

# **An Overview of A Proposed Census Bureau Guideline for the Translation of Data Collection Instruments and Supporting Materials**<sup>1</sup>

**Manuel de la Puente and Yuling Pan, U.S. Census Bureau**

**Deborah Rose, National Center for Health Statistics**

Manuel de la Puente, U.S. Census Bureau, Washington, D.C. 20233 (email: [Manuel.de.la.Puente@census.gov](mailto:Manuel.de.la.Puente@census.gov))

*The Census Bureau developed guidelines for the translation of data collection instruments and supporting materials in order to ensure that such documents translated from a source language into a target language are reliable, complete, accurate, and culturally appropriate. In addition to meeting these criteria, guidelines were developed to ensure that translated Census Bureau data collection instruments also have semantic, conceptual, and normative equivalence. The guideline recommends that the translation of data collection instruments from a source language into a target language be conducted using a translation team. The guideline relies on the cross-cultural and survey methodology research literature and specifies and describes five steps that comprise the translation process. These steps are: Prepare, Translate, Pretest, Revise, and Document.*

## **Background**

The increasing need for non-English language data collection instruments and other survey materials is clear given recent figures. The Census 2000 Supplementary Survey<sup>2</sup> revealed that, nation-wide, there are approximately 45 million people aged 5 years and older who speak a language other than English at home. This represents about 18 percent of persons in this age group. Of these 45 million people, over 10.5 million either speak English "not well" or "not at all."<sup>3</sup>

It is important to note that the national figures tell only part of the story. For example, according to data from the Census 2000 Supplementary Survey there are counties where the percent of people aged 5 and over who speak a language other than English at home is far greater than the national figure of 18 percent. For instance, in San Francisco County, California the comparable figure is almost 46 percent, in the Bronx Borough in New York it's slightly over 54 percent, and in Starr and Zapata Counties in Texas the comparable percentage is 93 percent.<sup>4</sup>

While Spanish speakers account for about 60 percent of the 45 million individuals who reported speaking a language other than English at home,<sup>5</sup> the necessity for data collection instruments in languages others than English and Spanish, especially for the decennial census, cannot be ignored. Other than size, factors such as differential undercoverage and geographic concentration of non-English speaking population should be considered in our dialog regarding the need for non-English language data collection instruments. It is also essential, where appropriate, to develop and pretest translations of data collection instruments and other survey materials into languages other than Spanish.

---

<sup>1</sup> This paper reports the results of research undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. It is preliminary in nature and in the early states of development. As such, it is subject to revision. Our intent in making this working document available at this time is to inform ongoing discussions related to language research at the Census Bureau.

<sup>2</sup> The Census 2000 Supplementary Survey universe is limited to the household population and does not include the population living in institutions, college dormitories, and other group quarters.

<sup>3</sup> Li R.M., McCardle, P., Clark, R.L., Kinsella K. and Berch, D. eds. *Diverse Voices- Inclusion of Language-Minority Populations in National Studies: Challenges and Opportunities*. National Institute on Aging and National Institute of Child Health and Human Development. Bethesda, MD: 2001.

<sup>4</sup> These data were obtained from standard tables: *QT-02. Profile of Selected Social Characteristics: 2000* generated by the American FactFinder, September 18, 2001.

<sup>5</sup> Li R.M. et al. 2001

Obtaining high quality data from households where English is not the home language and where some or all of its members have a limited knowledge of English requires more than just having a correctly translated data collection instrument. Recent research on the Spanish language American Community Survey computer assisted personal interview instrument indicated that some correctly translated questions still pose conceptual problems and other difficulties for Spanish speaking respondents.<sup>6</sup> This finding supports the need for conducting cognitive testing on all household data collection instruments, not just English language instruments as is currently the practice at the Census Bureau.

## Research Conducted to Support Guidelines

An important aspect of the research undertaken by the Census Bureau to develop the translation guideline was a search of documents pertaining to the translation of data collection instruments available on the World Wide Web. This search revealed that most of the information available is limited. Available guidance is generally included or incorporated in documents developed for other purposes such as delivery of social or health services and the translation of text in general (surveys are not usually mentioned explicitly). Other advice on the conduct of translations come from for-profit and non-profit organizations and professional translators engaged in activities such as the translation of legislation.

The results of this search also indicate that a handful of key international statistical agencies such as Statistics Canada, Statistics New Zealand, the Australian Bureau of Statistics, and the Center for Survey Research and Methodology (ZUMA) in Mannheim, Germany provide some guidance for translating questionnaires.

The cross cultural survey literature describes a number of approaches to develop questionnaires in multiple languages. The choice of approach used depends on whether there is an existing data collection instrument, or whether the questionnaire needs to be developed from scratch. If there is an existing questionnaire then there are two major approaches that can be followed. The first approach is to *adopt* the instrument to the target language and the second is to *adapt* the instrument to the target language.

*Adoption* calls for the direct translation of the data collection instrument from the source language to the target language without regard to linguistic and cultural subtleties that impact the intended meaning of the question.<sup>7</sup> The second approach, *adaptation*, uses the existing questionnaire as the basis, but adaptation allows for components of the survey question to be modified or altered (independent of changes made as a result of the translation) in order to make the survey question suitable for fielding in the target language. Adaptation acknowledges and accounts for semantic, conceptual and other differences that exist across languages.

An important task in the development of data collection instruments in multiple languages is the actual conduct of the translation, or, taking words and their meaning in the source language and identifying words in a target language that convey the same or similar meaning. This task can be performed using a variety of approaches or techniques. Translation techniques discussed in the context of questionnaire translation talked about in the research literature include *simple direct translation*, *modified direct translation*, *back-translation*, and the *committee approach*.<sup>8</sup>

A *simple direct translation* is conducted by a single bilingual individual who translates the questionnaire from the source language into the target language. *Back-translation* is another technique with wide appeal because it's relatively low cost and quick. Back translation is an iterative process that entails three basic steps (1) Translation of the source language instrument

---

<sup>6</sup> Carrasco, Lorena, *The American Community Survey (ACS) en Espanol: Using Cognitive Interviews to Test the Functional Equivalency of Questionnaire Translations*. Statistical Research Division Study Series Report (Survey Methodology #2003-17). Issued August 5, 2003. U.S. Census Bureau, Washington, D.C.

<sup>7</sup> For example see Harkness, J., Van de Vijver, F.J.R., and Mohler, P. (Eds) *Cross-Cultural Survey Methods*. New Jersey: John Wiley and Sons, Inc. 2003 for a description and discussion of the use of this model in the translation of surveys.

<sup>8</sup> For example, see Harkness, et al. *Cross-Cultural Survey Methods*. New Jersey: John Wiley and Sons, Inc. 2003 and O. Behling and K.S. Law *Translating Questionnaires and Other Research Instruments: Problems and Solutions*. London: Sage Publications, Inc. 2000.

by a bilingual individual. (2) Translation of the target language instrument back to the source language instrument by a second bilingual individual. (3) Comparison of the original source instrument with the back translated source language instrument. If substantial differences arise when the source instrument is compared to the back translated source instrument in Step #3, then a new target language instrument is produced that incorporates modifications designed to eliminate differences. This process is repeated until the back-translated instrument is "like" (similar to, comparable to) the original source language instrument.

Back-translation has several inherent flaws that preclude the production of data collection instruments with semantic, conceptual, and normative equivalence, and that are reliable, complete, accurate and culturally appropriate. For this reason the guideline does not recommend this approach for producing final data collection instruments.

An approach for survey translation that has recently gained exposure in the literature is the *committee approach* to survey translation and assessment. The committee approach is more comprehensive and collaborative than the other approaches described and discussed in the guideline because the committee approach calls for the pretesting of the translated instrument and relies on input from a team whose members have skill sets that go beyond the skills of a translator.

In general, the committee approach entails convening a group of individuals with complementary skills and who are assigned specific roles. The committee includes several translators, at least one adjudicator, translation reviewers, subject matter specialists, and someone with knowledge and experience in questionnaire design and pretesting. In this approach the actual translation is conducted by a group of translators in a parallel fashion. Several translators independently translate the instrument from the source language to the target language. A meeting is then held with the translators, the translation reviewers, and other members of the team, to discuss the translated versions of the instrument. From this meeting, a reconciled version of the translated instrument is produced which then goes to the adjudicator(s) who made final decisions are made and the data collection instrument is then pretested. After the pretesting is complete the adjudicator and other committee members convene again to decide on final decisions that will result in the final data collection instrument ready for the field. The committee approach is the recommended method of translation.

A more detailed and comprehensive review of the literature is presented in Attachment B of the guideline and a list of the literature consulted appears on the reference section of the guideline.<sup>9</sup>

### **Proposed Translation Guideline<sup>10</sup>**

The Census Bureau developed guidelines for the translation of data collection instruments and supporting materials<sup>11</sup> in order to ensure that such documents translated from a source language into a target language are *reliable, complete, accurate, and culturally appropriate*. When the translated text conveys the intended meaning of the original text the translation is deemed reliable. Translations that *do not* add any new information to the translated document and do not omit information provided in the source document are said to be complete. An accurate translation is one that is free of spelling and grammatical errors. Cultural appropriateness is achieved when the message conveyed in the translated text is appropriate for the target population. In addition to meeting the aforementioned criteria, translated Census Bureau data collection instruments and related materials should also have *semantic, conceptual, and normative equivalence*.

The matter of equivalence and the extent to which there is equivalence (to the source language) in the translated text is central

---

<sup>9</sup> The title of this draft document is: *Census Bureau Guideline: Language Translation of Data Collection Instruments and Supporting Materials*. Methodology and Standards Directorate, U.S. Census Bureau. Draft under review dated August 8, 2003.

<sup>10</sup> This section is based on a longer draft document on the proposed Census Bureau translation guidelines.

<sup>11</sup> In this document "data collection instruments and related materials" refers to questionnaires as well as documents such as cover letters, flash cards, reminder notices, and any other text presented to respondents. In this document we use "questionnaire" and "data collection instrument" interchangeably and for simplicity when we use these terms we are also referring to "related materials" (e.g., reminder notices, flash cards, etc.).

to the quality and appropriateness of a translation. If the translation lacks equivalence then the intended meaning of the information in the source language text is not appropriately conveyed in the translation. The literature points to several types of equivalence. *Semantic equivalence* refers to the extent to which the terms and sentence structures that give meaning to the information presented in the source language is maintained in the translated text. *Conceptual equivalence* concerns the degree to which a given concept is present in both the source and target cultures, regardless of the words used to express the concept. The third main type of equivalence discussed in the research literature is *normative equivalence*. This form of equivalence refers to the extent to which the translated text successfully addresses the difficulties created by differences in societal rules between the source and target culture.<sup>12</sup>

Census Bureau research conducted for the development of the translation guideline showed that there is considerable evidence in the field of survey methodology and cross-cultural research that translated questionnaires which lack the features noted above are *not* of the highest possible quality and that data obtained from such instruments may not necessarily be comparable to data collected using the source language instrument.

The guideline recommends that the translation of data collection instruments from a source language into a target language be conducted using a *translation team*. This recommendation is based on the research literature. This body of knowledge suggests that the task of translating a data collection instrument from a source language into a target language is *not* a solo activity or an exercise performed in relative isolation. Rather, it is a *process* that entails the participation and cooperation of a number of individuals with complementary skill sets and professional experience. They all make valuable contributions that result in high-quality translated data collection instruments.

Translation team members include subject matter specialists and program managers as well as individuals with knowledge of questionnaire design and pretesting. In addition to these individuals, every translation team assigned to produce final versions of Census Bureau translated data collection instruments and supporting materials should involve three different sets of people: *translators*, *translation reviewers*, and *translation adjudicators*.

The guideline relies on the cross-cultural and survey methodology research literature and specified and describes five steps that comprise the *translation process*. These steps are: **P**repare, **T**ranslate, **P**retest, **R**evise, and **D**ocument.

By making the translation guideline, the Census Bureau aims to improve the quality of translated data collection instruments and supporting materials, as well as to ensure that these documents are of a quality comparable to the quality of their English language counterparts for which the Census Bureau is known world wide. This guideline will support the Census Bureau's capabilities in providing high-quality data for its sponsors and data users from respondents with no knowledge or limited knowledge of English.

---

<sup>12</sup> For a discussion of these various forms of equivalence see O. Behling and K.S. Law *Translating Questionnaires and Other Research Instruments: Problems and Solutions*. London: Sage Publications, Inc. 2000.