

The Place of Expert Review in Translation and Questionnaire Evaluation for Hard-to-Count Populations in National Surveys

Goerman, Patricia; Meyers, Mikelyn; García Trejo, Yazmín

Veröffentlichungsversion / Published Version

Konferenzbeitrag / conference paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Goerman, P., Meyers, M., & García Trejo, Y. (2018). The Place of Expert Review in Translation and Questionnaire Evaluation for Hard-to-Count Populations in National Surveys. In D. Behr (Ed.), *Surveying the Migrant Population: Consideration of Linguistic and Cultural Issues* (pp. 29-41). Köln: GESIS - Leibniz-Institut für Sozialwissenschaften. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-58537-2>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC Lizenz (Namensnennung-Nicht-kommerziell) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nc/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC Licence (Attribution-NonCommercial). For more information see: <https://creativecommons.org/licenses/by-nc/4.0>

The Place of Expert Review in Translation and Questionnaire Evaluation for Hard-to-Count Populations in National Surveys

Patricia Goerman, Mikelyn Meyers, & Yazmín García Trejo⁷

Abstract

Many researchers consider it a best practice to include respondent pretesting as part of the survey translation process (Survey Research Center, 2016). The U.S. Census Bureau has a Pretesting Standard that delineates requirements to ensure that any data collection instrument “works,” by verifying that it can be administered as intended by interviewers and understood and responded to appropriately by respondents (U.S. Census Bureau, 2015). Expert review is included in the standard but it is described as a sort of evaluation method of last resort due to its lack of inclusion of respondent pretesting in the process. In large survey organizations, there can be a variety of types of materials in need of translation ranging from actual survey questionnaires to other support materials. Due to resource limitations, expert review is a method that often comes into play in the absence of resources for respondent testing. Recent discussions have involved defining and modernizing the Census Bureau’s approach to translation methodological expert reviews and looking at how they can best fit into the overall translation process. This paper provides a review of the literature on expert review, a description of how this type of review is currently done at the U.S. Census Bureau, along with limitations, challenges, and a plan for future research in order to further develop the method.

1 Introduction

Methodological expert reviews are often used to revise survey materials before they are shown to respondents either for pretesting or when an instrument is fielded. A methodological expert review can have various goals: to incorporate a team-based review of an instrument that did not include review as part of the translation process, or to do a joint methodological review along with a review of the source text. A methodological expert review gives the opportunity to methodologists, translators, and other experts to join forces and provide feedback based on empirical evidence, literature, and experience

⁷ *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.

that will ultimately improve an instrument. Some best practices for incorporating review into the translation process itself have been documented (Survey Research Center, 2016).

Ideally, expert review of survey materials (whether source text or translated) is a preliminary step that precedes respondent testing, but when funding, resources or time do not allow for respondent testing, expert review may be the only “pretesting” method possible. As such, establishing sound best practices for implementing methodological expert reviews of translated survey materials is vital, and this paper seeks to fill a gap in the cross-cultural survey methodology literature in this regard.

2 Review of the Literature

2.1 Hard-to-Count Populations in National Censuses and the Need for Multilingual Questionnaire Development and Evaluation

The fundamental goal when developing surveys for use across linguistic and cultural groups is to ask the same questions across these diverse types of survey respondents. Survey designers strive for comparable understanding by respondents. For example, ideally survey methodologists aim to use questions and terms that an average respondent would be able to interpret and respond to, which will ultimately contribute to collection of more parallel data across diverse populations and will increase the likelihood that respondents understand questions and concepts as methodologists intended.

This topic is particularly relevant when it comes to including hard-to-count populations such as immigrants in national censuses. In a population census, the goal is to survey the entire population and it is therefore important to design survey instruments and supporting materials in as many relevant languages as possible. In order to include as many people as possible in the count, the U.S. Census Bureau, for example, has historically divided languages into “tiers” for which varying levels of support have been provided. In 2010, the highest level of support was provided for languages for which American Community Survey data showed that there were 100,000 or more occupied U.S. housing units with no persons aged 15 or older who spoke English “very well.” Spanish was the most commonly spoken language followed by Chinese, Vietnamese, Korean, and Russian.

In large-scale surveys, there can be a variety of types of materials in need of translation, including actual survey questionnaires, respondent letters and brochures, interviewer instructions, instructional videos, and training manuals for interviewers. In any survey life cycle, there will inevitably be constraints on resources. Materials may need to be prioritized in terms of data collection instruments versus supporting materials and supplementary materials that may not even be seen by respondents. Due to resource limitations, expert review is a method that often comes into play in the absence of ability to conduct respondent pretesting. Basically, if one cannot pretest a survey instrument with respondents through methods such as cognitive interviews or focus groups, researchers are left to rely on expertise and feedback from survey methodologists and language experts. This

paper discusses the ideal role and methodology for expert review in the context of questionnaire translation.

2.2 Review as a Part of the Survey Translation Process

Many researchers consider it a best practice to include review as part of the survey translation process itself (Survey Research Center, 2016). In fact, it is often recommended that translation be done via a team approach as opposed to being done by one translator or through methods such as back translation (Survey Research Center, 2016; Harkness, 2008a; Harkness, 2008b; Harkness, Pennell, & Schoua-Glusberg, 2004; Pan & de la Puente, 2005; Willis et al., 2010). The key players in the team translation method are translators and/or reviewers working together to come up with the final product. The process can involve multiple translators each translating parts of the instrument or each translating the whole instrument independently. The work then includes a review of the draft instrument amongst the group of translators. The team assigns an adjudicator who makes final decisions in the event of disagreement once the team compares their work.

Many researchers recommend that translation and review be embedded in a larger survey development process. Methods such as TRAPD (Translation, Review, Adjudication, Pretesting and Documentation) recommend respondent pretesting as the next step in the development of a translated survey instrument. The Cross-Cultural Survey Guidelines (CCSG) provide a TRAPD graphic (see Figure 1). Many survey researchers point to the critical importance of including respondents from various social locations or backgrounds through pretesting. After all, it can be difficult for questionnaire designers or translators to imagine how respondents with differing characteristics might interpret a survey question.

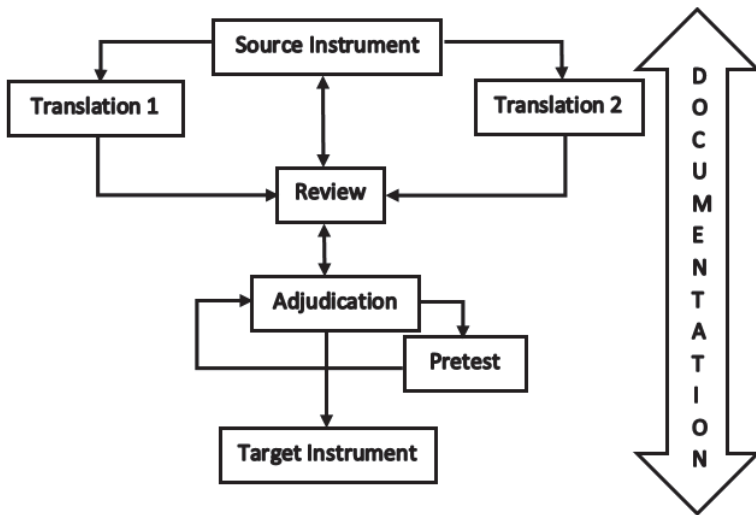


Figure 1 Illustration of the TRAPD model from the Cross-Cultural Survey Guidelines (Survey Research Center, 2016)

Despite team translation and pretesting being the industry standard, there are many agencies that do not have the resources to use the method for every type of survey material needed. In the event that an agency or organization has not been able to employ the team translation method followed by respondent pretesting, some agencies undertake an expert review of translations as a separate step involving multiple reviewers. In some cases, standalone reviews of translations are conducted in preparation for respondent pretesting rather than instead of respondent pretesting. There is a lack of consensus about whether and how to best conduct an expert review of a translation as a questionnaire evaluation method.

2.3 Methodological Expert Review of Translations as a Questionnaire Evaluation Method

Expert review can be defined as a method by which “questionnaire design experts appraise the questionnaire, applying generally accepted questionnaire design principles and knowledge based on their own pretesting experiences” (Willimack, Lyberg, Martin, Japac, & Whitridge, 2004). The goal is to predict interviewer and/or respondent difficulty with the questionnaire items and recommend ways to improve the instrument (U.S. Office of Management and Budget (OMB) Inventory, 2016b).

Many researchers discuss shortcomings of the expert review method (Presser & Blair, 1994; Tourangeau, 2004). In particular, Tourangeau points out that expert review tends to be carried out in a “nonrigorous, even subjective way” (p. 210). Many expert reviews involve assessment by experts but there is a lack of consistency in methods used across experts. Some researchers go so far as to say that “expert review, however systematic, does not provide transparent, empirical, or analyzable data, and cannot be considered a scientific method” (U.S. OMB Inventory, 2016b, p. 13). However, several cognitive appraisal coding assessments have been created which attempt to apply such consistency to the work of expert reviewers. Willis and Lessler’s (1999) Question Appraisal System and Lessler and Forsyth’s (1996) coding system are two such examples.

At the same time, expert review also has its advantages. Many researchers recommend that expert review be done early in the questionnaire design process since it can identify large issues prior to extensive development (Willis, 2005). In addition, it is a low-cost method compared to many other pre- and post-field testing methods (Tourangeau, 2004). Finally, expert review can be the only feasible pretesting for some languages and materials, and while team translation followed by respondent pretesting is the preferred approach, an expert review is the next best method available for ensuring high-quality translations, particularly when the expert review is conducted systematically by multiple reviewers with specialized expertise.

The U.S. Census Bureau has a Pretesting Standard that delineates requirements to ensure that any data collection instrument “works” by verifying that it can be administered as intended by interviewers and understood and responded to appropriately by respondents (U.S. Census Bureau, 2015). Most of the pretesting methods addressed in the Census Bureau standards involve respondent testing. In fact, the Census standards specify that respondent testing is both critical and required in order to identify and resolve any issues with content,

context effects, skip patterns, formatting, and navigation. Expert review is included in the standards but it is described as a method that does not generally satisfy the pretesting requirement. However, it is also described as meeting minimum testing requirements in the event that time and resources do not allow for respondent testing, which is often the case for translated survey instruments and supplementary materials.

3 Expert Review of Translations at the U.S. Census Bureau

The U.S. Census Bureau came out with a Translation Guideline in 2004 and a recommended translation review process in 2009 (Pan & de la Puente, 2005; Pan, 2009). The translation guideline was not a required procedure but rather was developed by Census Bureau staff who had previously collaborated in the development of the CCSG translation guidelines with the Comparative Survey Design and Implementation (CSDI) organization. The Census Bureau guideline was the work of linguists and survey methodologists and details “best practices” to be implemented by the Census Bureau when circumstances permit. Resource constraints have led to translations being conducted using methods besides the team approach, which is a departure from the “best practices” that would ideally be implemented. In response to these resource constraints, we have incorporated both team translation guidelines and survey methodology expertise into the expert review process.

The Census Bureau guidelines list types of people who should ideally be involved in the review process: subject-matter specialists, program managers, survey methodologists with knowledge of questionnaire design and pretesting, translators, and translation adjudicators.

4 Current Practices at the Census Bureau Center for Survey Measurement

The U.S. Census Bureau has a research center called the Center for Survey Measurement (CSM) that is made up of survey methodology experts who work on questionnaire design and pretesting. The Language and Cross Cultural Research (LCCR) group is a sub-team that works on questionnaire design and pretesting of translated instruments. The LCCR group has been receiving an increasing number of requests to do expert reviews of translated survey instruments in recent years.

Methodological expert reviews are somewhat different from review as part of the translation process itself in that they are sometimes done as a substitute pretesting method both for original English source wording as well as for translations. However, the line can get blurred when it comes to the methodological expert review of a translation. The current LCCR process involves having bilingual survey methodologists with language/cultural expertise and a background in survey methodology, social sciences, linguistics and/or translation review the translation. If not included on the official review team, subject-

matter experts and translators are often consulted with questions during the review process. Reviewers conduct an independent review of the materials in question and then meet for a consensus meeting, at which time an adjudicator makes any final decisions in the event that the team is unable to come to agreement. The types of participants involved in a particular project vary based on availability of staff, budget, timeline, and the types of materials being reviewed.

5 The LCCR Translation Methodological Expert Review Method

Step 1: Getting the Expert Team Together

The LCCR team is typically given a translation as a starting point. Some translations are done by an internal Census Bureau translation office and some are conducted by an outside agency. Typically, we do not have information about how the translation was done and whether any review has already taken place. In some cases, the translation may have been done by a single translator with no additional review. The LCCR team's expert reviewers begin by identifying relevant team members (typically two researchers) and a team lead who acts as both a reviewer and the adjudicator. The adjudicator is usually an experienced bilingual survey methodologist. While our teams often consist of a team lead and two supporting reviewers, the decision regarding how many people to staff on an expert review is often driven by staff availability and how many people are needed to fulfill all needed roles: survey methodologist, subject-matter expert, certified translator, program manager, and adjudicator. While it is ideal for bilingual program managers and independent translators, who did not complete the original translation, to participate as official members of the expert review team, in practice many surveys do not have a bilingual program manager, and it is not always possible for a translator who did not complete the original translation to participate. In those cases, the team lead works closely with program managers and translators as questions arise during the review process, but the program managers and translators do not act as official expert reviewers.

Step 2: Kick-off Meeting and Independent Reviews

The team lead creates an electronic repository for the documents under review and leads a kick-off meeting to provide the reviewers with background information about the content of the survey, the intended audience, any research conducted in the past and a thorough explanation about the process of the expert review along with deadlines and assignments. At such meetings, the team lead typically walks the reviewers through the source text and translations, the scope of the review, applicable context regarding survey operations, information about how the translation was produced, and any applicable reference materials (e.g., online medical dictionaries, existing translations of common batteries of survey questions, etc.).

Regardless of the format used to record reviewer comments, the review process should be completely independent, such that individual reviewers are blind to the comments of other reviewers until their review is complete. This ensures that reviewers, including the adjudicator, are not biased by referencing the comments of others.

Step 3: The Consensus Meeting

In the LCCR team, the team lead is typically responsible for combining reviewers' comments prior to consensus meetings into one document to facilitate group discussion. When time permits, consensus meetings are conducted with all reviewers, and comments are discussed one-by-one. Reviewers reach a consensus on comments during the meeting itself or decide on next steps, e.g., designating a particular reviewer to research a given term and report findings back to the group. The team lead rarely adjudicates disagreements between reviewers, as most disagreements can be resolved by coming to agreement through discussion or by further research in reference materials.

When the length and complexity of the materials as well as conflicting staffing demands do not permit all reviewer comments to be discussed individually during consensus meetings, the team lead prioritizes comments, looking for those that require group discussion: Where reviewers are in disagreement, or where a single reviewer had concerns about a particular item not shared by other reviewers, the team lead flags the item for group discussion. This process of prioritizing comments for discussion is necessarily subjective, and as such requires a team lead with familiarity with the target language, survey methodology, and subject-matter expertise.

Step 4: Tracking Decisions and Preparing the Recommendations

The team lead tracks decisions on reviewer comments during consensus meetings and distributes the document for team review either following each consensus meeting or after all consensus meetings are complete. Final documentation includes the source text, the original and proposed translations, back translations⁸ of new recommendations in the event that the client does not speak the target language, along with an explanation for proposed revisions. Typical explanations document the reasons that revisions are being recommended. For example, there may be 1) missing or additional concepts in the source text compared to the target language text, 2) mismatches between source text and target

8 The term “back translation” is best known as a description of a translation method where one translator translates source version A into target language B. A different translator then translates version B back into source language A so that the two source language versions can be compared by monolingual stakeholders to look for errors. This is not a highly recommended translation method (see Behr, 2017). We use the term “back translation” here to describe a sort of literal or word-for-word translation that can be used to explain to monolingual survey sponsors what newly recommended target language wording means. Our “back translations” of newly recommended wording are often provided along with an explanation of the meaning and context and reasons for recommending the change. This method enables a survey sponsor who does not speak the target language to participate in the decision-making process in terms of whether to accept recommendations.

language text register (i.e., formality or complexity of terms), 3) concerns that target language terminology may not be well understood by respondents of a variety of national origins when applicable (e.g., Spanish or Arabic), 4) grammatical errors, 5) readability problems, etc. In addition to including the recommendations from the expert review, final documentation also outlines the methodology used to conduct the expert review.

Challenges of conducting such expert reviews are manifold. Given the amount of information that must be conveyed about each individual recommendation (original source text, original target language text and back translation, recommended target language text and back translation, rationale for recommending a change), documentation is a major hurdle in any expert review. Establishing procedures for each reviewer to provide comments in a format that facilitates the combination of comments is vital, while careful note-taking during consensus meetings is also necessary to producing a high-quality final product. When possible, documentation from sponsors tracking which changes were implemented and why can be helpful for targeting future research on problematic items when schedules or resources permit respondent testing.

Step 5: Final Recommendations

Presenting final recommendations to clients is also challenging given that many stakeholders are not proficient in the target language and they may lack experience with survey translations. LCCR typically walks clients through our recommendations to ensure that all relevant information has been conveyed in a manner that is understandable and that allows stakeholders to make informed decisions about potential changes to their survey instruments. There are some types of findings, such as problems detected in both the original and source language versions of the questionnaire that can be out of scope for our translation reviews. While recommendations maybe valid, they often cannot be implemented due to constraints in the timing of the translation expert review during the survey life cycle, where the original English wording might already be final.

While expert review of a translation is a valuable step during the development of any translated survey instrument, it is particularly valuable when the translation was not conducted via the team approach. The expert review process allows an approximation of the team approach to be included in the review process itself. Additionally, expert review can improve translated survey instruments in advance of conducting respondent pretesting so that respondents do not waste time grappling with issues that could be addressed prior to respondent pretesting. When changes to translated survey instruments cannot feasibly be made prior to respondent pretesting (e.g., programming schedules do not allow updates to online instruments), expert review of translated survey instruments can inform the design of pretesting protocols to target suspected problems with the translation prior to conducting respondent pretesting. Finally, while testing translated survey instruments with respondents is the gold standard of respondent pretesting, when such testing is not possible due to timing or resource constraints, expert review is a vital step in the survey life cycle.

Selected Examples of Translation Expert Review Recommendations

Expert reviews can catch many important potential questionnaire issues, both in materials that are newly translated as well as in materials that were previously translated and have recently undergone additional pretesting. In a recent study, we conducted an expert review of translated Census Bureau Group Quarters materials in Spanish. Group quarters are places such as college dormitories, prisons, shelters or residential treatment centers. Reviewers noticed that the name of the form itself may present some issues. The form in English was called “Individual Census Report.” Spanish-speaking reviewers were concerned that the term “report” in both English and Spanish might sound threatening to respondents and recommended instead “Individual Census Questionnaire.” In addition, the term “facility” was translated as “facilidad” throughout the questionnaire materials. Reviewers pointed out the fact that “facilidad” in Spanish is a noun that means “ease” and “ability” and it does not mean facility and/or group quarters as in English. They felt that the use of “facilidad” in the instrument in Spanish in this context was an informal Anglicism, which is inappropriate in an official government survey. The review team recommended changing these terms and then conducting respondent testing as a next step.

In another recent study, LCCR conducted an expert review of a Spanish translation prior to cognitively testing the translation with Spanish-speakers. The English and translated wording were as follows (Table 1):

Table 1 English – Spanish translation of “break the law”

Source Text	Original Translation	Back Translation
Is the census used to help the police and FBI keep track of people who break the law, or is it not used for this?	¿El censo se usa para ayudar a la policía y al FBI a mantener un registro de las personas que infringen la ley o no se usa para este fin?	[Is the census used to help the police and the FBI maintain a registry of people who infringe on the law, or is it not used for that purpose?]

Expert reviewers commented that the term “infringir” [infringe] in the translation was comparatively a higher register term than “break” in the English source text, and recommended using a lower register term like “romper” [break] for consistency with the English. This change was not implemented prior to cognitive testing. Findings from cognitive testing indicated that three of ten Spanish-speakers expressed concerns about the term “infringir” and felt that it would be difficult for many Spanish-speakers to understand. Findings from this cognitive testing project validated the recommendations made by expert reviewers.

6 Limitations to the Current Process

There are several limitations to our current process. For example, we do not always have the same number and types of staff members available to perform an expert review. Some surveys are about complex, specialized topics such as health conditions or social programs. It can be difficult to find bilingual survey methodologists, linguists or translators with particular knowledge of the subject-matter.

Some instruments under review are “legacy” instruments involving time series that are not easily changed. Similarly, in many cases, expert reviews of a translation occur separately from a review of the English source language version and at a time when only minimal changes can be made to the English version. This limits the scope of recommendations that can be implemented.

7 Current Issues/Challenges

While this paper outlines the procedures that LCCR uses to carry out methodological and translation expert reviews, more work remains to be done comparing practices within and across agencies and organizations, and more importantly evaluating those models to develop more formal best practices.

Separately from the issue of which procedures result in the most successful expert reviews is the question of why expert reviews are often the only form of pretesting for many translated materials. While this decision is often driven by resource and time constraints, new project schedules and budgets should ideally address the needs of translated materials at the outset rather than fitting the design, review, and testing of these materials into schedules and budgets after the design, review, and testing of English source materials have already been scoped out. Ideally, source and target language materials should be developed and tested in parallel in order to allow for a higher quality final product and to ensure that testing of target language materials is not reduced due to cost or schedule over-runs.

8 A Program of Research for the Future

LCCR has recently begun experimenting with our expert review procedures in an effort to formalize best practices. In one recent project, we experimented with the inclusion of bilingual field interviewers as expert reviewers. In this evaluation, we coded the interviewers’ comments to determine at what point additional reviewers provided diminishing returns as well as the extent to which interviewers who have contact with the general public but lack survey methodology, program, and subject-matter expertise are able to contribute actionable comments to the review process (Goerman, Meyers, & García Trejo, 2018).

As it becomes more common and necessary for surveys to provide translations of their materials in order to cover more of the population, LCCR is being asked to do more expert reviews of translated survey materials. Staffing constraints sometimes require that reviews be completed without at least three independent reviewers and a series of consensus meetings. LCCR is experimenting with models of expert review that require less staff participation while meeting a minimum quality standard for those projects with smaller budgets or tighter schedules. CSM continues to advocate for respondent testing as the gold standard of evaluating translated survey materials, while acknowledging that real-world constraints make establishing best practices for expert reviews a much-needed initiative in cross-cultural survey research.

References

- Behr, D. (2017). Assessing the use of back translation: The shortcomings of back translation as a quality testing method. *International Journal of Social Research Methodology*, 20, 573–584.
- Survey Research Center (2016): Guidelines for best practice in cross-cultural surveys. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. Retrieved from http://ccsg.isr.umich.edu/images/PDFs/CCSG_Full_Guidelines_2016_Version.pdf
- DeMaio, T. J., & Landreth, A. (2004). Do different cognitive interview techniques produce different results? In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 89–108). New Jersey: Wiley.
- Forsyth, B. H., & Lessler, J. T. (1991). Cognitive laboratory methods: A taxonomy. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 393–418.) New York: John Wiley & Sons.
- Goerman, P., Meyers, M., & Garcia Trejo, Y. (2018). Survey translation and the place of expert review in the question evaluation tool kit. Presentation given at the Federal Committee on Statistical Methodology (FCSM) Research and Policy Conference. Washington, D.C. March 7 – 9, 2018
- Groves, R. M., Fowler, Floyd J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Hoboken, NJ: Wiley and Sons.
- Hansen, S. E., & Couper, M. P. (2004). Usability testing to evaluate computer-assisted instruments. In S. Presser, J. M. Rothgeb, M. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 337–360). Hoboken, NJ: John Wiley & Sons.
- Harkness, J. A. (2008a). Round 4 ESS translation strategies and procedures. Retrieved from http://www.europeansocialsurvey.org/docs/round4/methods/ESS4_translation_guidelines.pdf

- Harkness, J. A. (2008b). Comparative survey research: Goals and challenges. In E. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 56–77). New York, NY: Psychology Press Taylor & Francis Group.
- Harkness, J. A., Pennell, B.-E., & Schoua-Glusberg, A. (2004). Survey questionnaire translation and assessment. In S. Presser, J. Rothgeb, M. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 453–473). Hoboken, NJ: John Wiley & Sons.
- Lessler, J., & Forsyth, B. (1996). A coding system for appraising questionnaires. In N. Schwartz, & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco: Jossey-Bass.
- Pan, Y., & de la Puente, M. (2005). *Census Bureau guidelines for the translation of data collection instruments and supporting materials: Documentation on how the guideline was developed*. Research Report Series #2005-06. Statistical Research Division, Washington, DC: U.S. Bureau of the Census.
- Pan, Y. (2009). *A process for reviewing translations of data collection instruments and related materials*. U.S. Census Bureau Survey Methodology Study Series #2009-09. Retrieved from <https://www.census.gov/srd/papers/pdf/ssm2009-09.pdf>
- Pan, Y., Sha, M. M., Park, H., & Schoua-Glusberg, A. (2010). 2010 Census language program: Pretesting of Census 2010 questionnaires in five languages. Research Report Series (Survey Methodology #2009-01) Retrieved from <http://www.census.gov/srd/papers/pdf/rsm2009-01.pdf>
- Presser, S., & Blair, J. (1994). Survey pretesting: Do different methods produce different results? In P. Marsden (Ed), *Sociological methodology* (pp. 73–104). San Francisco: Jossey-Bass.
- Ramirez, C. (2002). Strategies for subject matter expert review in questionnaire design. International Conference on Questionnaire Development, Evaluation and Testing Methods (QDET), Charleston, SC.
- Tourangeau, R. (2004). Experimental design considerations for testing and evaluating questionnaires. In S. Presser, J. Rothgeb, M. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 209–224). Hoboken, NJ: John Wiley & Sons.
- U.S. Census Bureau Statistical Quality Standard A2: Developing data collection instruments and supporting materials (2015). Retrieved from <https://www.census.gov/about/policies/quality/standards/standarda2.html>
- U.S. Office of Management and Budget (2016a). *OMB Statistical Policy Directive No. 2 Addendum: Standards and Guidelines for Cognitive Interviews*. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/omb/inforeg/directive2/final_addendum_to_stat_policy_dir_2.pdf
- U.S. Office of Management and Budget (2016b). *OMB Statistical Policy Working Paper 47: Evaluating Survey Questions: An Inventory of Methods*. Statistical and Science Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget. Retrieved from <https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/spwp47.pdf>

- Willimack, D. K., Lyberg, L., Martin, J., Japac, L., & Whitridge, P. (2004). Evolution and adaptation of questionnaire development evaluation, and testing methods for establishment surveys. In S. Presser, J. Rothgeb, M. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 385–407). Hoboken, NJ: John Wiley & Sons.
- Willis, G., & Lessler, J. T. (1999). *The BRFSS-QAS: A guide for systematically evaluating survey question wording*. Rockville, MD: Research Triangle Institute.
- Willis, G. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage Publications, Inc.
- Willis, G. B., Kudela, M. S., Levin, K. L., Norberg, A., Stark, D. S., Forsyth, B. H., ... Hartman, A. M. (2010). Evaluation of a multistep survey translation process. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. Ph. Mohler, B.-E. Pennell, & T. W. Smith (Eds.), *Survey methods in multinational, multicultural and multiregional contexts* (pp. 137–152). Hoboken, NJ: John Wiley & Sons.