rules, Cultural Norms, and Social Practices, as well as Production Errors and Respondent Errors. We briefly summarize the categories in table 2.

How the coding scheme works. Through the examples of issues uncovered in the review of the pretesting of the 2010 U.S. Census questionnaire that follow in this section, we demonstrate how the coding scheme works. In order to reach the goal of producing appropriate and accurate translations, it is necessary to understand the unique properties of each individual issue that arises in a given language, in a given questionnaire. These examples are presented for the purpose of demonstrating how we coded translation issues, based on cognitive interview summaries from the language teams; they are intended to be illustrative, not exhaustive.

Coding of linguistic rules issues. Based on cognitive interview summaries, we identified evidence that indicated problems with respondents' understanding of the translated questions due to the subtle differences in word usage or sen-

Table 2
The Coding Scheme

Codes	Explanations
Linguistic Rules	This category classifies issues in a translation that are due to off-target lexical items (e.g., antiquated or obsolete terminology for which modern synonyms exist), overly complex or incorrect morphology, incorrect or unnatural syntax (e.g., syntax that mirrors English patterns rather than the word order of the target language), etc.
Cultural Norms	This category refers to issues that arise when concepts that are expressed one way in English (the source language of the survey) are expressed in a different way in the target language (e.g., personal address conventions, numbering/counting, kinship terms, time references, conversational norms, etc.).
Social Practices	this category classifies issues with concepts that can be described in English but cannot be translated into the target language because either the concept does not exist in that culture, or respondents have no experience with the concept (e.g., educational systems; while the general concept of "education" may exist cross-culturally, certain subtypes like "home schooling" may not).
Production Errors	this category refers to survey production problems that are simple mistakes (e.g., word omissions, typographical errors) that can be easily corrected.
Respondent Errors	this code refers to actions taken by respondents while reading or answering a questionnaire that they themselves identify to be easily-corrected mistakes (e.g., those caused by inattention).

Were there any additional people staying here April 1, 2010 that you did not include in Question 1?
Mark ☒ all that apply
☐ Children such as newborn babies or foster children
☐ Relatives, such as adult children, cousins, or in laws
☐ Nonrelatives, such as roommates or live-in baby sitters
☐ People staying here temporarily
☐ No additional people

Figure 1. Question 2 from the 2010 U.S. Census

tence structure between the target language and the English language rules. An example to illustrate this is shown in Figure 1.

Regarding this question, there were translation issues in three of the four languages that were classified as related to Linguistic Rules. In Korean, "live-in baby sitter," seen here as part of the third response option, was translated as "always-staying baby sitter," which sounded as awkward, and confusing, in Korean as it does in English. The concept of a live-in nanny or baby sitter does exist in Korean culture, per interviews with respondents and language experts, thus the problems that respondents had with "always-staying baby sitter" indicate that the lexical items chosen to translate the term were not optimal. While a respondent could perhaps guess what "always-staying baby sitter" might mean, the phrase does not convey its intended meaning as clearly as necessary. In Vietnamese, the translation of the question included repetitive auxiliary verbs, an overly-complex and convoluted syntactic construction that obscured the meaning of the question, in particular the reference time of April 1, 2010; the auxiliary verbs needed to be deleted and the syntax clarified to make the sentence sound natural and the meaning clearer in Vietnamese. These are issues classified as Linguistic Rules issues because they implicate lexical items and syntax; the resulting awkwardness of the translations was due to the failure to adapt the English original to the linguistic features of the target language, rather than unfamiliar concepts or usage conventions.

It is important to bear in mind that Linguistic Rules issues are not necessarily errors. The translations may be grammatically correct, but they are not natural in the target language, in that the translations sound to respondents like translations, rather than materials written in their language. Put another way, these issues are examples of words or phrases that respondents would not be likely to hear or read in their language. These issues can hinder respondents' comprehension of the translated questions and answer categories.

Coding of cultural norms issues. The problems in this category usually resulted from the different ways of expressing a similar concept in different cultures. This means that a concept included in the English questionnaire exists in the target culture, but due to different conventions of expression or a different focus on certain values in American culture and the target culture, the concept is conveyed differently through language. For another example, see Question 7 in Figure 2:

An example of a Cultural Norms issue was present in the translation of this question into Korean. More than two thirds of Korean speakers experienced confusion writing their age because the Korean convention of counting age is different from the American one. In Korean culture, newborn babies are considered one year old, so someone who would be 50 years old as Americans count age would be considered 51 by Korean speakers. In order to address this issue, the translation must specify that the form is asking about the American

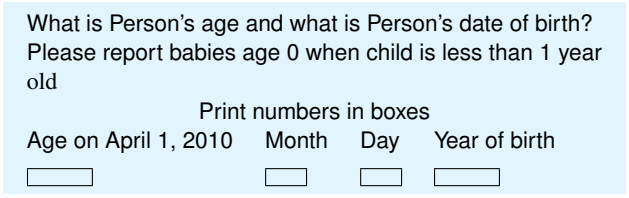

Figure 2. Question 7 from the 2010 U.S. Census

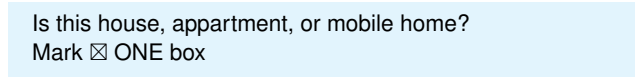

Figure 3. Question 3 from the 2010 U.S. Census

way of indicating age. This is a Cultural Norms issue because the Korean speakers' interpretation arises not from the lexical items used, or the syntax of the translation, but from the cultural background of the speakers that allows them to interpret the code; this means that when they respond to the question, they are answering based on an interpretation of the question that was not intended by the American English-speaking survey designers.

Coding of social practices issues. There were many examples of questions on the 2010 U.S. Census questionnaire that gave rise to Social Practices issues. In Question 2 cited above in Figure 1, the term "foster children" proved problematic for Chinese, Korean, and Vietnamese speakers. There are existing terms similar to "foster children" in these languages, and so these terms were used in the translations, but the terms refer to concepts that are quite different from the concept of "foster children" in the United States. In these languages, the terms for "foster children" were understood as meaning something closer to "children under the temporary care of relatives or friends." The concept of a foster program administered and supported by the government was not retained in the translation. This type of program does not exist in China, Korea, or Vietnam, so there is no lexical item in these languages that can be used to indicate it; therefore, a more descriptive phrase that provides additional information is necessary.

Another example of a concept that appeared to be uniquely American is found in Question 3 (Fig. 3). Question 3 provides response options relating to the ownership of the residence, but the question itself proved to be difficult. Speakers of all four tested languages were unsure about what a "mobile home" was, as mobile homes are not common in other countries (and in fact, speakers of other languages such as Spanish will often adopt the American word "trailer" rather than describe the concept in their native language). A final example of a concept that could not be readily translated into another language clearly and unambiguously was "nursing home," found in the response options for Question 10 (Fig. 4).

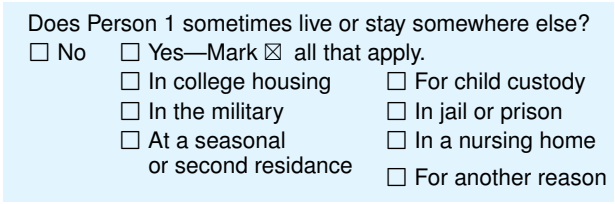

Figure 4. Question 7 from the 2010 U.S. Census

For the term "nursing home," the approximate translations were found to be inaccurate (or inadequate). The target languages have terms similar to "nursing home," but the meanings are different from the American concept. In Chinese, the translation meant, to various speakers, either a mental hospital (to Hong Kong Chinese) or a recreational resort (to Mainland Chinese); in Korean the translation meant a resting place; in Russian it meant a medical establishment; and in Vietnamese it meant a luxurious resort. Given that these cultures do not have experience with American-style nursing homes, the translations were unable to capture that meaning. A longer and more descriptive phrase is necessary in order to make sure that the individual filling out the questionnaire understands the question as it is intended.

Identifying overall patterns of problems. Once all of these issues were identified and categorized according to the coding scheme, it was possible to quantify them in order to determine which types of issues were most problematic in the 2010 U.S. Census translations. The chart below (Fig. 5) summarizes the findings of types of issues identified in the four target languages (excluding the English source material) during the cognitive testing process. There are four main types of issues coded: Linguistic Rules, Cultural Norms, Social Practices, and Production Errors. We did not find any Respondent Errors in this phase of testing, because the analysis was conducted using a composite summary of the issues in each language, rather than summaries of individual interviews with respondents. Any respondent difficulties were likely not significant enough to be included in the summary report; this may be explained by the fact that the 2010 Census questionnaire contains only ten demographic questions, and it is relatively simple in terms of form navigation.

Regarding these four types of issues, it quickly became apparent that Linguistic Rules and Cultural Norms issues were the most common, as they comprised 33% and 39% of the total 163 issues discovered, respectively. The remaining issues were divided between Social Practices and Production Errors, with 18% and 10% respectively.

Next, in order to deepen our understanding of the translation issues and how the coding scheme works, we wanted to determine whether each of the four target languages seems to have the same pattern of issues identified, or not. The chart below (Fig. 6) shows that in Chinese, the Linguistic

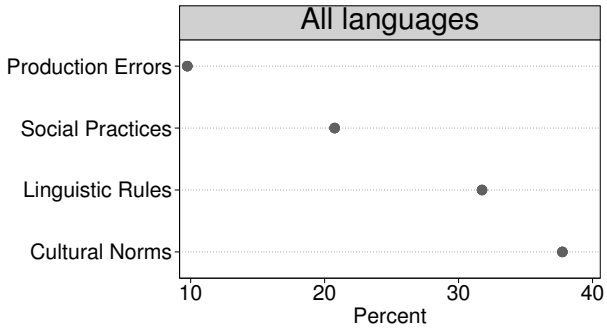

Figure 5. Types of Issues in all Four Languages for the 2010 Census Questionnaire

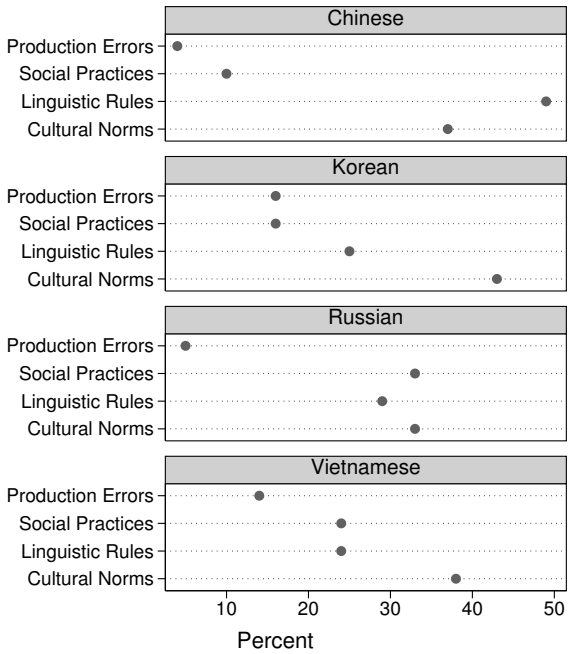

Figure 6. Issues as a Percentage of Total (Per Language) in the 2010 Census Questionnaire

Rules category comprises a much larger percentage of total errors. Cultural Norms issues present the biggest challenge for Korean and Vietnamese translations. Social Practices and Cultural Norms issues were equally difficult for Russian translation. In addition, Social Practices issues were identified across the four translations; in other words, concepts that were considered unique to American life were equally difficult to translate into Chinese, Korean, Vietnamese, and Russian.

After coding the translation issues into these categories and presenting them by language groups, we can use this information to focus our attention on explaining why this might be by discussing with the language experts and by conducting follow-up research. For example, the Chinese translation exhibited a particularly high rate of Linguistic Rules issues due to overly-complex syntactic structure. Discussion with the Chinese language experts in our team suggested that the complex syntax in Chinese translation resulted from word-for-word translations that mirrored English syntactic structures rather than adaptations to more typical Chinese structures. Follow-up research on Korean, Russian, and Vietnamese communication norms also gave us insights on cultural differences between English and target languages, and how to revise the translations to make it more culturally appropriate. Issues with Social Practices indicate that there might be elements of the survey questions, designed with English speakers in mind, that are difficult to translate effectively.

This project showed that the coding scheme worked well not only to identify what types of causes were at the root of translation issues, but also to show how the translations into different languages contained different levels and proportions of problems, as well as some similarities. This suggests that different languages might require different types of linguistic and cultural adaptations, and different skill levels managing such adaptations during the conduct of translation.

4 Solutions for Translation Problems

After the coding scheme was developed using data from the pretesting of the 2010 Census questionnaire, the coding scheme was applied to a different survey pretesting endeavour (a study of the American Community Survey) in order to examine how the scheme worked in a more detailed analysis of cognitive interview data, and to explore possible solutions for the translation issues uncovered.

4.1 Using the Coding Scheme with the American Community Survey

In this phase of the project, we applied the coding scheme to the cognitive testing of the Chinese and Korean translations of the American Community Survey (ACS) questionnaire. The ACS is the largest general survey conducted by the U.S. Census Bureau, and it covers many topics, such as demographic characteristics, housing, health insurance, education, income, and transportation. The cognitive testing project reported here was conducted as part of a larger project undertaken by the Census Bureau to evaluate the many subsections of the translated ACS questionnaire (which is too extensive to be thoroughly reviewed in one round of cognitive testing and analysis of results). The portion tested for this study included all of the instructions on how to complete the questionnaire, basic demographic questions, and the section of questions on housing characteristics.

Similar to the pretesting of the 2010 Census questionnaire, the ACS Chinese and Korean pretesting project followed the same methodology for carrying out the cognitive interviews

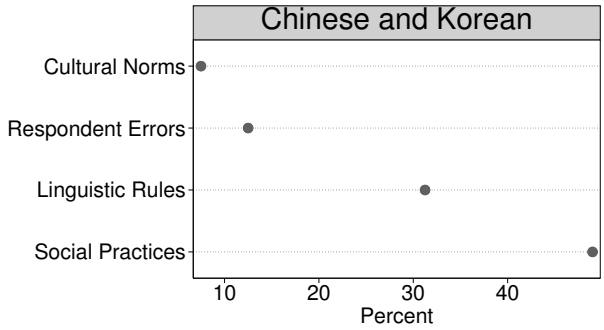

Figure 7. Types of Issues in Chinese and Korean in the ACS Questionnaire

and analysis. A total of 41 interviews were conducted (19 in Chinese, 22 in Korean). The analysis procedure for both multilingual projects was similar, but there were a few differences. For the 2010 Census questionnaire project, the results from the cognitive interviews, which were written up in a summary report by the language teams, were analyzed by Census Bureau researchers. Each issue uncovered in the interviews was noted and explained by the language experts, who then offered recommendations for possible solutions, based on their linguistic and cultural knowledge. Working with the explanations of the issues and the suggested solutions, Census Bureau researchers coded each issue. This review process was important to developing descriptions of each of these categories, and determining how clearly the issues fit into one or more of them; the coding scheme was operationalized, with sharper divisions between categories, as a result of this process.

The analysis of the ACS interviews proceeded slightly differently. Because the coding scheme was ready to be used after the completion of the 2010 Census project, language experts were able to use it while they were administering the protocol to each respondent and while writing up their summaries of the interviews. Therefore, the respondents' answers to each cognitive probe within the protocol were coded by the language expert, who also provided an explanation of the code chosen. Then, a Census Bureau researcher reviewed each cognitive interview summary, as well as the codes and justifications for each issue, and confirmed the assessment (or in some cases, queried the code if the explanation did not seem to match the code assigned). In this way, each issue was reviewed by two coders: one language expert and one research analyst.

4.2 Uncovering Overall Patterns of Translation Issues

In applying the coding scheme to evaluate the Chinese and Korean translations of the ACS questionnaire, we see the distribution of the types of issues that were uncovered in the 41 cognitive interviews in Figure 7.

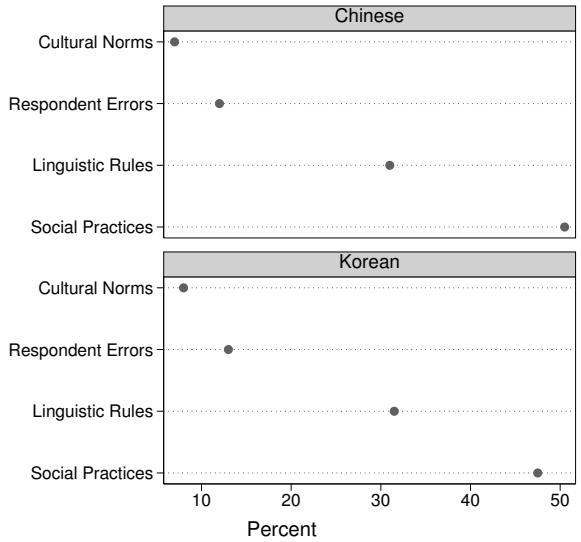

Figure 8. Issues as a Percentage of Total (Per Language) in the ACS Questionnaire

From this chart, we can immediately notice that in this ACS study, Social Practices issues constitute the majority of the issues (49%) followed by the Linguistic Rules issues (31%). The predominance of Social Practices issues is not unexpected, because the translations used in this study had undergone numerous rounds of translation review before the cognitive interviews were conducted. Even when translation protocols are followed closely, Social Practices issues are the most difficult to resolve through translation reviews, so it is not surprising that most of the problems found involved these issues.[6] As for the Linguistic Rules issues, respondent comments indicated that these translations tended to use long and complex sentences that mirror the original English sentences without taking into consideration Chinese or Korean language-specific structures, and employed some outdated as well as high-register terms (e.g., terms more familiar to highly-educated respondents). No Production Errors were identified in the cognitive testing process, due to the fact that these translations were very carefully reviewed numerous times prior to the cognitive testing, in contrast to the 2010 Census questionnaire translations.[7] Small propor-

[6]This is posited because Social Practices issues identify concepts that do not exist or are not closely equivalent in English and in the non-English target languages. In fact, terminology for certain concepts might differ in different countries in which the same language is spoken. Thus, translating such terms is challenging because respondents' social experiences are diverse.

[7]Regarding the different translation processes for the 2010 Census and the ACS, for both surveys a translation company was contracted to complete the translation. After they did so, the 2010 Census was cognitively tested by a research company contracted by the

tions of Respondent Errors (13%) and Cultural Norms issues (7%) were identified.

In figure 8, we can see the issues as a percentage of the total number of issues uncovered in each language for the ACS questionnaire translation.

In this chart, it is clear that the proportions of issues in Chinese and Korean are similar. This was expected because the translations had gone through expert review many times and the differences in quality among translation teams (which can be significant) were tempered; also, most of the issues that remain are Social Practices issues common to both Chinese and Korean societies (which do not have simple translation fixes) and residual Linguistic Rules issues (e.g., complex sentences or vocabulary that respondents interpreted differently from the translator's intention).[8]

Trying to uncover patterns in respondents' comments during cognitive interviews is important because if we see that multiple respondents interpreted a question in a way that was not intended by the survey designers, then we know that there is a serious deficiency in the translation. Also, if we see that the respondents interpreted the question in a way that was not intended because of the same general reason, then this not only makes solving the problem easier, but it adds to the bank of information that a survey research organization has about what works in survey questions.

This is why the ability to quantify our results is so important. We are able to see, in one glance, where the most severe problems lie. We can identify the scope of the problem in a given language, or in a given question across languages, allowing us to flag problematic questions for follow-up, as well as to determine if the translation issue was unique to a single language or culture or if it might even be related to the English original (because the English versions are constantly being revised as a result of cognitive interview data). The quantitative results are an important additional tool to aid in translation review; they are not significant in themselves. This is to say that for survey methodologists working on cross-cultural surveys, it is not necessary to know whether, for example, the difference between the number of Linguistic Rules issues and Social Practices issues in the Chinese version of the ACS was statistically significant, but it is interesting to know which questions had the most Social Practices issues so that they can be modified.

4.3 Resolving Translation Issues, Advocating for Flexibility

In addition to identifying issues, which a number of coding schemes do quite thoroughly (Levin et al., 2009), the coding scheme can also help to expedite the process of fixing questions that proved to be problematic by engendering productive conversation with survey designers and sponsors. For example, Production Errors require careful review and edit, but they do not need additional cognitive pretesting or new translations; these are uncontroversial changes to make. A team-based approach to translation review and careful review procedures involving multiple parties are important for catching these problems early in the process.

Is there a business (such as a store or a barber shop) or a medical office on this property?
☐ Yes
☐ No

Figure 9. Question 6 from the ACS Questionnaire

For Linguistic Rules issues, this is the category that requires a close examination of the subtle meanings associated with a term and attention to whether the translation sounds natural, in the sense that it uses the syntactic structure of the target language instead of that of English. Oftentimes, translations tend to follow the English sentence structure too closely, which can create unnecessary comprehension difficulties for non-English speakers. Figure 9 is an example of the Chinese translation of Question 6 in the ACS questionnaire which has two issues in the Linguistic Rules category.

In this translation, there are two issues with Linguistic Rules: the translated term for "property" and the term for "business." The respondents did not grasp the intended meaning of these terms in the context of the survey question. The translated term "property" was interpreted by Chinese respondents as "real estate property for sale" or "realtor's office." The translated term "business" implied a large-scale multinational business rather than a small local business. Therefore, the two terms were coded as Linguistic Rules issues. Based on these findings, the Chinese language team suggested another term "housing unit and yard land" to translate the English term "property" to lessen the confusion. For the term "business," the Chinese language team suggested restructuring the translation to say "*do* business" instead of "*a* business." This change required restructuring the question's syntax and using a verb "do business" instead of a noun phrase "a business." This solution fixed the problem, and the new structured sentence sounded natural in Chinese, based on the language experts' opinion and the respondents' feedback.

This example demonstrates a way in which classifying

Census Bureau. However, in the case of the ACS, the research company hired to perform the cognitive testing conducted an additional review and revision of the translation before the testing commenced. For this reason, the ACS was more thoroughly reviewed by a diverse range of language experts.

[8]Linguistic Rules, while less of a translation obstacle than Social Practices that have no clear solution, remain persistent for a variety of possible reasons. In particular in the ACS, the questions tend to be complex and multi-part, meaning that respondents might have some degree of difficulty processing the language of the questions and response options regardless.

Is this house, apartment, or mobile home part of a condominium?
☐ Yes → What is the monthly condominium fee?
For renters, answer only if you pay the condominium fee in addition to your rent; otherwise, mark the "None" box.
Monthly amount – Dollars
☐
or
☐ None
☐ No

Figure 10. Question 13 from the ACS Housing Section

translation issues through the coding scheme can provide a direction of how to find solutions, and the solutions become feasible and more systematic. By coding the issues as related to Linguistic Rules, based on interviews with respondents and input from language experts, the scope of the translation revision was refined. In this case, respondents understood the general concepts under discussion, and had other ways of describing them, but the problems arose in the particular terms used in the survey. Thus, translators need to search for different terminology that would be more commonly understood, rather than to attempt to add additional content to the survey in order to explain a foreign concept. In addition, coding these lexical items as Linguistic Rules issues allowed the language experts to explain the scope of the changes that needed to be made to the translation, and assure the survey designers and other stakeholders that while changes were necessary to avoid off-target responses, they could be handled within the scope of the concepts that were already present in the source questions (i.e., no additional content was required).

For Cultural Norms issues, it is necessary to ask whether a certain concept exists in the target culture, and if it does, how it is expressed. These questions will help to identify early on what the cause of the translation difficulty is. The goal is to identify and use culturally- appropriate expressions, employ culture-specific communication styles, translate the discourse structure of English into that of the target language, and incorporate politeness strategies where appropriate. For example, due to the different politeness practices between American and Korean cultures, the Korean translation of the instructions in the ACS questionnaire was found to be too direct, which violated the politeness norms of Korean culture (Pan, 2011; Pan, Landreth, Hinsdale, Park, & Schoua-Glusberg, 2010). In one instruction, seen here in Figure 10 incorporated into the "yes" response option, the English wording is: "For renters, answer only if you pay the condominium fee in addition to your rent."

The original Korean translation of this instruction reads: "Answer only if you rented and pay money in addition to your rent." This translation was correct and followed the politeness norms in the English original by using a direct expression of a command. Based on reactions from the Korean respondents in cognitive interviews, this was coded as an issue in the Cultural Norms category for a lack of appropriate polite expressions and awkward discourse structure. The revised Korean translation becomes: "If you are renting a condominium, please write the amount of condominium fee that you pay in addition to your rent." The revised translation took into account the necessary expression of politeness ("please") and re-structured the sentence so it flowed better in the Korean language (by placing the "if" clause at the beginning, before the command). In this case, the language experts were able to advocate for a change to the translation that, while not extensive, might appear unnecessary because the terms were correct and familiar to respondents. Within this simple coding scheme, however, the politeness and discourse structure issues could be positioned as equal in importance to issues that relate more directly to content or concepts.

How many acres is this house or mobile home on?
☐ Less than 1 acre →SKIP to question 6
☐ 1 to 9.9 acres
☐ 10 or more acres

Figure 11. Question 4 from the ACS Housing Section

While issues in Linguistic Rules and Cultural Norms categories can hinder comprehension and increase difficulty for respondents, issues in the Social Practices category indicate that a construct being measured in the questionnaire may be a foreign concept to the target population. This can very likely lead to measurement errors or item non-response. For Social Practices issues, it is important to ask if the translated question might measure a concept or experience that respondents have no knowledge of. If a certain concept or practice does not exist in the target culture, how can the concept or practice be translated, and how can respondents come to understand the new concept quickly and clearly? We may need to think of creative, descriptive ways to translate the concept, or it may be necessary to revisit the source materials or source questionnaires to collect as much background and contextual information as possible. From there, translations can include explanations, examples, or notes, plus clear instructions, which are culturally appropriate and helpful to speakers of target languages. Also, we recommend flexibility when translating Social Practices issues, and we encourage the use of descriptive phrases instead of existing terminology. As an example, see the ACS Question 4 about land measurement (Fig. 11).

The land measure "acre" is not used in Chinese- or Korean-speaking countries, so "acre" is not a salient concept. The idea of land measurement exists, but the specific unit "acre" is not in the Chinese or Korean measuring system. To

Table 3
The Coding Scheme: Solutions Proposed

Codes	Examples of Solutions
Linguistic Rules	Changes in sentence structure are necessary to promote understanding among native speakers of the target language(s); syntax that matches the English source questionnaire is not effective. Lexical items must not be cognates of English terms if they are not understood in the target language(s); rather, different terminology must replace these terms.
Cultural Norms	While the concepts discussed might be accurate and understood by native speakers of the target language(s), questions must be presented in culturally-appropriate ways, using politeness markers, introducing topics according to cultural conventions, and indicating instances in which cultural schemata might not match.
Social Practices	These concepts and questions cannot be better translated through additional rounds of review and pretesting; the concepts do not exist in the target society, thus greater flexibility is required to convey meaning in the target language(s). To ensure data quality, significant changes to question wording or formatting might be required to provide the background necessary for understanding.

overcome this Social Practices issue, the solution was to add supporting information appropriate for each target language. In Chinese, a note was included that read "One acre is about 4,000 square meters" (as meters are commonly used to describe land area in Chinese). In Korean, the supporting note read "One acre is about 1,230 pyeong" (as the measurement unit "pyeong" is unique to Korean, and commonly used).

Social practices issues are perhaps the most crucial obstacles to overcome in survey translation, and these issues require the most radical changes in order to achieve functional equivalence. Characterizing Social Practices issues in contrast to Linguistic Rules and Cultural Norms makes the need for flexibility in translating these concepts clearer and more compelling.

5 Conclusion

One of the problems faced in the survey translation process is the evaluation of translated surveys in terms of functional equivalence. A translation is "successful" when it is functionally equivalent to the source questionnaire, and this means that not only is the information presented accurately, but that it is understood as intended. The translation process used in survey research so far can address the former, but not always the latter. The latter can be determined through cognitive testing with respondents so that survey methodologists can understand better how their questions are being interpreted; the next critical step is to aggregate those results so that a translation can be evaluated at a glance, by individuals who may not speak the language(s) into which the survey was translated.

However, we need a way to describe the results of the cognitive interviews that takes all insights from all interviews conducted into account, and convey these insights to survey sponsors who may be reluctant—due to cost, or version equivalency, etc.—to approve changes in translations that may promote functional equivalence, but cause the versions of a survey to appear different from one another. Convincing sponsors of the need for functional equivalence is a challenge, and this simple coding scheme aims to address this challenge. It is not useful to have dozens of cognitive interviews summarized; they must be interpreted and translated into actionable problems that individuals who are not speakers of the target languages, or who are not language experts in general, can readily see. This coding scheme allows us to produce a clear, articulated description of what problems are observed in a translation and what the causes or roots of the problems are so that they can be addressed. Rather than reviewing each one as an isolated event, we can see what questions seem to be inadequate, and see where the issues lie, for each language.

For example, in Table 3, we describe how each coding category calls for a distinct type of solution.

Beyond individual issues or examples, there are also more global solutions to these translation issues. One possibility is to train translators and reviewers to identify different types of issues (Linguistic Rules, Cultural Norms, Social Practices, Production Errors) so that they can be addressed more swiftly with survey sponsors, providing translators and reviewers a framework within which to present their recommendations. With a broader goal in mind, findings can be shared with survey questionnaire designers and sponsors where appropriate, so that the original surveys can be constructed in a way that makes them more easily adaptable to other languages. Also, it would be useful to develop a bank of terms and concepts that are commonly used but difficult to translate and to include tested solutions so that future issues can be avoided.

Our next steps will be to use the coding scheme in additional survey and survey material translations and document the results of the translation review, to determine if the coding scheme is effective in producing final translations that are more functionally equivalent. We will also continue to

refine our coding categories based on feedback from survey sponsors with regard to the explanations that make the most sense logistically to implement. Also, we will implement a committee approach to coding (to find out where the areas of contention are, if any), to determine inter-rater reliability, and to adjust the coding scheme as necessary based on future cognitive interview data. The anticipated applications of this coding scheme are to better evaluate the quality of translated material, measure how quality improves over time, with training, etc., and better communicate the goal of functional equivalence outside of the translation team.

Acknowledements

We would like to acknowledge the support from the U.S. Census Bureau for this research. We are grateful to Barbara Lazirko (U.S. Census Bureau), Mandy Sha (RTI International) and Hyunjoo Park (RTI International) for their assistance in collecting the cognitive interviewing data for analysis. We would also like to thank the anonymous reviewers for reviewing an earlier version of this paper and for their constructive comments. However, any errors remain the sole responsibility of the authors. This report is released to inform the interested parties of research and to encourage discussion. The views expressed in the paper are those of the authors and not necessarily those of the U.S. Census Bureau.

References

Behling, O. & Law, K. S. (2000). *Translating questionnaires and other research instruments: problems and solutions*. London: Sage.

Carrasco, L. (2003). *The American Community Survey (ACS) en Español: using cognitive interviews to test the functional equivalency of questionnaire translations*. Statistical Research Division Research Report Series, Survey Methodology #2003-17. Washington, DC: US Census Bureau.

Conway, K., Patrick, D. L., Gauchon, T., & Acquadro, C. (2010). Enhancing cross-cultural appropriateness for newly developed patient-reported outcome (pro) instruments: the use of translatability assessment. *PRO Newsletter*, *44*, 9–12.

Dept, S., Ferrari, A., & Wäyrynen, L. (2010). Developments in translation verification procedures in three multilingual assessments: a plea for an integrated translation and adaptation monitoring tool. In J. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, ... T. W. Smith (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 157–173). New York: Wiley.

Dorer, B. (2011). *Advance translation in the 5th round of the European Social Survey (ESS)*. FORS Working Paper Series, 2011-4. FORS, Lausanne.

Fitzgerald, R., Widdop, S., Gray, M., & Collins, D. (2011). Identifying sources of error in cross-national questionnaires: application of an error source typology to cognitive interview data. *Journal of Official Statistics*, *27*, 569–599.

Forsyth, B., Kudela, M. S., Levin, K., Lawrence, D., & Willis, G. (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. & Caspar, R. (2007). A new methodology for the cognitive testing of translated materials: testing the source version as a basis for comparison. Proceedings of the Joint Statistical Meetings (Section on Survey Research Methods). Retrieved from https://www.amstat.org/sections/srms/proceedings/y2007/Files/JSM2007-000719.pdf

Gumperz, J. J. (1999). On interactional sociolinguistic method. In S. Sarangi & C. Roberts (Eds.), *Talk, work and institutional order: discourse in medical, mediation and management settings* (pp. 453–471). Berlin: de Gruyter.

Gumperz, J. J. (2001). Interactional sociolinguistics: a personal perspective. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 215–249). Malden, MA: Blackwell Publishing.

Harkness, J. (2003). Questionnaire translation. In J. A. Harkness, F. J. R. van de Vivjer, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 35–56). Hoboken, NJ: John Wiley & Sons.

Holmes, J. (1992). *An introduction to sociolinguistics*. London: Longman Group.

Levin, K., Willis, G. B., Forsyth, B. H., Norberg, A., Stapleton Kudela, M., Stark, D., & Thompson, F. E. (2009). Using cognitive interviews to evaluate the Spanish-language translation of a dietary questionnaire. *Survey Research Methods*, *3*, 3–25.

McKay, B., Breslow, M., Sangster, R., Gabbard, S., Reynolds, R., Nakamoto, J., & Tarnai, J. (1996). Translating survey questionnaires: lessons learned. *New Directions for Evaluation*, *70*, 93–104.

Pan, Y. (2008). Cross-cultural communication norms and survey interviews. In H. Sun & D. Kádár (Eds.), *It's the dragon's turn – Chinese institutional discourse(s)* (pp. 17–76). Bern: Peter Lang.

Pan, Y. (2011). Methodological issues in east asian politeness research. In D. Kádár & S. Mills (Eds.), *Politeness in East Asia: theory and practice* (pp. 71–97). Cambridge: Cambridge University Press.

Pan, Y. & de la Puente, M. (2005). *Census bureau guideline for the translation of data collection instruments and supporting materials: documentation on how the guideline was developed*. Statistical Research Division

Research Report Series, Survey Methodology #2005-06). Washington, DC: US Census Bureau.

Pan, Y., Landreth, A., Hinsdale, M., Park, H., & Schoua-Glusberg, A. (2007). Methodology for cognitive testing of translations in multiple languages. Proceedings of the Joint Statistical Meetings (Section on Survey Research Methods). Retrieved from https://www.amstat.org/sections/srms/proceedings/y2007/Files/JSM2007-000246.pdf

Pan, Y., Landreth, A., Hinsdale, M., Park, H., & Schoua-Glusberg, A. (2010). Cognitive interviewing in non-english languages: a cross-cultural perspective. In J. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, ... T. W. Smith (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 91–113). Hoboken: Wiley.

Pan, Y., Sha, M., Schoua-Glusberg, A., & Park, H. (2009). *2010 Census language program: pretesting of Census 2010 questionnaire in five languages.* Statistical Research Division's Research Report Series, Survey Methodology #2009-01. Washington, DC: US Census Bureau.

Potaka, L. & Cochrane, S. (2002). *Developing bilingual questionnaires: experiences from new zealand in the development of the 2001 maori language survey*. Paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing Methods (QDET), Charleston, SC.

Schiffrin, D. (1988). Sociolinguistic approaches to discourse: topic and reference in narrative. In K. Ferrera, B. Brown, K. Walters, & J. Baugh (Eds.), *Linguistic change and contact* (pp. 1–17). Austin: Department of Linguistics, University of Texas.

Schoua-Glusberg, A. (1992). *Report on the translation of the questionnaire for the National Treatment Improvement Evaluation Study*. Chicago: National Opinion Research Center.

Schoua-Glusberg, A. (2006). *Eliciting education level in Spanish interviews*. Paper presented at the American Association of Public Opinion Research Conference, Montreal, Canada.

Tannen, D. (2005). Interactional sociolinguistics as a resource for intercultural pragmatics. *Intercultural Pragmatics*, *2*, 205–208.

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

Willis, G., Lawrence, D., Hartman, A., Kudela, M., Levin, K., & Forsyth, B. (2008). Translation of a tobacco survey into Spanish and Asian languages: the tobacco use supplement to the Current Population Survey. *Nicotine & Tobacco Research*, *10*, 1075–1084.

Willis, G. & Zahnd, E. (2007). Questionnaire design from a cross-cultural perspective: an empirical investigation of Koreans and non-Koreans. *Journal of Health Care for the Poor and Underserved*, *18*, 197–217.