Translation of country-specific programs and survey error: Measuring the education level of immigrants

Patricia L. Goerman
U.S. Census Bureau
Patricia.l.goerman@census.gov

Leticia Fernández
U.S. Census Bureau
leticia.esther.fernandez@census.gov

Rosanna Quiroz
RTI International
quiroz@rti.org

Abstract: The difficulty of translating country-specific programs for use in surveys has been well documented. Questions about educational attainment offer a good illustration of this difficulty, particularly amongst Spanish-speaking immigrants in the United States, who come from a variety of countries where education systems are different in both name and structure. This article presents results from cognitive testing of Spanish education-level questions in the U.S. Census Bureau's American Community Survey. We conducted two iterative rounds of testing with 46 Spanish-speaking respondents from 11 different countries. Respondents had differing interpretations of the Census Bureau's education-level categories because they differed, either in meaning or by the terms used, from the categories in their countries of origin. For example, Mexican-origin respondents interpreted '*escuela secundaria,*' or 'high school,' to correspond to nine years of schooling, while in the U.S. completing high school corresponds to 12 years of schooling. This type of misinterpretation could result in upward biases in reports of educational levels. We discuss various approaches tested to deal with this type of response error.

Keywords: Education system differences, measuring immigrant respondents' education levels, multi-lingual surveys, translation of country specific programs.

1. Introduction

Various studies have shown the difficulty of translating concepts related to country-specific programs for use in surveys (Carrasco, 2003; Fernández et al., 2009; Goerman, Caspar, Sha, McAvinchey, & Quiroz, 2007; Schoua-Glusberg, Carter, & Martinez-Picazo, 2008). Examples of concepts that are difficult to translate range from respondent participation in the U.S. foster care program to specific types of health insurance coverage. Educational level has posed translation challenges to many U.S. agencies, particularly in the case of groups such as Spanish-speaking immigrants. These respondents come from a variety of different countries, many of which have educational systems that are different from the U.S. system and from each other. This has been an issue in European

countries as well and recent studies have focused on developing methods to best translate and adapt educational credentials of immigrants being surveyed in Europe (Schneider, 2007, 2009; Schneider, Joye, & Wolf, 2016).

In this article we present results of two iterative rounds of cognitive testing on a series of American Community Survey (ACS) educational-level questions with 46 Spanish-speaking respondents from 11 countries. The educational-level categories were initially translated without any attempt at adaptation or differences across languages to account for different educational systems. The first round of cognitive testing showed that Spanish speakers interpreted several of the categories differently from what was intended. Some terms were unfamiliar to respondents. Additionally, respondents from different countries used some of the same terms to refer to different levels of education, using the educational system in their countries of origin as a reference. These mismatches in terminology often resulted in upward biases in reports of educational levels.

This article provides an overview of previous research on the translation of country-specific programs. We then examine the Spanish translation of the ACS education questions. Based on cognitive interview results, we developed and tested strategies for improving the comparability of the educational-level terms across languages. Our results were mixed; but none of the approaches eliminated the translation problems completely. We conclude with a discussion of unresolved issues and recommendations for further research.

2. Review of the literature on survey translation

Comparative research across disciplines strives to evaluate key indicators of socioeconomic status across different populations. When measuring across different cultural and linguistic groups, translation quality and methods are of key importance. Much of the translation literature in the field of survey methodology focuses on recommended steps in the development of a survey instrument and its translation(s) (Dean, Caspar, McAvinchey, Reed, & Quiroz, 2007; Harkness, 2003; Harkness, Villar, & Edwards, 2010; Pan, 2009; Potaka & Cochrane, 2004; Schoua-Glusberg & Villar, 2014). Many researchers emphasize the importance of instrument adaptation, which includes tailoring instruments to fit culture-specific concepts in target populations, such as use of preferred terms, use of unambiguous concepts, and use of culturally relevant examples (Dean et al., 2007; Harkness et al., 2010; Hunt & Bopal, 2004; Mohler, Dorer, de Jong, & Hu, 2016). There may be terms or concepts that do not exist in respondents' countries of origin and researchers should address this possibility during instrument development (Behr & Shishido, 2016; Goerman, 2010; Schoua-Glusberg et al., 2008; Schoua-Glusberg & Villar, 2014). Additional challenges arise when the adaptation of instruments must include several subpopulations that speak different dialects of a language (Cortés-Martinicorena, Rosel-Gallardo, Artázcoz-Osés, Bravo, & Tsakos, 2010; Domínguez et al., 2006).

Cognitive testing is a pretesting method involving face-to-face interviewing, in which respondents are probed about their responses, interpretations, and mental processes associated with survey questions (Willis, 2005). Careful analysis can identify questioning strategies that will yield more accurate answers, thus reducing measurement error (Biemer, Groves, Lyberg, Mathiowetz, & Sudman, 1991). The method can also be used to test whether translated instruments measure the intended concepts (Behr & Shishido, 2016; Schoua-Glusberg & Villar, 2014).

In measuring educational attainment across different populations in the U.S., survey instrument translation must include not only translation to non-English languages but also the recognition that there are large differences in the way educational systems are structured in immigrants' countries of origin. Schoua-Glusberg et al.'s (2008) review of educational systems in Latin America found large variations in the number of educational categories used across countries to measure levels of education, from six categories in Costa Rica to 14 in Chile and Puerto Rico. Finding a common classification system is increasingly challenging when the U.S. education system is added to the comparisons. Adaptation of educational level categories to work across educational systems would be ideal. However, due to a lack of resources and mechanisms to examine the validity of major adaptations, this is often not possible.

2.1 Background on the ACS Cognitive Testing Project

The ACS is a monthly U.S. Census Bureau survey that samples about 3.5 million addresses each year and contains questions about households, including the number of occupants and type of dwelling, and questions about each person living in the household, such as age, sex, race, education, and employment.

There are two slightly different versions of the Spanish-language questionnaire, one for use in the continental U.S. (the 'stateside' version) and one for use in Puerto Rico. This makes it possible to tailor the Puerto Rican version to contain terms that are more familiar to Puerto Rican Spanish speakers. The ACS is a multi-mode survey: respondents first receive an invitation to participate via the internet and then receive a follow-up paper questionnaire. Until 2017 non-response to these modes was handled first through a telephone interviewing CATI operation and then by a face-to-face CAPI operation. Due to cost and respondent burden considerations, the CATI operation was ended in 2017. A Spanish internet option was implemented in 2013 but response in Spanish is still overwhelmingly concentrated in the interviewer administered, CAPI mode.

During several years before and after 2010, the Census Bureau underwent a process of cognitively testing segments of the ACS Spanish CATI and CAPI instruments to verify that the English and Spanish versions were collecting parallel data.

3. Methodology and respondent characteristics

Researchers conducted cognitive testing of the Spanish ACS instrument through semi-structured cognitive interviews, which focused on interpretation and comprehension of key words and phrases. In testing the education-level questions, we first asked respondents the full series of questions without interruption. Interviewers then probed about the meaning of particular terms and questions. We conducted two iterative rounds of testing (see Fernández et al., 2009 for more details).

We interviewed twenty-three Spanish-speaking respondents and five English speakers in each of two rounds of testing. Testing of the questions with a small number of English-speakers was included to help assess whether the issues discovered in Spanish were due to translation problems or conceptual difficulties with the original English wording (see Goerman & Caspar, 2010; Schoua-Glusberg & Villar, 2014). A better method would have been parallel testing and revisions to the English and Spanish versions, including testing with non-native English speakers educated outside of the U.S. However, the current project focused only on testing and making revisions to the Spanish instrument. Because

of time and resource constraints, the study sponsor was only seeking ways to verify the quality of the translation and was not interested in introducing major adaptations or changes that would require further verification and/or changes to the English version.

We recruited respondents at two sites (Puerto Rico and North Carolina). We used a non-random sample and recruited respondents to represent the diversity of the U.S. Hispanic immigrant population. Respondents were people with limited or no English proficiency from a variety of Spanish-speaking countries who also represented diverse levels of educational attainment. About a quarter of the respondents (24 percent) had completed less than high school; 28 percent had a high-school level diploma; and 48 percent had at least some college or technical courses.

Respondents were from Argentina, Colombia, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Peru, Puerto Rico, and Venezuela. About 44 percent of the respondents were ages 30 and younger; slightly over half were female; and most of the stateside respondents had lived in the U.S. for fewer than 10 years (77 percent). Consistent with the socioeconomic characteristics of U.S. immigrants from Latin America (Grieco et al., 2012), a higher proportion of respondents with low levels of education came from Mexico and Central America than from other regions. The Puerto Rican respondents were all born and residing in Puerto Rico.

4. Results

The ACS education level series begins with a question about recent school enrolment. The second question asks for details about the type of school (public, private or home school) and the level of education of those who were enrolled in school at any time in the last three months. Next, all respondents are asked, 'What is the highest degree or level of school you have COMPLETED?' For respondents who answer that their completed education is between grade 11 and 'some college no degree' (that is, grade 11, grade 12, high school, a GED[1] or some college), there are several follow-up questions. Finally, respondents who answer that they have completed a four-year college degree or higher are asked about their main fields of study.

In this article we focus on cognitive testing findings among Spanish speakers. Detailed results for English speakers, who generally had no difficulty with these questions, are reported elsewhere (Fernández et al., 2009). The first round of Spanish-language cognitive interviews revealed three types of response burden or error. First, respondents had difficulty understanding some questions due to their length and complexity. Secondly, some of the terms were confusing to Spanish speakers because they referred to concepts that do not exist in Latin American countries, such as 'home schooling' or the 'GED' (see footnote 2). Finally, Spanish-speakers often found it difficult to select a response option from among the education levels listed.

Overall, in the first round of testing we observed that complex and ambiguous Spanish question wording was not only imposing unnecessary cognitive burden on respondents but also potentially impacting data quality, as we observed some respondents choosing responses that did not correspond to the number of years of schooling they had actually completed. Below we discuss this

---

[1] GED refers to General Educational Development tests. GED tests are a group of exams about different topic areas. They were designed by the American Council on Education to allow people to

response burden and error in greater detail.

4.1 Problem Type 1: Overly complex question structure
One type of problem we identified was that some of the questions were overly long and complex. This issue was magnified in Spanish, since the translation was even longer than in English. The first question in the education series is a good example of this. The original question wording in both languages was as follows:

> *Las siguientes preguntas son sobre instrucción y educación. En cualquier momento <u>durante los últimos 3 meses</u>, ¿asistió usted a una escuela o universidad? Incluya sólo guardería infantil o preescolar, kindergarten, escuela elemental, enseñanza en el hogar y escuela que conduce a un diploma de escuela secundaria o un título universitario.*

> [The next questions are about schooling and education. At any time <u>in the last 3 months</u>, have you attended school or college? Include only nursery or preschool, kindergarten, elementary school, home school, and schooling that leads to a high school diploma or a college degree.

After hearing the question, over half of the Spanish-speakers asked for clarification before offering a 'yes' or 'no' answer, and two answered 'yes' incorrectly. The respondents' questions suggested that they had difficulty keeping track of the multiple conditions in the question. Respondents asked things like: "…in the last three months?", "for me or a child?", and "to study or to visit?"

Compounding the issue of question length, there were also terms in the question that proved confusing. For example, the term '*asistió*' (attended) can mean both 'attended' as in attending school (enrolled) and 'went to,' as in physically going to the school building. For some respondents, the term took on this second meaning, particularly when it was combined in the same sentence with the examples 'nursery or preschool, kindergarten.' Some respondents thought that the question was asking about 'going to' any of these places for any purpose, including dropping off a child.

Other respondents heard the list of examples and thought it was meant to be exhaustive because of its length. This was evident in debriefing when some respondents pointed out that there were options "missing", such as computer courses and English classes.

One obvious recommendation was to shorten the question or divide it into multiple questions. However, due to the fact that the English-language instrument was not open for edits and the survey sponsor wanted to avoid adaptations, we were only able to test minor terminology modifications in the second round, such as changing the term '*asistió*' (attended) to '*estudió*' (studied).

All minor terminology changes tested in the second round were successful in improving question administration and fewer respondents asked for clarification when they heard the revised question. However, we continued to see issues related to the length of the question. In the second round we found that respondents were having difficulty remembering the reference period. In fact, seven out of 18 Spanish-speaking respondents (39 percent) did not notice or remember the phrase 'in the last 3 months.' Moreover, two of the four respondents who answered 'yes' to this question had actually not been studying in the last three months.

Ultimately, we found that the issue with the question being overly long and complex could not be resolved in the translated version alone. This issue would require joint research and testing of the original English and translated versions.

4.2 Problem Type 2: Conceptual mismatch across cultures

Another issue identified with the education questions was that respondents were not familiar with some types of U.S. education, such as 'home schooling,' and the 'GED or alternative credential.' These programs do not exist in many Latin American countries. Therefore, respondents interpreted the terms within their own cultural contexts.

Example 4.2.1: Home Schooling, 'Enseñanza en el hogar'

Interviewers probed all respondents in the first round of cognitive testing about their interpretation of the question below, even if they reported not attending school in the last three months:

> *¿Fue esa una escuela o universidad pública, una escuela o universidad privada o era enseñanza en el hogar?*
>
> [Was that a public school or college, a private school or college, or home school?]

Although Puerto Rican respondents did not immediately recognize the phrase '*enseñanza en el hogar*,' we found that many of them were familiar with the English phrase 'home school,' perhaps because of their familiarity with the mainland U.S. educational system.[2] This suggests that providing a literal translation, even in the case of a concept that people are familiar with, can be problematic. The educational system in Puerto Rico is structured similarly to the U.S. mainland system, with three levels: primary or elementary school (grades 1-6); 'intermediate' or junior high school (grades 7-9); high school (grades 10-12); and college or higher education. Furthermore, home schooling is an educational option in Puerto Rico.

Other Spanish speakers, however, were not familiar with the concept of 'home school' and they interpreted the common terms in the phrase '*enseñanza en el hogar*' as less formal concepts such as online classes, bible school, and lessons taught by parents related to culture, table manners, and moral principles. To address this issue, we proposed different modifications to the two Spanish versions. In the Puerto Rican version, we recommended the addition of the English term (home school) at the end of the question. We initially recommended an adaptation to resolve the stateside issue: dropping the home school concept from Spanish or the use of an altogether different translation in the Spanish instrument. Project guidelines did not allow for this type of deviation, however, so our next recommendation was to include a short definition of home school. We based the definition on the wording used by the National Household Education Surveys Program (NHES, 2007). In English, the definition reads: "Home schooling is when parents decide to educate their children at home instead of sending them to school."[3] The English 'home school' was added in parentheses to the Puerto Rican version of the question and both the Puerto Rican and stateside versions were tested including the new definition. However, both versions maintained the original home school translation: '*enseñanza en el hogar*.'

Round 2 testing revealed that many Puerto Rican respondents (six of eight)

---

[2] Puerto Rico is an unincorporated territory of the United States. As such, Puerto Rican people are U.S. citizens by birth and many U.S. programs, English language terms and concepts are familiar to people living in Puerto Rico.

[3] Due to the fact that the English language instrument was not open for edits, thereby restricting major restructuring to the Spanish instrument, the definition was under consideration only as possible optional help text to be added to the interviewer instructions in the instrument.

recognized the term 'home school' in English. Among the other Spanish speakers, only one person (a Guatemalan) understood the unchanged term '*enseñanza en el hogar*' as intended. The other respondents, with diverse educational levels and national origins, heard both the Spanish term *and* the definition in Spanish, and still interpreted '*enseñanza en el hogar*' to mean learning on one's own such as from videos, TV or internet or by distance learning or private tutoring.

The fact that home schooling does not exist in many Latin American countries makes it challenging to explain the concept in a simple sentence. Of course, we have no way of knowing whether or what proportion of Spanish speakers would answer in the negative when they hear a question involving an unfamiliar term. In this case, however, the terms used in the translation sounded familiar to respondents and they had a particular meaning in their cultural context. The risk in maintaining this wording is that respondents may confuse '*enseñanza en el hogar*' (home school) with online learning or private tutoring and this would cause response error in the form of inaccurate 'yes' responses, with effects on data quality. In the absence of the requirement to maintain consistency with the English version, our recommendation would have been to drop and/or replace the term home schooling on the stateside Spanish version of the question. Further research into whether an adaptation could collect parallel data across languages is warranted.

4.3 Problem Type 3: Ambiguous terms used in translation

The first question in the education-level series is the same for Puerto Rican and other Spanish speakers. The question reads:

> *¿Cuál es el título o nivel escolar más alto que usted ha COMPLETADO?*

> [What is the highest degree or level of school you have COMPLETED?]

Since this research was done during the time when the ACS had both CATI and CAPI operations, we tailored the testing to both modes. Interviewers using the CAPI instrument were to show respondents a show card containing a long list of response options while CATI interviewers read from the same list of options with the instruction to stop when the respondent interrupted with an answer. In cognitive testing, roughly half of respondents were interviewed simulating the CAPI mode and saw the options listed on a show card. The other half were interviewed simulating the CATI mode and they heard the interviewer read the list of options. A number of terms were found to be ambiguous and in all of these cases some respondents over-reported their educational levels, an issue with strong implications for data quality.

Example 4.3.1: Regular high school diploma ('Diploma de escuela secundaria')

In the first round of testing both Puerto Rican respondents and those from Nicaragua, Peru and Venezuela understood the response option '*diploma de escuela secundaria*' to mean high school, or 12 years of schooling, as intended. This was because those countries all have educational systems in which education after grade school is blocked as a 5-year grouping, similar to the structure in the U.S. education system, in which students complete the secondary level at 17 or 18 years of age. However, despite the fact that they understood the original wording, Puerto Rican respondents expressed a preference for the phrase '*diploma de escuela superior.*'

Mexicans and Colombians, on the other hand, understood '*diploma de escuela secundaria*' as a diploma earned after nine years of schooling, not twelve.

In their countries, '*preparatoria*' or '*bachillerato*' are the terms used for 12 years of schooling.

Based on these observations, we tested two different Spanish versions in the second round of interviews. The original wording '*diploma de escuela secundaria*' for 'regular high school diploma,' was modified for Puerto Rico to read '*diploma de escuela superior* (high school).' We included the English term ('high school') in parenthesis to take advantage of the familiarity of Puerto Ricans with the U.S. mainland educational system. In the second round of testing, we confirmed that Puerto Ricans had no difficulty recognizing the phrase to mean the diploma obtained after 12 years of schooling.

Choosing revised terms for the stateside Spanish instrument was more difficult since the target population is composed of Spanish speakers from many countries. Our recommendation was to change the translation for 'Regular high school diploma' to '*Grado 12, CON DIPLOMA* (high school)' [Grade 12, WITH DIPLOMA (high school)]. This was deemed too great a deviation from the English to be made based on cognitive testing alone.

Instead, we tested the strategy of grouping terms that would be meaningful to different national origin respondents together in the same category. We tested the phrase '*diploma de escuela secundaria o preparatoria* (high school).' The hope was that respondents might recognize at least one of these terms as equivalent to high school and report their education level accurately. In addition, we added the English-language term in parentheses; since respondents are living in the U.S., they may be familiar with the English terms.

Changes to the stateside Spanish version were not as successful as the Puerto Rican changes. Mexican respondents were confused by the use of the terms '*escuela secundaria o preparatoria* (high school)' grouped together because both Spanish terms are used in their educational system and they mean different levels of education. In fact, because they now interpreted the response option to be asking if they had either 9 *or* 12 years of schooling, respondents from Mexico who had less than a high school education still mistakenly chose this option instead of choosing '*Grado 9*' (Grade 9). Respondents from Mexico specifically recommended not grouping '*secundaria*' and '*preparatoria*' together since to them they are different levels of schooling.

The term 'high school' in parentheses was of limited help since some respondents did not know how many years it takes in the U.S. to complete high school. Part of the explanation for this may be that our recruitment criteria for this project included that participants speak very little English (i.e., recent immigrants or those relatively isolated from mainstream U.S. interactions). Based on our findings, we recommended that the tested changes not be made to the stateside Spanish version. There was a need for further research into whether an adaptation that would enable the collection of parallel data across languages could be identified and validated.

Example 4.3.2: Bachelor's degree (*Título de bachiller universitario*)
Another translation that was problematic for many non-Puerto Rican Spanish-speakers was 'bachelor's degree,' a term meant to capture a four-year college education. In the first round of testing we confirmed that Puerto Rican respondents understood '*Título de bachillerato universitario*' as a 4-year college degree as intended. However, this was not the case for respondents from many Latin American countries, particularly Mexico.

For example, a Mexican respondent with a high school education incorrectly selected '*Título de bachiller universitario*' (bachelor's degree) from the list of options because he thought it meant '*bachillerato*' (high school in Mexico). Two

other respondents with a bachelor's degree selected '*Título asociado*' (Associate's degree) and '*Título profesional más allá de un título de bachiller*' (professional degree beyond a bachelor's degree) because they also understood the term '*bachiller*' to refer to high school in their countries (Mexico and Colombia). Another college-educated Mexican respondent said that she was finally able to select the right answer after reading the list of options *three times* when she noticed 'BA' and 'BS' as examples under the bachelor's degree category.[4] This respondent had some familiarity with the U.S. educational system.

Moreover, college-educated Spanish speakers from Nicaragua and Mexico thought that '*bachiller universitario*,' referred to a 'technician with a specialty,' and that the degree was 'equivalent to high school.' The terms seemed contradictory to them since it included '*bachiller*' (high school) and '*universitario*' (university level) together in the same category. In all, using the term '*bachiller*' for bachelor's degree had the effect of pushing some of our respondents to overstate their educational level. Those with a high school degree tended to choose this option and those with a bachelor's degree tended to select options corresponding to higher levels of education in their efforts to convey that they had studied beyond high school.

We tested a new phrase for 'bachelor's degree' in the second round of interviews. However, the term '*bachillerato*' was also kept in the category because of the survey sponsors' concerns that some respondents may look for a term that looks similar to the English 'bachelor's degree.' The new wording was: '*Título de licenciatura o bachillerato universitario (por ejemplo: BA, BS)*' (Bachelor's degree (for example: BA, BS)).

The main observation from the second testing round was that college-educated Spanish-speakers from some countries (Colombia, Nicaragua and Mexico) had difficulty understanding the new terms as intended. These respondents understood '*título de licenciatur*a' as intended but said that including the term '*bachillerato*' confused them because it refers to a lower level of schooling. Based on these findings the term '*bachillerato*' was ultimately dropped from the stateside Spanish version.

With only two rounds of cognitive interviewing, the limitation that the Spanish instrument needed to be similar to the English, and the fact that changes to the English version were not possible at the time of the research, our ability to make changes in wording was limited and several issues remained unresolved. However, we identified strategies that enabled us to address a number of issues successfully. These strategies, which we discuss in the next section, could be applied in other projects where there is more flexibility to make changes to a translation.

5. Discussion and recommendations

This article contributes to the literature on translating survey questions across cultural groups and languages with the ultimate goal of reducing measurement error. The measurement of educational level in countries with high rates of immigration, such as the U.S., can introduce response error in survey data for a number of reasons. In particular, the educational systems in respondents' home countries differ not only from the U.S. system but from each other as well. A further complication is that labels assigned in some countries may be the same but

---

[4] BA stands for Bachelor of Arts, and BS for Bachelor of Science. Both are often used as abbreviations to refer to four-year college degrees in the U.S.

refer to different levels of education or a different number of years of study (e.g., '*secundaria'* corresponds to a five-year block in Peru, a four-year block in Colombia, and a three-year block in Mexico).

As our findings show, in some cases there is no easy solution that will apply to respondents of all national origins. However, our research provides some translation strategies that can be used when developing questions for use with culturally diverse populations.

In this study, we observed many differences in the terms used by different cultural and national origin groups. The strategies we used to deal with this included tailoring the wording to just one group; grouping multiple terms or synonyms that would be meaningful to respondents with different backgrounds together; the inclusion of English terms in parentheses to assist respondents who are more acculturated; and the use of optional help text to explain confusing terms.

Tailoring translations to use terms specific to a country's educational system, such as in the case of Puerto Rico, was a successful adaptation strategy. The fact that we had a separate Spanish instrument for use in Puerto Rico made this possible, and may substantially reduce the potential for biased data.

A second strategy used was grouping multiple terms or synonyms together in one category. The strength of this strategy is that when terms from different countries are grouped together, respondents should be able to recognize at least one of them; however, in this case the fact that the same educational terms have different meanings in different countries made the strategy unsuccessful.

Thirdly, our respondents' lack of familiarity with the U.S. educational system and English language limited the usefulness of English terms in parentheses. However, monolingual Spanish speakers are not the only respondents who use Spanish survey instruments. It may be that this strategy would work well for Spanish speakers who know some English but who choose to complete a survey in Spanish.

A final strategy we used was the inclusion of a definition or optional help text when we could not find a way to better translate a concept or program (i.e., home schooling). In our study, including a short definition of home schooling did not improve respondent comprehension. Even if the definition had been helpful there would be no guarantee that it would be read regularly to respondents in the field. The ideal solution in this type of situation is to use an adaptation, or different, more culturally relevant, example. This is only possible when the term is used as one possible example as opposed to being part of an exhaustive list of categories that must be included to keep the question parallel across languages.

Our inability to include and validate adaptations at the time of this study prevented us from resolving several of the issues identified. One such issue was how to translate terms in the case where Spanish-speaking respondents from different countries use the *same* terms to refer to *different* concepts. In this type of case, it may be that a search for 'terms' is not the most productive avenue. Instead, obtaining information about the *number* of years of education completed *in addition to* the names of specific degrees could prove fruitful. There is previous research showing that number of years of schooling alone can be misclassified as a degree earned (Kominski & Siegel, 1993). Reporting of degrees earned provides more information about educational attainment, such as the ability to differentiate between technical and academic tracks (Kominski & Siegel, 1993; Schneider, 2007; Schneider et al., 2016).

Given our findings, however, asking a follow-up question about number of years of schooling of Spanish speakers who report that their level of school completed is '*diploma de escuela secundaria o preparatoria* (high school)'

(regular high school diploma) may help to separate those who have completed the equivalent of high school from those who have completed less than high school. Cognitive testing could help to determine whether this is the case.

This would represent a departure from the wording in the English version, i.e., an adaptation. However, when the goal is to collect parallel data from different cultural and linguistic groups there are some cases where adaptation may be the only viable solution. Validation testing could be employed to ensure that parallel data can be collected via different question wording. Ideally, during questionnaire development both the source and translation(s) should be open to revisions, joint testing, and cultural adaptations. The goal should be to develop instruments that can collect parallel data without having to be identical in structure and/or content.

Areas for future research

This study suggests several areas for future research. First, it is becoming more possible to tailor survey questions to national origin groups given CAPI/CATI, internet, and new mobile technology. In theory, an automated instrument could contain a screening question asking in what country a respondent completed his or her highest level of schooling. The respondent could then be sent down a survey path containing educational level questions specific to the system in that country. This solution is currently not an option in the ACS because the survey is also administered through paper questionnaires. However, this may be a future solution for automated surveys.

Further research is also needed to understand how translations and/or adaptations affect data quality. We can use this knowledge to construct new techniques to translate questions involving unfamiliar concepts among specific populations, such as recent immigrants and respondents with limited English proficiency.

This article has focused on several possible solutions to the problem of how to best translate survey questions about country specific programs. We have discussed a number of possible solutions to add to researchers' 'tool kits' of translation techniques. While we hope that some of these techniques would be useful with country-specific programs aside from education, we advise that any of these possible solutions be tested with respondents of various backgrounds to be sure that they will work in a given context.

Disclaimer: This article was written to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau or RTI International.

References

Behr, D., & Shishido, K. (2016). The translation of measurement instruments for cross-cultural surveys. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The SAGE handbook of survey methodology* (193-209). London: SAGE Publications.

Biemer, P. P., Groves, R. M., Lyberg, L.E., Mathiowetz, N.A., & Sudman, S. (Eds.). (1991). *Measurement errors in surveys*. New York, NY: John Wiley & Sons, Inc.

Carrasco, L. (2003). *The American Community Survey (ACS) en Español: Using Cognitive Interviews to Test the Functional Equivalency of Questionnaire Translations*. U.S. Census Bureau Statistical Research Division Study Series (Survey Methodology #2003-17). Retrieved July 3, 2018 from https://www.census.gov/srd/papers/pdf/ssm2003-17.pdf

Cortés-Martinicorena, F. J., Rosel-Gallardo, E., Artázcoz-Osés, J., Bravo, M., & Tsakos, G. (2010). Adaptation and validation for Spain of the Child-Oral Impact on Daily Performance (C-OIDP) for use with adolescents. *Med Oral Patol Oral Cir Bucal*, *15*(1), 106-111.

Dean, E., Caspar, R., McAvinchey, G., Reed, L., & Quiroz, R. (2007). Developing a low-cost technique for parallel cross-cultural instrument development: The Question Appraisal System (QAS-04). *International Journal of Social Research Methodology, 10* (3), 227-241.

Domínguez, R., Vila, J.F., Augustovski, F., Irazola, V., Castillo, P.R., Rotta Escalante, R., … Meschia, J.F. (2006). Spanish cross-cultural adaptation and validation of the National Institutes of Health Stroke Scale. *Mayo Clinic Proceedings, 81*(4), 476-480.

Fernández, L., Clifton, M., Quiroz, R., McAvinchey, G., Reed, L. & Rodriguez, S. (2009). *American Community Survey Spanish CAPI/CATI cognitive testing, phase III – rounds 1 & 2 - final report*. Statistical Research Division internal report. Washington, DC: U.S. Census Bureau.

Goerman, P. L. (2010). Number and nationality of Spanish-speaking immigrant respondents for U.S. cognitive testing studies. *2010 JSM Proceedings,* 5924-5937.

Goerman, P. L., & Caspar, R. A. 2010. A preferred approach for the cognitive testing of translated materials: Testing the source version as a basis for comparison. *International Journal of Social Research Methodology, 13*(4), 303-316.

Goerman, P. L., Caspar, R. A., Sha, M., McAvinchey, G., & Quiroz, R. (2007). *Census bilingual questionnaire research final round 2 report*. Statistical Research Division Working Paper Series. Washington, DC: U.S. Census Bureau.

Grieco, E. M., Acosta, Y. D., de la Cruz, G.P., Gambino, C., Gryn, T., Larsen, L. J., Trevelyan, E. N., & Walters, N. P. (2012). The foreign-born population in the United States: 2010. *American Community Survey Report #ACS-19*. Washington, DC: U.S. Census Bureau.

Harkness, J. A. (2003). Questionnaire translation. In J. A. Harkness, F. J.R. Van de Vijer, & P. Mohler (Eds.), *Cross cultural survey methods* (pp. 35-36). Hoboken, NJ: John Wiley & Sons, Inc.

Harkness, J. A., Villar, A., & Edwards, B. (2010). Translation, adaptation and design. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. Mohler, … T. W. Smith (Eds.), *Survey methods in multinational, multiregional and multicultural contexts* (pp. 117-140). Hoboken, NJ, John Wiley & Sons, Inc.

Hunt, S. M., & Bhopal, R. (2004). Self-report in clinical and epidemiological studies with non-English speakers: The challenge of language and culture. *J Epidemiol Community Health, 58*, 618-622.

Kominski, R., & Siegel, P.M. (1993). Measuring education in the Current Population Survey: A revision of educational attainment questions. *Monthly Labor Review,* September, 34-38.

Mohler, P., Dorer, B., de Jong, J., & Hu, M. (2016). *Adaptation: Guidelines for best practice in cross-cultural surveys* (4th ed.). Ann Arbor, MI: Survey Research Center, Institute for Social Research. University of Michigan.

NCES. (2007). Parent and family involvement in education questionnaire module, 2007. National Household Education Surveys Program. Washington, DC: U.S. Department of Education.

Pan, Y. (2009). A process for reviewing translations of data collection instruments and related materials. *Statistical Research Division Working Paper Series #2009-09*. Washington, DC: U.S. Census Bureau.

Potaka, L., & Cochrane, S. (2004). Developing bilingual questionnaires: Experiences from New Zealand in the development of the 2001 Maori language survey. *Journal of Official Statistics*, *20*(2), 289-300.

Schneider, S. L. (2007). *Measuring educational attainment in cross-national surveys: The case of the European Social Survey*. Paper presented at the 8th Conference of the European Sociological Association (ESA) in Glasgow, 3-6 September 2007. Retrieved from http://www.equalsoc.org/uploaded_files/publications/schneider 2007 -1.pdf

Schneider, S. L. (2009). *Confusing credentials: the cross-nationally comparable measurement of educational attainment: DPhil thesis*. (Doctoral dissertation, University of Oxford, UK. Retrieved from http://ora.ouls.ox.ac.uk/ objects/uuid%3A 15c39d54-f896-425b-aaa8-93ba5bf03529

Schneider, S. L., Joye, D., & Wolf, C. (2016). When translation is not enough: Background variables in comparative surveys. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 288-307). London: SAGE Publications.

Schoua-Glusberg, A, Carter, W., & Martinez-Picazo, E. (2008). *Measuring education among Latin American immigrants in the U.S.: A qualitative examination of question formulation and error*. Paper presented at the International Conference on Survey Methods in Multinational, Multiregional and Multicultural Contexts (3MC). Berlin, Germany.

Schoua-Glusberg, A., & Villar, A. (2014). Assessing translated questions via cognitive testing. In K. Miller, S. Wilson, V. Chepp, & J. L. Padilla (Eds.), *Cognitive interviewing methodology* (pp. 51-67). Hoboken, NY: John Wiley & Sons.

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