



City Research Online

City St George's, University of London

Citation: Schoua-Glusberg, A. & Villar, A. (2014). Assessing Translated Questions via Cognitive Testing. In: Miller, K., Willson, S., Chepp, V. & Padilla, J. L. (Eds.), Cognitive Interviewing Methodology. (pp. 51-67). Hoboken, USA: John Wiley & Sons. ISBN 9781118383544 doi: 10.1002/9781118838860.ch5

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/14428/>

<https://doi.org/10.1002/9781118838860.ch5>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

5

Assessing Translated Questions via Cognitive Testing

ALISÚ SCHOUA-GLUSBERG

Research Support Services

ANA VILLAR

City University London

INTRODUCTION

This chapter presents arguments showing that cognitive testing provides crucial information regarding how translated questions (as well as other texts used in surveys) are understood, the underlying cognitive processes that respondents undergo to answer them, and how this relates to respondents interpretation of the same questions in the source language. This information can help improve wording choices in each language used in a study, increase cultural appropriateness of the text, as well as evaluate and improve comparability between translated questions and the original text in ways that cannot easily be achieved by requesting input from bilingual expert reviews.

The first section describes challenges found in multilingual research and different approaches to instrument production in multilingual research. The next section provides some context by describing translation procedures and translation assessment techniques for survey research. The third section reflects on the need for pretesting as part of the translation assessment procedures. The fourth section discusses idiosyncrasies of cognitive testing of survey translations and the fifth section presents examples of problems discovered through the use of cognitive testing of survey translations.

WHY USE COGNITIVE TESTING IN MULTILINGUAL SURVEY RESEARCH

As it has been described in previous chapters, cognitive testing of questionnaires is an indispensable step in the process to assure collection of high quality survey data. Yet, most often, in multilingual studies cognitive testing only occurs at best in the language in which the instrument was designed. The underlying assumption is that, if a question performs as intended in the source language (tapping into the construct of interest, eliciting a codable, accurate and unbiased response that is supported by the narrative elicited from a respondent) then translators can produce a question version in the target language that also "works" well. However, just as questionnaire designers cannot always anticipate problems with questions that arise when questions are asked of actual respondents, translators and survey translation experts cannot fully anticipate how respondents will interpret and use the translated questions. This section describes the context surrounding multilingual survey research and how the characteristics of this context justify the need for cognitive testing.

Multilingual Research Settings

Multilingual surveys appear in a number of different research settings. Comparative cross-national or cross-cultural research projects, for example, often require asking questions of samples that speak different languages. In research within one country this can also be true if the target population includes speakers of different languages, whether these are countries where native born populations are speakers of

varied languages (e.g., Ghana, Switzerland, the Philippines) or where sizable immigrant populations reside (e.g., the United States, France, Germany). The European Social Survey (ESS) is an example of cross-national research where translation is necessary both because it involves countries where different languages are spoken and countries where more than one language is spoken. Each country in the ESS, therefore, produces a version of the questionnaire in each language spoken as first language by at least 5% of the national population: in some countries this 5% represent “official” languages (such as Catalan in Spain), whereas in other countries this 5% of the population speaks a foreign language (such as Hungarian in Slovakia or Russian in Israel) (European Social Survey 2012). In both types of research, comparability of the survey instruments is crucial for the goals of the project¹.

Comparability can be threatened, among other things, by a number of problems related to how a translation performs. The main reasons why questions might not perform as intended are: a) problems arising from translation choices, b) cultural factors influencing interpretation even in “perfectly” translated questions, and c) lack of construct overlap between the language/culture of the translation and the language(s) for which the questions were designed. We will describe these types of errors in more detail in the section on analyzing cognitive interviews of translated instruments.

Comparability, thus, cannot just be assumed; it needs to be actively pursued through careful research design and assessed by gathering evidence of comparability (or lack thereof) where possible (Van de Vijver and Leung 1997). Cognitive testing of translations is one way to assess how questions will work in the target language(s), what types of errors are present in each language version, and whether questions perform comparably across language versions. Cognitive testing will also provide information about the causes of those problems in question performance that can guide recommendations and point to potential remedies. Thus, failing to conduct cognitive testing in multilingual settings means that the fielded question may not perform in the new language version(s) as intended.

Instrument Production in Multilingual Settings

Depending on the model followed to produce the different language versions, translation may start before or after the source questionnaire is finalized. The most commonly found approach in survey research follows a sequential development, where a source questionnaire is produced and finalized before other language versions are produced (Harkness, van de Vijver and Johnson 2003). Traditionally, this model did not involve translation during source questionnaire production.

More recent applications of this model, however, try to incorporate cross-cultural input earlier in the process, which can involve translation before the source questionnaire is completed. For example, Harkness and Schoua-Glusberg (1998) and Braun and Harkness (2005) advocated the use of advance translation, a procedure where rough translations are produced before the source questionnaire is finalized as a means of early detection of comparability issues in the questions. Two international survey programs, the International Social Survey Programme (ISSP) and the ESS, have successfully

¹ Survey translation might also take place in two settings where instrument comparability may or may not be a research goal. The first is monolingual research where existing questions that were designed in a different language are borrowed. The second is survey research commissioned by researchers who are not proficient in the language(s) of the target population(s), and thus design the instrument in their own language (as in studies conducted by U.S. or European researchers in the developing world). Even when the goal of such studies might not be comparability with another population, cognitive testing of translations would serve the same functions that it has in monolingual research (see chapters 3 and 7, this volume), making sure that the questions as rendered in the language(s) of administration measure what the survey designer intended.

implemented this approach in the past, leading to important changes in the source questionnaire (Dorer 2012). Also in the ESS, draft questions are translated and tested through cognitive interviews and piloting in languages other than the source language before the source questionnaire is finalized (European Social Survey 2012).

The focus of this chapter is on the advantages of using cognitive testing for language versions translated from a *finalized* source questionnaire. At the same time, we argue that the use of cognitive testing across cultures and languages *while the source questionnaire is still being developed* can help design an instrument that is easier to translate, and avoid complicated situations if similar stimuli cannot be conveyed in a way that grants measurement equivalence unless the source question is modified. This could prevent problematic situations where cognitive testing of translations reveals problems with source questions that are considered finalized and researchers reluctant to make changes to it based on recommendations arising from cognitive testing or to allow for the necessary adaptations of the target questions.

Other models of multilingual instrument production such as decentering (Werner and Campbell 1970) use translation during the source question design stage. In the decentering approach, text is translated back and forth between two (and potentially more) languages until all versions are found to be equivalent; the goal is to avoid questions that are biased towards one of the cultures or languages involved. In this approach, the concepts of a “source” and a “target” question becomes unnecessary. However, comparison of how questions work with the intended population through pretesting remains necessary in order to overcome the limitations related to expert reviews.

TRANSLATION AND TRANSLATION ASSESSMENT PROCEDURES

While the focus of this chapter is not to review existing practices for instrument translation, a brief look at survey translation methodology provides a good background for a discussion on testing translated questions. Most current approaches to survey translation advocate the use of some type of team collaboration and pretesting as part of the translation assessment process (e.g., Acquadro et al 2004, Guillemin et al 1993, Harkness 2003, Harkness and Schoua-Glusberg 1998, U.S. Census Bureau 2005, Van Widenfelt et al 2005). In particular, the TRAPD model has been discussed and implemented in cross-national survey research programmes such as the ISSP, the ESS, and the Survey for Health, Ageing and Retirement in Europe (SHARE).

Team Translation Approaches

As mentioned before, in this chapter we focus on research projects where comparability is key and which follow an “ask-the-same-question” model. This model is based on keeping the “meaning” of translated questions as similar as possible to the source text and maintaining the stimulus constant across languages, while keeping the question format and measurement properties constant (Harkness et al 2010).

Harkness (2003) describes the five functions that form the TRAPD model: **T**ranslation, **R**eview, **A**djudication, **P**retesting and **D**ocumentation. In this model, two or more individuals translate the instrument into the desired target language. The translators and at least one reviewer then meet to review the original translation(s) and make comments on issues they find or changes they recommend. An adjudicator (who may or may not be at the review meeting) will ultimately decide whether to adopt the changes or recommendations or make other changes based on reviewer(s)' findings. Then the reviewed translated document is pretested. Throughout the process, decisions made at every step are documented

to inform designers and analysts about how the final translation was reached. While other approaches may include all the steps in TRAPD, team or committee approaches have the translation, review and adjudication steps built into the method.

Translation Assessment Procedures

In ask-the-same-question models, translation is required to render the same questions in all languages and offer the same response options as the original questionnaire (Harkness and Schoua-Glusberg 1998). But how do we test for that 'sameness' across languages?

Assessment of translations is embedded in each of the steps of the TRAPD model. In doing their task, for instance, translators worry about a number of factors:

- Is x the best way to translate this question (or this term)?
- Will respondents understand this translated question? Is the type of language appropriate for the intended audience?
- Are the nuances of the original question text maintained in the translation?
- Does the wording of the question feel natural?
- Are the response choices likely to be understood and used as in the original question?

Such considerations lead to numerous changes to wording during the translation stage.

At the review meeting, once again, discussion of challenges in the translation process as well as expectations of how questions will perform among respondents happen in a real-time discussion among translators and other experts, and comments will be considered and used to guide review of the translated text.

These *review* steps give translators and researchers answers to translation challenges but only from the point of view of experts (such as linguistic, methodological, substantive, cultural experts). Expert review, however, can yield many different and opposing opinions, leaving the researcher with the difficult task to reconcile them. Thus, translation review is akin to expert review in question design; it will tell us how bilinguals evaluate a translation (Survey Research Centre 2010) but the “true” answers to problems with question wording can only be appropriately explored by talking to members of the target population for the study, that is, by interviewing respondents similar to those who will be administered the translated questions. *Monolingual feedback* (Harkness and Schoua-Glusberg 1998) or monolingual review can be useful to know whether a translation sounds idiomatic and conveys the intended meaning in the target language, as evaluated through a population similar to those in the study. This makes pretesting an essential step in adequate survey translation approaches. Yet, by virtue of being reviewed by monolinguals, it cannot shed any light as to whether the translation seems, even at face value, to be asking the same as the original questionnaire. Therefore input from both bilinguals and monolinguals are necessary for proper translation assessment.

Pretesting as Part of Translation Assessment

Respondents of a translated questionnaire in many cases cannot speak the source language of the questionnaire (or they speak it but live in a country where a different language version is available); we will refer to these individuals as monolinguals, even though they might speak other language(s). Given that they do not know the source language, monolinguals cannot be asked about the extent to which a translated question matches the intent of a source question. Even if respondents happen to speak the

source language, their assessment of the translation adequacy would be a matter of opinion, and would leave us no better (probably worse off) than relying on translation expert review. However, respondents speaking the target language can actually provide much richer information by allowing the researcher to peek into their cognitive processes as they answer the translated questions, as they elaborate on their answers thereby letting us understand how they interpret each item.

As when testing is done in a monolingual study, the main and primary goal of pretesting of translated questions is to uncover patterns of interpretation and respondents' answering strategies, and to evaluate the adequacy of the survey instruments. When testing translations, however, we also need to make sure that other factors besides those strictly semantic are comparable in all language versions. Specifically, pretesting needs to investigate many issues that can arise when translating survey questions, related to how these function pragmatically across all languages. Some of these issues are:

- *Differences in social desirability bias.* Questions could be received as sensitive in one language and not in the other, whether due to linguistic or cultural reasons.
- *Differences in level of diction.* One language version might be cognitively more complex to process than the other. For example, in the 2002 ISSP module on Family and Changing Gender Roles the word "parent" was translated as "progenitor" in Spain in the statement "One parent can bring up a child as well as two parents together". The word was rendered in Mexican Spanish as "padre", which has a stronger male connotation (and hence could be understood as "father") than the term used in Spain. "Progenitor", however, is a much less frequent word and has a higher register than the word "padre", making the question more difficult to understand.
- *Differences in naturalness of language.* Translated questions, while semantically faithful, might not sound idiomatic and be perceived as stilted (Harkness, Pennell and Schoua-Glusberg 2004). This may happen because the translator followed the structure or expressions of the original version too closely, making it sound like a translation, and therefore changing how the question as stimulus is perceived.
- *Differences in how answer scales and response options are used.* A specific culture's comfort or discomfort with selecting extreme response categories, with expressing disagreement to an unknown person, or with admitting not knowing the answer to a question can affect how questions are interpreted and answered, which may, in turn, influence data quality.

Whether or not these nuances have an impact on measurement is an empirical question. Research looking at response distributions of different question versions, both in monolingual and multilingual settings, suggests that such factors matter sometimes (e.g., Smith 1995; Villar et al 2006). Therefore, resources need to be assigned to testing translated questions and uncovering potential problems before fieldwork starts. In other words, pretesting methods need to be used for translated instruments as much as for newly developed questions. In the remainder of this section we present two possible pretesting techniques that can be used in testing translations, and the next section will focus on the use of cognitive testing for pretesting translated questions.

Focus Groups

Focus groups (Globe et al 2002) can also be used to involve members of the target population in the assessment process. There are three main advantages of using focus groups as compared to cognitive interviews: 1) they allow the researcher to talk to more people in a shorter time; 2) they permit the inclusion of a varied socio-demographic group of the target population and this provides an opportunity to

see if they can reach consensus as to interpretation of items during the group discussion; and 3) they provide a chance to see what is idiosyncratic and what is cultural in language choices and opinions. Focus groups also allow the researcher to listen to how respondents use language among themselves, which can itself illuminate issues of translation. However, focus groups are limited in the number of questions that can be discussed. More importantly, focus groups do not give us the glimpse into a respondent's cognitive process in answering questions that we can get from a cognitive interview, given that they are not an adequate forum for eliciting respondent-level narrative about the questions they just answered.

Open-Ended Pretests

Some pretesting approaches provide an opportunity to involve monolingual members of the target population and expose them to translated instrument items to examine how these work for them. In *Open-ended pretests*, for instance, the instrument is administered to a respondent without providing response choices, that is, allowing the respondent to answer the question in their own words. This approach is useful in detecting translation comprehension problems, and is a good way to elicit possible response choices particularly when researchers are not sure their own lists of choices are comprehensive. But this approach does not uncover interpretation patterns, and this may hide the fact that a translated question is being interpreted in a way different from that intended, even if respondents can easily answer.

COGNITIVELY TESTING TRANSLATIONS OF SURVEY QUESTIONS

Ideally, cognitive testing of a translated question is meant to uncover patterns of interpretation of a question or text comparable to those in the original language, and both of these should match the interpretation that the researcher intended. Testing should help establish the translated question's validity, that is, how well the question measures the concept(s) it is intended to measure. Even though most times fewer resources are devoted to pretesting survey translations than questions designed in the language of the researcher(s), researchers sometimes agree to conduct cognitive testing of translated questions even though they have not conducted such testing of the original language version. The rationale behind this approach seems to be that while they feel equipped to judge if a question works as intended in their own language, they want independent assessment of the translation quality.

Even though from a practical standpoint testing a translated version should not be very different from testing the source language version, a few particular aspects must be considered when preparing cognitive testing of translated survey texts. This section presents a review of some of the considerations that must be taken into account.

Cognitive Interviewers

The same qualities are needed in a cognitive interviewer in any language. However, it is important to remember a few additional factors. Cognitive interviewers of a translated version would ideally have experience as qualitative researchers and be native speakers of the target language or at least extremely fluent speakers who have lived immersed in the culture of the target population for some extended period of time: they should have near native speaker proficiency and “intuition”, so that they are sensitive to subtle nuances that other fluent speakers of the language might not understand. The first level of analysis of a cognitive interview is conducted by the interviewer in the course of the interview itself. Thus, a native speaker without adequate qualitative researcher training, or an experienced researcher without such strong language skills, will likely be unable to probe responses adequately to satisfy the needs of that first

level of analysis in which inconsistencies across responses or across a response and the corresponding narrative are best explored (see chapter 2, this volume).

Respondent Selection

Cognitive testing participants should be as similar as possible to the target population. An important element in translation and testing of a translation is knowing who the intended target population for the translated questionnaire is. That is, the population for the original instrument may differ in more than just language from those who will answer the translated version. For example, in the European Social Survey, where the original questionnaire is designed in English, the English version used in Great Britain with general population will differ in language and potentially in cultural aspects from the version used in French with general population in France, yet both versions need to be aimed at a general population of varied socio-economic status. In the United States, a Spanish version of a national survey questionnaire will be primarily targeting low income and low education immigrants from Latin America, predominantly Mexican. Thus, it needs to be translated taking this into account and utilizing language that can be understood by respondents with those characteristics. Other language groups might have a different profile. For example, in the 2010 U.S. Decennial Census the Spanish version needed to work well with a lower education immigrant population, while the Russian version was targeted for a more highly educated group.

To the extent possible, participants should be monolingual. Bilinguals are—by definition—different from the monolinguals for whom a translation is intended; given that culture is transmitted through language, bilinguals may exhibit different cultural traits when using different languages. As they live across two worlds, they are exposed to the culture of the source language while the target population of the translation may only be tangentially so, or may find it altogether foreign. This is particularly important in populations where some respondents might have a decent ability in the source language. Their understanding of the target language might be different from that of respondents who do not speak the source language, and they might be more tolerant (and less sensitive) to interferences across languages, such as wrong collocations (order or words in sentences) or prepositions.

Within monolingual testing researchers know that subgroups in the population—subcultures, one may call them—may tend to select some response choices over others, either because their reality warrants so, or because of their own cultural interpretation of the meaning of response categories. Therefore, when selecting participants for cognitive testing special care must be put into including members from all the relevant subgroups (see chapter 2, this volume).

Introduction, Protocol, and Implementation

Development of the cognitive testing protocol can start from the same protocol as is used for the original language version. This way the same aspects of question interpretation and processing will be considered in all language versions, and it will be possible to examine whether interpretation patterns for questions and response choices are similar across versions. Additional probes can be used as necessary to investigate specifically the interpretation of items/terms that presented particular difficulties or raised uncertainty during the translation process, or that researchers feel might pose cultural or cognitive problems in the target language.

A number of researchers have pointed to challenges faced when implementing cognitive interviewing protocols designed for the general population with immigrants in the United States (Goerman 2006,

Goerman and Caspar 2010, Pan et al 2010), suggesting that cognitive interviewing procedures need to be culturally adapted. However, the Budapest Initiative successfully implemented the same protocol and procedures across seven countries (Fitzgerald et al 2009; Miller 2008), as do the ESS Question Design Team routinely in the process of designing the source questionnaire (see, for example, Widdop et al 2011). In addition, Goerman (2006) acknowledged that: a) some of the interviewers for the target language had less experience those in the source language, which might partially explain why probes did not work with respondents who spoke the target language, and that b) native English speakers with lower levels of education also appear to have difficulties with some of the procedures involved in cognitive interviewing. If lower education is a driver of difficulties in participating in cognitive interviews, strategies may need to be followed that improve task comprehension and willingness to cooperate during the interview for the low educated. It is thus possible that to help put respondents at ease, one may need to spend more time explaining the rationale behind the procedure and what it consists of, or providing more examples of what you want respondents to do.

Analyses

As mentioned before, whether or not the source questionnaire is open to change will determine how to handle problems found in cognitive testing: cognitive testing of a translation can be used as part of the overall instrument design process or inform only the translation. In studies where the source instrument is being tested simultaneously to other language versions, all versions should be tested against the researchers' intended meaning of each question, and results from all cognitive interviews can inform the source questionnaire. If the source questionnaire has been finalized, then the translated text will also be checked against the source text.

One of the ways in which cognitive testing can help us evaluate questions (whether original or translated) is through textual assessment. In the original language, when respondents interpret a question in a way different from the intent of the question designers, this signals the presence of a specification error, where the "translation" of the intended meaning into the question wording did not work well. As a consequence, the source question wording needs a revision. In the translated version, when respondents interpret a question differently from the intended interpretation, this could signal either a problem in the original item or a problem with the translation.

Most research on using cognitive testing in multilingual settings uses these three broad categories to classify the types of problems that are found in analysis, albeit with slightly different names. Table 5.1 shows three papers that assign similar labels to the types of problems that are encountered in cognitive testing of translated questions. This classification informs the type of recommendation that will result from the problem: whether the source wording needs to be modified, the translation reviewed, or the entire concept to be measured reconsidered.

	Problems are identified as related to translation choices made	Problems are identified as related to the source question	Problems are identified as related to culture
Willis et al., 2005	Translation errors	General problems	Culturally specific issues

Budapest Initiative (Miller 2008; Fitzgerald et al 2009)	- Translation error - Interaction between source question and translation	Source question problem	Cultural issues
European Social Survey, 2013	Translation problems: - Resulting from translation error - Resulting from source question design	Poor source question design	Cultural portability

It may not be easy to disentangle in a field test of a questionnaire whether an immigrant population responds differently to a translated question about their general health status because they actually are in poorer health than those responding in the original language, because response categories have been poorly translated, or because of cultural factors that make them stay away from self-reporting very positive health status. It is cognitive testing that allows us to examine each of these possible explanations through the analysis of the narrative elicited in testing.

PROBLEMS UNCOVERED BY COGNITIVE TESTING OF TRANSLATIONS

In the next paragraphs we will present some examples of issues that have been uncovered through cognitive testing in translated instruments. Some issues are strictly related to the language version tested, others are likely to relate to problems with the original language version, and others seem to be related to cultural differences between the target populations. We will illustrate these issues with examples from real cognitive testing of questions either from the literature, or from our own experience translating survey instruments into Spanish.

Uncovering Translation Problems

Uncovering Translation Mistakes

Numerous mistakes in translation can be uncovered in cognitive testing: unidiomatic expressions, additions, omissions, wrong terms can all be revealed in the process of investigating how respondents interpret and answer the question. Example 5.1. illustrates the case of a Swiss French translation where an addition (“par rapport au revenue”) made the response options too similar and thus confusing for respondents (Fitzgerald et al 2009). The added words “in relation to income ratio” made the second response option closer in meaning to the first response option. The recommendation was to add a note for translators in the source question to remind them to check that the two response options were clearly distinct.

Example 5.1

Original Item:

Using this card please tell me which of the three statements on this card, about how much working people pay in tax, you agree with most?

1. Higher earners should pay a greater proportion in tax than lower earners
2. Everyone should pay the same proportion of their earnings in tax
3. High and low earners should pay exactly the same amount in tax

Translated Item:

Veillez lire les trois affirmations figurant sur cette carte, qui portent sur les impôts payés par les gens qui travaillent. Avec laquelle de ces affirmations êtes-vous le plus d'accord ?

1. Les hauts revenus devraient payer une plus grande proportion d'impôt que les bas revenus.
2. Tout le monde devrait payer la même proportion d'impôt par rapport au revenu.
3. Les hauts revenus et les bas revenus devraient payer exactement le même montant d'impôts.

Uncovering Regional Variations in Terms

Languages are sometimes spoken in different varieties depending on regional differences. For that reason, it has been recommended that the person who translates a questionnaire has the target language as their native language, and that they live in the same region as those who compose the target population (ISR, CCSG guidelines). For example, to produce a translation of Austrian German it would be best to hire translators and reviewers who speak that variety of German, rather than, say, Swiss German. Even in projects where the language of the source questionnaire is spoken in more than one country, adaptation to regional differences in how the language is spoken might be necessary. For example, differences are found in the American, British, and Australian questionnaires for ISSP modules.

Other projects, however, seek *harmonization* of language versions across countries (see Harkness, Pennell et al 2008 for a review of the challenges involved in this approach). If one language version will be used for all countries speaking that language, cognitive testing of the harmonized instrument in each of those countries will be crucial to uncover problematic terms that are not understood or are understood differently.

Regional variations are present also within countries, and linguistic and lexical differences across regions can be more or less pronounced. When translating instruments into Spanish for Latin American immigrants in the United States, for example, researchers encounter varied national origins; language differences across Spanish from different countries are larger than regional variations within those countries, and yet one version in the United States is expected to be used for all respondents, regardless of the variation they speak.

Cognitive testing of Spanish translations in the United States has shown to sometimes uncover the use of terms that do not mean the same across respondents from different countries (Goerman et al 2013; Schoua-Glusberg et al 2008; Sha et al 2013). Conducting cognitive interviews with individuals who speak the different variations might also be more feasible than finding translation experts from each

regional variation. Example 5.2 uses an attitudinal question focused on the subject of smoking to illustrate this.

During a cognitive interview (Schoua-Glusberg 2005), female respondents from Guatemala explained that “bares” to them evoked the idea of drinking establishments where men go to drink alone or with prostitutes, and where “decent” women simply do not go. The translated item evoked for these women an environment that was not suggested in the original question, and therefore the translation was modified to include more options so as to soften the possible effect of the inclusion of the problematic term. It was not advisable to exclude the term because for speakers of other regional variations it conveyed the intended meaning precisely. In this case, a perfectly valid translation in semantic and linguistic terms, was not appropriate from the pragmatic point of view, and would have most likely elicited a different type of response from female respondents from Guatemala as compared to female respondents from other regions where Spanish is spoken and from respondents in the source language.

Example 5.2

Original Item:

In bars and cocktail lounges, do you think smoking should be allowed in all areas, some areas or not at all?

Translated Item:

En bares o barras ¿cree que se debería permitir fumar en todas las áreas, en algunas áreas, o no debería permitirse del todo?

Modified Item:

En bares, barras, cantinas o tabernas, ¿cree que se debería permitir fumar en todas las áreas, en algunas áreas o no debería permitirse del todo?

Uncovering Lack of Familiarity with Vocabulary and Idioms

Sometimes, a perfectly translated question that semantically means the same as the source language is not understood as intended by the target population. Example 5.3 illustrates this.

A question that we tested was a Spanish translation of a question asking respondents to say to what extent they agreed or disagreed with the statement “Smoking is physically addictive” (Schoua-Glusberg 2005). After asking the question in the cognitive interview, the interviewer probed respondents to provide more information about their answer choice. In doing so, some respondents revealed that they had answered even though they were not sure what “adicción” meant. Others indicated they knew what that term meant but not when qualified with “física”. Yet others respondents thought the phrase “adicción física” meant that the body is somehow addicted to the movement of the arm to bring the cigarette back and forth to the mouth, rather than the physiological reaction as the researcher intended. While Latin Americans with higher education are typically familiar and comfortable with this phrase, the translation choice did not work adequately for the study’s sample of general population of Latino immigrants in the United States.

Example 5.3.

Original Item: Smoking is physically addictive

Translated Item: Fumar causa adicción física

Modified Item: Fumar causa adicción a la nicotina

Question not Understood as Expected by the Experts

Example 5.4. illustrates how cognitive testing can uncover interpretation issues in translated instruments. Initially “investigational drug” was translated as “droga investigacional” to keep the stimulus as complex as translators felt it would be for respondents to the English version, and under the understanding that this would be the phraseology that might be utilized in a consent form prior to mass vaccinations for community residents in neighborhood schools in case of a hypothetical outbreak of a serious and contagious disease. That term posed problems for a few of the Mexican respondents who interpreted the term ‘droga’ as referring to illegal or street drugs. The recommendation was made to change ‘droga’ to ‘medicamento’ (medication or drug—without the connotation of illegal drugs). After broader consultation with the researchers, and to make the full phrase clearer in Spanish, the term was changed to “medicamento bajo investigación”.

Example 5.4.

Original Item: If you were asked to sign a piece of paper at the school saying that the smallpox vaccine is an “investigational drug” that has not yet been completely tested, would you be not at all worried, slightly worried, moderately worried, very worried, or extremely worried?

Translated Item: Si en la escuela le pidieran que firmé un papel que diga que la vacuna de la viruela es una “droga investigacional” que todavía no ha sido completamente probada, ¿no estaría nada preocupado(a), estaría un poco preocupado(a), moderadamente preocupado(a), muy preocupado(a) o extremadamente preocupado(a)?

Modified Item: Si en la escuela le pidieran que firmé un papel que diga que la vacuna de la viruela es un “medicamento bajo investigación” que todavía no ha sido completamente probado, ¿no estaría nada preocupado(a), estaría un poco preocupado(a), moderadamente preocupado(a), muy preocupado(a) o extremadamente preocupado(a)?

Uncovering Problems with the Source Question

A review of published reports from multilingual projects where cognitive testing is carried out suggests that problems with the source question might be more common than translation errors or issues of cultural portability. We present here just one of many issues that can be discovered through cognitive testing of translations.

Uncovering Cognitively Difficult Questions

Questions that are cognitively difficult to process include those where something in the item impairs processing, such as text length, syntactic complexity, or grammatical structure. Questions in all languages should elicit a similar level of cognitive difficulty. However, translation sometimes introduces elements that make the question in the target language more complex than the source question. Other times, items that function well in the general population who speaks the source language may not work well among the population that speaks the target language because the latter have little or no formal education (such as immigrants from Latin America to the U.S.).

Example 5.5 illustrates an example where the complexity of the word did not change in translation, but where the question proved very difficult to process for some Spanish language respondents. During the cognitive interviews, several respondents asked for re-reads and visibly made efforts to comprehend the question fully, such as frowning. The recommendation after testing was thus to simplify the source question.

In fact, sheer length and packing of too much information in the item seemed to cause the trouble with this question, both in the translated and the source version. The original English item had not been cognitively tested, and it was only through testing the translation that designers became aware of the question's cognitive complexity.

Example 5.5

Original Item: How much do you think the people making plans to deal with terrorist attacks in your community know about the concerns you would have and the information you would want in these sorts of situations? Do you think that planners know a great deal about your concerns and information needs, a lot, a moderate amount, a little, or nothing at all?

Translated Item: ¿Qué tanto cree que las personas que están haciendo planes para enfrentar ataques terroristas en su comunidad saben acerca de las preocupaciones que usted tendría y la información que usted querría tener en esos tipos de situaciones? ¿Cree usted que las personas a cargo de los planes saben muchísimo, mucho, algo, un poco o nada con respecto a sus preocupaciones y a la necesidad de información que usted tendría?

Uncovering Problems Related to Cultural Differences – Need for adaptation

Uncovering Culturally Unknown Concepts

In cognitive testing it became apparent that although the wording of the question presented no interpretation problems, the question expects respondents to have some background/knowledge about the fact that part of the price of a pack of cigarettes are taxes, how taxes relate to services, and what a tobacco control program is. This was found not to be true for all the immigrants the question was tested with.

Answering the question was therefore difficult for those whose degree of acculturation/education did not allow them to connect the question with any prior understanding of how taxes work in the new context. If there is no possible translation that could fix this difference in how the question is understood, researchers need to resort to other strategies that in most cases will imply either modifying the source question or adapting the question so that it is processed and understood in similar ways in all languages. In example 5.6 (Schoua-Glusberg 2005) this might mean adding an explanation so that respondents learn about taxes in the culture. Because the source question was not open to changes and the goal of the researchers was to keep the questions as similar as possible, the researchers chose to break down the long question and set up the hypothetical situation in the first sentence. The question itself then became simpler to process.

Example 5.6

Original Item: How much additional tax on a pack of cigarettes would you be willing to support if some or all the money raised was used to support tobacco control programs?

Translated Item: Si los impuestos que se cobran en cada cajetilla de cigarrillos se usaran en parte o totalmente para contribuir a programas de control del tabaquismo, ¿qué aumento en el impuesto apoyaría usted?

Modified Item: Supongamos que los impuestos que se cobran por cada cajetilla de cigarrillos se usaran en parte o totalmente para programas de control del tabaquismo. ¿Qué aumento en los impuestos estaría usted dispuesto(a) a apoyar?

Uncovering Relevance of Concepts

Some questions travel better than others. Oftentimes adaptation is necessary for a question to be pragmatically appropriate and comparable in a culture other than the one it was designed for. As mentioned before, adaptation can also be necessary in the absence of translation, which points to the need to also conduct cognitive testing in new cultures that speak the language of a source questionnaire.

Questions that include examples most often need to be adapted so that the examples are as relevant to the target culture as the ones mentioned in the source question are to the source culture. Cognitive testing provides a great opportunity to test whether respondents from the target population interpret and process the adapted questions in a similar way to the population for whom the source language version was designed.

Willis et al (2005), for example, report testing an adaptation in a question about food consumption; they found that using “quesadilla” instead of “lasagne” as an example of food containing cheese worked adequately for the target population.

CONCLUSION

In this chapter we have presented the goals and aims of cognitively testing translated questions. Just as in cognitive testing of source questions or other survey materials, uncovering interpretation patterns and response processes is crucial to assess the validity of translated items. Furthermore, cognitive testing provides a means for assessing a translation’s performance and quality, in semantic terms as well as in terms of how it works in the culture and context for which it is intended. In designing and testing a question, a great deal of effort and attention goes into making sure that the phrasing and terms are interpreted as intended and that the item functions well in eliciting the information it seeks. Yet this same focus and care are often not exhibited in producing a version of the question in a different language (and therefore, for use in a different cultural context). Cognitive testing of every language version that will be fielded is unquestionably best practice. It is then possible to answer whether a question is being interpreted as intended in every language in which data will be collected.

The dangers of not using appropriate pretesting techniques to evaluate translated survey tools are not just the same as the dangers of not pretesting questions in monolingual studies. In addition to uncovering unexpected and undesired patterns of interpretation, cognitive testing of a translated instrument can uncover translation issues that might go unnoticed in other translation assessment processes such as expert reviews. Furthermore, it can be used to test and try out different solutions to translation challenges when the review meeting has not led to an unequivocal choice.

A naive observer may wonder if such care should go into the translation process when the questions will undergo cognitive testing anyway. They may argue that if there are question translation issues, these will be uncovered in the elicitation of interpretation patterns; for instance, if a question has a term that has been translated erroneously, the interpretation of the item will point at that error. This is a true and a valid point. In a world with unlimited resources, both time and money, cognitive testing could begin with a very rough draft of a translation. As interpretation patterns are uncovered that point at issues in the draft translation, revisions could be made to the translation and a new round of testing could follow. This process would likely take a number of cognitive testing rounds until all translation problems were uncovered and resolved. In the real research world, however, a better use of resources calls for having the best possible drafted questions before testing begins, both in the original language and in the translated version. Research is needed to investigate what each assessment step and pretesting technique can contribute towards ensuring comparability of different language versions: what kind of mistakes can each uncover, prevent, and help fix.