Survey Translation and the Place of Expert Review in the Question Evaluation Tool Kit

Patricia Goerman, Mikelyn Meyers and
Yazmin Garcia Trejo
Center for Survey Measurement, U.S. Census Bureau

Presentation given at the Federal Committee on Statistical Methodology (FCSM) Research and Policy Conference
Washington, D.C.
March 7 – 9, 2018

Roadmap of the talk

1. Background on expert review / review of literature

2. Discussion of current challenges/research goals

3. A case study: NHIS 2019.

    – Trying out a coding scheme/method of evaluation across reviewers

4. A Program of Research for the Future

Introduction: Expert Review

- Method by which survey methodologists evaluate a questionnaire applying questionnaire design principles and knowledge based on pretesting experiences (Willimack et al 2004)
- Goal: to predict respondent difficulty with questionnaire items and recommend ways to improve instrument
- Ideal to follow with respondent pretesting
  - Pretesting Inventory-OMB 2016
  - Pretesting Standards-Census Bureau 2015

Multilingual Surveys

- Fundamental goal of surveys that have multiple language versions: elicit the same type of information from diverse types of respondents. Parallel data and inclusion.
- Variety types of materials in need of translation
  - Actual survey instruments, interviewer instructions
  - Respondent letters, brochures, supplementary materials
  - Instructional videos and other training manuals for interviewers
- Question: when is expert review alone enough?

Review of the Literature

Two types of translation review:

1. Review as part of the survey translation process (TRAPD)
2. Methodological expert review as an evaluation method

Review as Part of the Survey Translation Process: the TRAPD Model

- TRAPD model: Translation, Review, Adjudication, Pretesting and Documentation
- Translation via team or committee approach with independent review and adjudication in case of disagreement (Cross cultural Survey Guidelines, 2016).
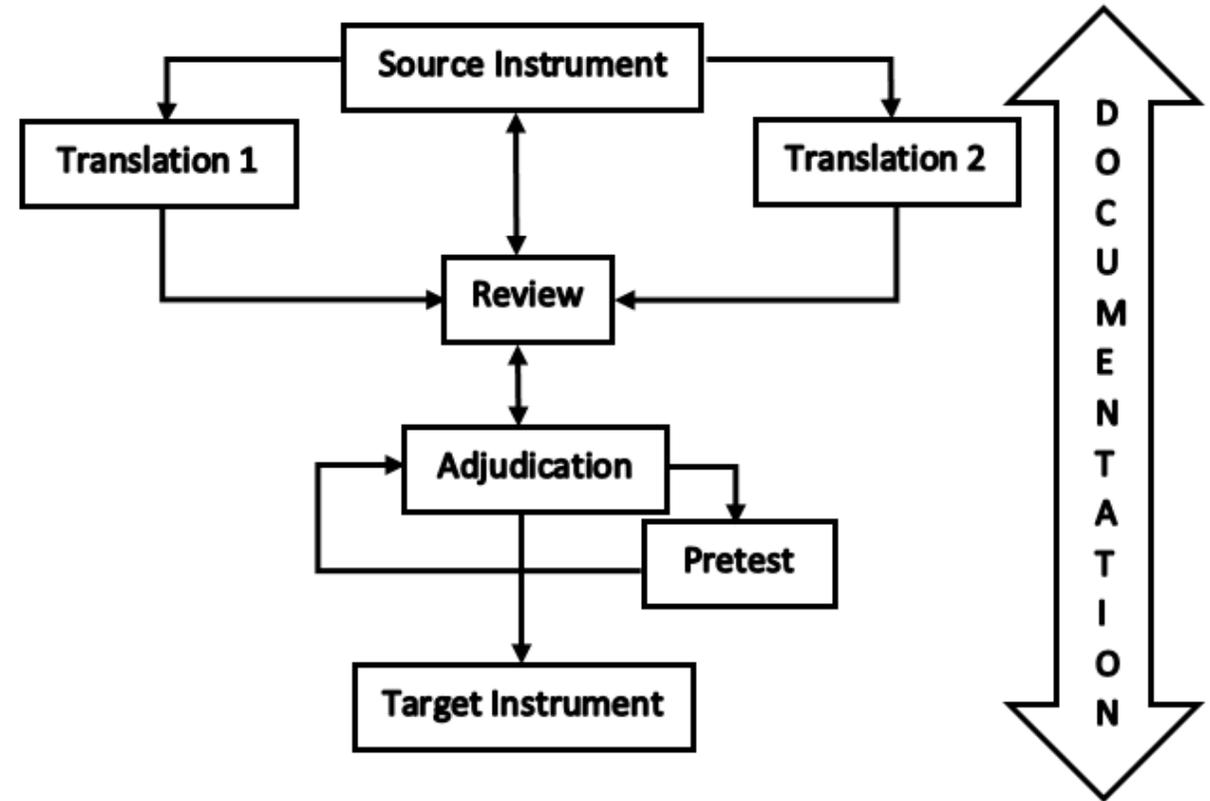
Methodological Expert Review of Translations as a Questionnaire Evaluation Method

- Goals:
  - Incorporate team based review of instrument that did not include review as part of translation process
  - Receive feedback from variety of experts: survey methodologists, translators, subject matter experts, language/cultural experts
- Limitations: can be nonrigorous, subjective (Tourangeau, 2004)
- Lack of consistency in methods across experts

Benefits of Expert Review of Translations

- Possible to be systematic, e.g. Cognitive appraisal methods such as Willis and Lessler (1999); Lessler and Forsyth's (1996)

- Useful early in questionnaire design cycle

- Low cost method compared to many others

- Maybe only choice in situations of limited time/resources

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

- Center for Survey Measurement: survey design and pretesting
- Variation in expert review processes even within one center
- Language and Cross Cultural Research (LCCR) Group
  - Sub-team focused on non-English survey design and pretesting
- Documentation:
  - Translation guideline (2004)
  - Translation review process (2009)

The Language and Cross Cultural Research (LCCR) Group Translation Review Method
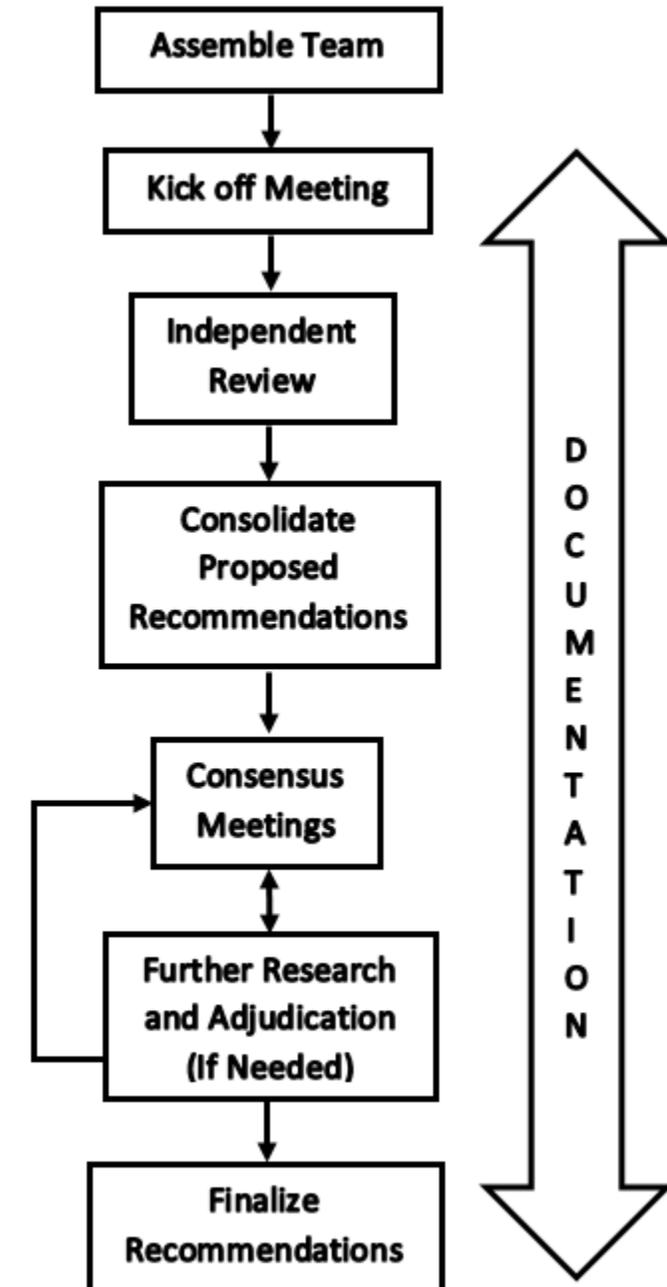
Step 1: Assemble the team

  – Team lead plus 2 additional researchers

Step 2: Kick-off meeting

Step 3: Independent review

Step 4: Consensus meeting(s)

Step 5: Final recommendations

Assemble Team

Kick off Meeting

Independent Review

Consolidate Proposed Recommendations

Consensus Meetings

Further Research and Adjudication (If Needed)

Finalize Recommendations

DOCUMENTATION

Current Issues/Challenges

- Same number and types of staff not always available
- Instruments can have source language "frozen."
  - Cross-cutting issues cannot be addressed
- Lack of respondent testing and desire to encourage sponsors to build in time and money for respondent testing of translations
- Need for research on best practices/methodology
  - Number and types reviewers
  - Best process to follow

Evaluation of Translation Expert Reviews:
A Case Study from NHIS

- The Translation
  - The 2019 National Health Interview Survey (NHIS) instrument was redesigned, resulting in new translations
  - Respondent testing of the translation was cost prohibitive
  - Expert review requested
- Goals of review: Incorporate LCCR and sponsor processes together: Inclusion of interviewers as reviewers
  - Gather feedback from bilingual Census Bureau interviewers on the Spanish translation
  - Gain buy-in from interviewers to ensure compliance with new translation

Methods

- Interviewer Participants
  - 12 bilingual interviewers trained to participate in electronic expert review
  - 7 of 12 ultimately provided comments
- Interviewer Review Process
  - Interviewers trained by bilingual survey sponsor SME over the phone with a manual, including practice exercise
  - Recommendations submitted to LCCR in excel spreadsheet
- LCCR Review Process
  - LCCR conducted internal expert review independently using our usual process (without the training)
  - LCCR team lead coded interviewer recommendations and compared to LCCR expert review recommendations

Limitations

- LCCR staff did not lead interviewer training but reviewed training materials
- Interviewer recommendations evaluated by only one coder
  - No interrater reliability check
- Interviewer drop-off in participation (for unclear reasons)
  - From 12 recruited to 7 who completed review

Coding Scheme

Code	Description
Incorrect or non-parallel recommendation	Inaccurate due to adding or removing information compared to source text, or grammatically incorrect
Stylistic recommendation	Non-substantive change that LCCR reviewers recommend not implementing (e.g., a preference for one term over another when they are functionally equivalent)
Confirmative recommendation	A valid recommendation but one that LCCR reviewers had independently made
Actionable/ additional recommendation	A valid recommendation that LCCR had not made ourselves and that was incorporated into final recommendations

Findings

- Interviewers submitted a total of 301 recommendations:
  - 43% were coded as Incorrect or Non-Parallel
  - 23% were coded as Stylistic recommendations
  - 29% were Confirmative of LCCR reviewers' recommendations (i.e., redundant)
  - 6% were coded as actionable/additional

Discussion

- Field interviewers may have varying levels of formal education in Spanish, which can lead to some recommendations that are ungrammatical

- Interviewers may have little familiarity with survey research methodology, which can lead to recommendations that could contribute to measurement error (e.g., unnecessarily altering common response scales)

- Interviewers tend to provide more stylistic comments

Findings by Interviewer

Interviewer	Incorrect or non-Parallel Recommendation	Stylistic Recommendation	Confirmative Recommendation	Actionable-Additional Recommendation
Interviewer 1	33	3	4	5
Interviewer 2	6	5	12	3
Interviewer 3	4	18	20	1
Interviewer 4	19	15	6	1
Interviewer 5	65	17	40	4
Interviewer 6	1	9	4	3
Interviewer 7	1	1	0	1
Total (N = 301)	**129**	**68**	**86**	**18**
% of Total	**43%**	**23%**	**29%**	**6%**

Discussion

- Interviewers are essentially laymen, and some were better at this task than others
- Low number of actionable comments (6%) indicate little added value for the cost of including so many additional sets of eyes
  - time to train interviewers
  - time to format their recommendations in multiple documents
  - time to evaluate each recommendation
- But it is difficult to price out the value of gaining "buy-in" from field staff.

Incorporating Interviewer Feedback through Focus Groups

- For a later review, LCCR conducted focus groups with same interviewers to see if they could be included in expert review process more efficiently
  - Most of the feedback was not directly related to the translation
  - For example, interviewers shared respondent concerns about confidentiality
  - Regardless, interviewers appreciated the opportunity to provide feedback

A Program of Research for the Future

- Test out use of an expert review coding scheme such as Cognitive appraisal methods, Willis and Lessler (1999); Lessler and Forsyth (1996) in context of translation expert review
- Similar coding could be conducted to examine types of expert review comments made by experts with varying skill sets
  - Linguists
  - Independent translator reviewers
  - Bilingual survey methodologists
  - Subject matter experts

Survey Translation and the Place of Expert Review in the Question Evaluation Tool Kit

Thank you!

For further discussion/information:

Email: Patricia.L.Goerman@census.gov

Works Cited

- Cross-Cultural Survey Guidelines (CCSG): Guidelines for Best Practice in Cross-Cultural Surveys, Full Guidelines (2016). http://ccsg.isr.umich.edu/images/PDFs/CCSG_Full_Guidelines_2016_Version.pdf
- Lessler, J. and Forsyth, B. (1996). A Coding System for Appraising Questionnaires in Schwartz, N. and Sudman, S. (eds.) <u>Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research.</u> San Francisco: Jossey-Bass.
- OMB Statistical Policy Working Paper 47: Evaluating Survey Questions: An Inventory of Methods (2016). Statistical and Sciency Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget. https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/spwp47.pdf
- Tourangeau, R. (2004). "Experimental Design Considerations for Testing and Evaluating Questionnaires," in Presser, S., Rothgeb, J.M., Couper, M., Lessler, J.T., Martin, E., Martin, J., and Singer, E. (eds.). *Methods for Testing and Evaluating Survey Questionnaires* (pp. 209–224). Hoboken, NJ: John Wiley & Sons.

Works cited (cont)

- U.S. Census Bureau Statistical Quality Standard A2: Developing Data Collection Instruments and Supporting Materials (2015) https://www.census.gov/about/policies/quality/standards/standarda2.html

- U.S. Census Bureau Translation Review process citation: Translation review process (2009) https://www.census.gov/srd/papers/pdf/ssm2009-09.pdf

- U.S. Census Bureau Translation Guideline (2004) https://www.census.gov/srd/papers/pdf/rsm2005-06.pdf

- Willimack, D.K., Lyberg, L., Martin, J., Japec, L., and Whitridge, P. (2004). "Evolution and Adaptation of Questionnaire Development Evaluation, and Testing Methods for Establishment Surveys," in Presser, S., Rothgeb, J.M., Couper, M., Lessler, J.T., Martin, E., Martin, J., and Singer, E. (eds.). *Methods for Testing and Evaluating Survey Questionnaires* (pp. 385–407). Hoboken, NJ: John Wiley & Sons.

- Willis, G., and Lessler, J.T. (1999). *The BRFSS-QAS: A Guide for Systematically Evaluating Survey Question Wording*. Rockville, MD: Research Triangle Institute.