Computer-Assisted Migration Research: What We Can Learn About Source Questionnaire Design and Translation from the Software Localization Field
Behr, Dorothée

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Chapter 5
Computer-Assisted Migration Research: What We Can Learn About Source Questionnaire Design and Translation from the Software Localization Field

Dorothée Behr

5.1 Introduction

In the digital age, surveys are often conducted computer-assisted, which can either be administered by an interviewer or self-administered in the presence of an interviewer, or online without any interviewer present. In the following, I will use the term "computerized surveys" for all these scenarios. Among the great advantages of computerized surveys is the fact that the collected data is immediately digitized and does not require further processing. On the design side, the computerization comes with an increased interplay and interconnectivity between questionnaire and survey software (e.g., through the use of fills or placeholders, complex routing, automatic pop-up error messages, etc.), which has to be carefully considered when implementing computerized surveys.

The trend towards computerization also applies to migration research. While computerized surveys facilitate certain processes, such as reaching the target population (Pötzschke & Braun, 2017), the typically multilingual and multicultural character of migration surveys adds a layer of complexity to computerized implementation. This chapter aims to tackle the topic of questionnaire translation in migration research mainly from this technical perspective of multilingual survey implementation. General translation approaches and requirements can be found in various publications on questionnaire translation (e.g., Behr, 2018a; Behr & Shishido, 2016; Harkness et al., 2010b; Mohler et al., 2016). However, technical aspects related to translation often remain in the background, even though they can become quite crucial for data quality (Wang et al., 2017). It is only recently that challenges resulting from the interplay between translation and survey software have been brought to the fore (Pan et al., 2020; Wang et al., 2017). Although highly

D. Behr (✉)
GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany
e-mail: dorothee.behr@gesis.org

informative, these recent works do not attempt to build on and connect to related research in other disciplines, such as in translation studies. A notable exception is Upsing et al. (2011), who describe processes for large-scale competency assessment with reference to standards in software translation and adaptation – typically called "software localization" – using PIAAC as an example, the OECD Program for the International Assessment of Adult Competencies. Localization hereby refers to the processes of "adapting [digital] content linguistically, culturally, and technically" (Gambier, 2016, p. 891). It is the explicit aim of this chapter to draw on frameworks from software localization to foster knowledge transfer. I should add that the content of this chapter is applicable to multilingual computerized survey research in general, whether it is migration, cross-cultural or cross-national research.

In the following, I will provide a brief overview of good practices in questionnaire translation in general (Sect. 5.2). I will then introduce frameworks from software localization and transfer these to the survey research field (Sect. 5.3). With the help of these frameworks, I will outline major decisions in multilingual survey design and challenges that can arise when translating questionnaires for computerized surveys (Sect. 5.4, 5.6 and 5.7). Major players in multilingual research (e.g., Upsing et al. 2011; Dept et al., 2017; the consortia implementing the OECD studies PISA[1] and PIAAC) already operate according to the principles referred to in this chapter. Implementing computerized studies with a high level of comparability across diverse countries and cultures would otherwise not be possible. The author is not cognizant of computerized migration research that systematically addresses technical challenges right from the start. This can also be due, however, to a general scarcity of documentation on translation procedures and challenges in migration research.

5.2 Questionnaire Translation: What Aspects Lead to High Quality?

Migration research frequently requires the translation of a source questionnaire into various target languages in order to reach the intended respondent population. The term translation, as understood in this chapter, shall also include different levels of cultural adaptation (e.g., substantial modifications to items) to make an item work better in a new context (Behr & Shishido, 2016; van de Vijver & de Leung, 2011). Unfortunately, questionnaire translation is still too often seen as a step outside of the scientific process (Smith, 2004). Moreover, it is often regarded as a mere "word-by-word substitution, a problem of dictionaries" by those not familiar with translation (Gambier, 2016, p. 887). However, good translation significantly differs from such a view. It draws on interacting competencies, including linguistic and textual competencies in both the source and target language, as well as on cultural, substantive, information acquisition, tools, and general translation know-how (Behr, 2018a).

---

[1] PISA: Program for International Student Assessment.

When seen from this angle, it becomes clear that a simplistic view of translation leads to little consideration of translation needs, to small budgets for translation teams, and/or to unfeasible timelines. A reduced quality of the translation and ultimately of the research output is likely to result.

Much has been written in the cross-cultural survey methodology literature on good practice translation procedures. Seminal work is included in the edited volumes by Harkness et al., (2003), Harkness et al. (2010a), Johnson et al. (2018) as well as in the *Cross-Cultural Survey Guidelines* (Survey Research Center, 2016). The main guidance in cross-cultural psychology is provided by the *ITC Guidelines for Translating and Adapting Tests* (International Test Commission, 2017). The health research field is quite divers in the guidance it offers. Aquadro et al. (2008), Wild et al. (2005), Wild et al. (2009), and Eremenco et al. (2018) represent useful literature to get started in this field. Practices and challenges, specifically for migration research, are summarized in an edited volume by Behr (2018b).

The different disciplines coincide in emphasizing that a multi-step process is needed in order to ensure comparability and quality of the newly produced instrument. Using the example of the TRAPD model (Harkness, 2003), good practice calls for double translation of the questionnaire by two independently working translators (T), team reviews to arrive at a final version (R = Review and A = adjudication), pre-testing (P) among the target population, and a thorough documentation of all these steps for both internal and external quality control (D).[2] The various steps should include input from experts from different fields since a combination of expertise (in translation, questionnaire/survey design, and the substantive topic) is deemed crucial for producing a high-quality translation that fulfills both the needs of a good translation and those of a properly functioning measurement instrument (see Behr & Shishido, 2018a, b; Harkness, 2003).

Furthermore, translation teams need to be briefed on the task at hand. That is, they need to be given concrete information on the study and the translation goal (e.g., are cultural adaptations allowed or should the translation adhere to the source text?) so that they can make appropriate decisions in line with the overall objective of the study. Behr et al. (2018) also speak of "input documentation" in this context, as opposed to "output documentation." The latter includes the translated questionnaires and comments on these (e.g., in case of difficult decisions or needed cultural adaptations) and a description of the overall process implemented. At a minimum, the briefing – or "input documentation" – should include information on the study, the translation goal, the target group, the survey mode, the employed translation and assessment processes, and expectations linked to each of these steps. It can – and even should – be expanded by information on key terms, on the questionnaire structure in case of complex instruments, reference materials, etc. Information can be conveyed in written form and additionally through (web) trainings (Behr, 2018a; Behr et al., 2018; Dept et al., 2017). Input documentation intended for translators and

---

[2]Modifications of this double translation and team model are illustrated, for instance, in Dept et al. (2017), Martinez et al. (2006), and Goerman et al. (2018).

output documentation intended for research teams, alongside with open communication channels for all types of queries and issues, is particularly important in situations where research teams responsible for a study do not themselves speak the languages of a study and thus need to rely on external translators.

Another crucial cornerstone when it comes to ensuring translation quality and comparability is the source questionnaire itself. It is already during the development phase of a source questionnaire that the way is paved for comparability. Cross-national or cross-cultural research collaboration at the development stage of a questionnaire is vital to ensure that different cultural and linguistic realities are considered and sufficiently taken on board. The wording of source questions should be kept as simple as possible and allow a "relatively" easy transfer from one language to the other. Questions can also be annotated for translation or specifically earmarked for adaptation (Behr & Scholz, 2011). Furthermore, pretesting and translatability assessments – or alternatively advance translations – help to assess the questionnaire's suitability for a multilingual and multicultural study before the source questionnaire is finalized (Acquadro et al., 2018; Dept et al., 2017; Dorer, 2020; Smith, 2004). The translatability criteria summarized in Aquadro et al. (2018) or the advance translation scheme by Dorer (2011; later updated in Dorer, 2020) highlight what can be considered when reviewing a source questionnaire for translatability (e.g., issues pertaining to culture, language or item construction).

Challenges that need to be considered when implementing a computerized survey in more than one language and/or cultural group can partly be deduced from these schemes but they are not explicitly mentioned. Since the technical set-up of a computerized survey, in particular the way how it is programmed, will impact on translation and may lead to problems with the translation later on, this topic shall receive heightened attention in this chapter. Readers should bear in mind, however, that these more technical aspects always need to be considered alongside cultural, linguistic, and design issues that can impact translation and comparability.

Against this backdrop, I now want to introduce key frameworks and approaches from software localization. The ultimate goal is to transfer these to the survey research field. The software localization field has summarized the technical challenges with multilingual software that they have encountered over decades into best practice frameworks. Transferring this knowledge across disciplines – and adapting it where needed – avoids making the same mistakes again.

5.3   Software Localization: Frameworks and Approaches

With the advent of the personal computer in the 1980s, the localization industry began to develop, tasked with "adapting [digital] content linguistically, culturally, and technically" for new regional markets (Gambier, 2016, p. 891). The industry started with the localization of software and websites, and it has now moved on to also include the localization of mobile phones and video games (Gambier, 2016).

The activities related to providing such products for new linguistic and cultural markets largely exceeded the requirements linked to translation as exercised and known before. After all, extensive project management, software or graphics engineering, content management systems, etc. are all needed in this endeavor. Hence, a new term was coined: localization. Simultaneously to these developments, computer-aided translation (CAT) tools rose to prominence, which help, for instance, in the separation of programming code on the one hand and content for translation on the other hand, and in consistent use of reoccurring text elements. CAT tools are nowadays an integral part of the translation environment of professional translators (Sin-wai, 2016).

In a more fine-grained and process-driven view, the localization industry operates according to the GILT framework, which subsumes the four interdependent activities globalization, internationalization, localization, and translation. *Globalization* stands for all activities related to marketing a product in various regional markets. *Internationalization* stands for preparing a product at the technical level for localization. As such, it is the "process of generalizing a product so that it can handle multiple languages and cultural conventions without the need for redesign" (LISA cited by Esselink, 2000, p. 2). *Localization* stands for the process of modifying a product for a specific market. *Translation* is in fact already part of localization, because localization includes both adaptation and translation. The better a product is internationalized, the more cost- and time-effective localization can be carried out (Gambier, 2016; Sandrini, 2008; Valli, 2019). Overall, the GILT framework highlights that technology and its requirements are one decisive pillar besides language and culture.

The internationalization and localization processes can further be broken down into five core elements (Schäler, 2010): (1) *Analysis* refers to a set of key questions that need to be asked prior to localization, for instance, whether it makes sense to localize the content at all, whether all the text that needs to be translated is accessible for translators, or whether it is hidden in program code that cannot be modified? (2) *Preparation* refers to preparing a so-called localization kit for everyone working on the project, including source materials, reference materials, guidelines, milestones, etc. (3) *Translation* takes place in a highly computerized environment, which is nowadays standard for many translators. Translators work with CAT tools that include translation memories[3] (TMs), terminology databases, machine translation (MT) functionalities, automatic checking routines, etc. Sometimes, preview functions allow viewing the translated text in the actual software environment, which is even more important when text strings have to be translated out of context. (4) *Engineering and testing* involves assessing the content in terms of linguistic correctness, interface layout, and functionality. (5) A *review* closes the localization project. Lessons learned are collected for future projects.

---

[3]Translation memories display previously translated (similar) segments and thus both speed up the translation process and ensure consistency.

Table 5.1 Merging frameworks and approaches across survey methodology and localization

Survey methodology	GILT framework	Core elements of a localization process	Covered in this chapter
Design and implementation of the source questionnaire	Internalization	Analysis	Sect. 5.4
		Preparation	Sect. 5.5
Translation, incl. Adaptation	Localization (translation)	Translation	Sect. 5.6
Technical pretesting		Engineering and testing	Sect. 5.7
Substantive pre-testing. (Please note: For reasons of completeness, substantive pretesting that examines wording, content, and comparability (e.g., through cognitive interviewing and/or pilot testing) is listed here. It should complement the process of questionnaire translation since it evaluates to what extent the questions indeed measure what they are supposed to measure. However, due to the technical focus of this chapter, substantive pretesting is not further covered. More information on different forms of pretesting can be found in: Goerman and Caspar (2010), Pan et al. (2010) or Willis (2005, 2015, 2016))	n/a	n/a	n/a

In the following, these frameworks and approaches from the localization field are merged to outline steps that should be taken when preparing for and implementing translations in multilingual and multicultural computerized migration surveys (see Table 5.1).[4]

The chapter will focus on technical issues that can affect or interact with translation. For more comprehensive guidelines on technical questionnaire design, Hansen et al. (2016) should be consulted. In the subsequent descriptions, we will assume that survey programming is implemented centrally and that a source questionnaire, as programmed, can serve as a template for the intended target languages. I will not go into details of programming but rather point to general issues to consider.

---

[4]Please note: For reasons of completeness, substantive pretesting that examines wording, content, and comparability (e.g., through cognitive interviewing and/or pilot testing) is listed here. It should complement the process of questionnaire translation since it evaluates to what extent the questions indeed measure what they are supposed to measure. However, due to the technical focus of this chapter, substantive pretesting is not further covered. More information on different forms of pretesting can be found in: Goerman and Caspar (2010), Pan et al. (2010) or Willis (2005, 2015, 2016).

5.4 Design and Implementation of the Source Questionnaire

This section is dedicated to the design phase of a survey questionnaire. It highlights 'internationalization' decisions that need to be made prior to and during technical questionnaire design. In line with best practice in multilingual and multicultural research, namely that the design process is decisive for translation quality later on (Behr & Zabal, 2019), this preparatory process should receive special attention in a survey.

5.4.1 Software Fit

Migration research is oftentimes multilingual. The languages chosen depend on the target population. The German IAB-BAMF-SOEP Survey of Refugees, for instance, is implemented in German and (alongside it in a bilingual fashion) in English, Arabic, Farsi, Pashto, Urdu, and Kurmanji (Jacobsen, 2018). In such multilingual surveys, the first necessary technical clarification refers to the survey software. It needs to support the required scripts and character sets, language directions (i.e., left-to-right, right-to left, vertical, bi-directional, see Hansen et al., 2016), fonts, etc. that are needed for a specific survey. This means that the decision about the required target language(s) should precede any decisions about survey software, or at least it should go hand in hand with software decisions. The right-to-left implementation needed for Arabic[5] is particularly important nowadays given that major refugee studies around the world field their surveys in Arabic, amongst other languages (e.g., Jacobsen, 2018; AIFS, 2018). Sometimes, audio-assisted self-interviewing (ACASI) is meant to compensate for the lack of multilingual interviewers – this should then also be supported by the software (see the above-mentioned studies for examples).

In the following, we will look into more specific aspects that need to be considered when designing – and translating – multilingual and multicultural computerized surveys.

---

[5] For ease of communication, we simply refer to "Arabic" in these lines. Readers should be aware, however, that there are many different ways to translate a questionnaire into Arabic, taking into account different regional dialects and the difference between written and colloquial Arabic. The target group should be the decisive factor when determining which type(s) of Arabic should be used for a given survey project.

5.4.2   Culture-Driven Response Formats with Technical Implications

Some survey features are culture-dependent. Hence, the survey should be designed in such a way that it allows for cultural adjustments. These adjustments could affect the following (Hansen et al., 2016; Valli, 2019, Maroto & De Bortoli, 2001; Pym, 2011):

- *Date formats:* e.g., different positions regarding day, month, and year such as mm/dd/yyyy in the US and dd/mm/yyyy in many European countries;
- *Time formats:* e.g., 12-hour vs. 24-hour clock;
- *Name formats:* e.g., two surnames in Spanish-speaking countries;
- *Address formats:* e.g., different sequence of information or type of information required (state, province, etc.);
- *Telephone number formats*: e.g., including or excluding local prefixes;
- *Number formats*: e.g., different decimal, thousand, etc. separators, such as 20,5 in German (DE) vs. 20.5 in English (US);
- *Currency formats:* e.g., currency symbol after or before the relevant currency entry;
- *Measurement units*: e.g., metric vs. imperial units for distances, Celsius vs. Fahrenheit, different clothing units, etc.

These formats are not only relevant when it comes to programming individual questions but also when information from these questions is automatically inserted in follow-up questions in a survey (*fills*). The way how fills are programmed needs to ensure cultural particularities such as addressing a person with the appropriate order of names (Wang et al., 2017) or "piping in" the date in the culturally appropriate way (see Sect. 5.4.4 for more information on fills.)

These or similar formats also play a role when defining out-of-scope answers that automatically trigger error messages popping up on the screen. For instance, the allowable ranges for feet or meter, when it comes to size, will be different across cultural groups who use metric vs. imperial units (e.g. km vs mile); or the need for inclusion or exclusion of commas or full stops will vary depending on language. Wang et al. (2017) share their experiences from the Chinese 2016 Census Test Internet Instrument: Respondents had to set up security questions. A valid answer had to contain at least three characters. In Chinese, the names of people, locations or items often consist of only two characters. Hence, respondents entering those names were confused with automatic validation checks, which were appropriate for the English source language but not for Chinese.

Moreover, if pre-coded response lists are provided, these will need to be adapted to the respective needs. For instance, pre-coded response lists of time will look different in different languages and cultures.

Valli (2019), speaking for software localization, argues that during internationalization cultural assumptions should be removed from software design. In particular, software should not include *hard-coded* culture-specific formats (e.g., date

formats) that cannot be changed. Hard-coded stands for text that is directly part of a source code and typically not accessible for translators. Transferred to the survey world, we can say that anything that could require a cultural accommodation should be soft-coded and/or be made editable in some other way.

5.4.3 Non-linguistic Adaptations with Technical Implications

Graphics, icons or images can be embedded in a survey for different reasons. They may serve a measurement purpose when they are an integral part of questions. Additionally, they could serve to represent the survey sponsor, survey agency or the study itself, for instance in the header of an online survey. For measurement-related graphics, icons or images, their cultural suitability for the target population should be assessed to ensure that cultural norms are not transgressed. For instance, Hansen et al. (2016) show, using the 2007 International Social Survey Program (ISSP), that body shapes can be presented with figures wearing only boxer shorts or bikinis in Austria whereas in the Philippines they wear clothes covering larger parts of their body. In technical terms, such images should be soft-coded so that they can be replaced if needed.

Similarly for icons representing a sponsor, agency or the study itself: For instance, when conducting multilingual web surveys within the contexts of cross-cultural web probing studies (Behr et al., 2019), the survey icon representing our institute, *GESIS – Leibniz Institute for the Social Sciences,* was uploaded and integrated into the survey with the German institute name for the study conducted in Germany and with the English institute name for all other surveys.[6] At a minimum, graphics, icons or images should contain editable text in case translation teams need to implement a change (Valli, 2019).

Also, links to external websites (e.g., on the survey introduction page linking to further information) may need to be replaced so that they directly link to a website in the respective target language. Similarly, Sha et al. (2018) describe how entry pages (websites) to a multilingual survey should best be designed and also adapted in order to ensure participation across multilingual groups in a society. In their case study, the authors were interested in limited English speakers' entry to U.S. Federal Government internet surveys.

In technical terms, all of the information referred to in this section should be accessible for translation teams so that the content can be adapted, if needed.

I should add here that also colors (e.g., background colors of a survey or of a logo) should be thoroughly checked in terms of cultural meaning and associations (Hansen et al., 2016).

---

[6]We started from one common source questionnaire that needed to be translated and adapted; the survey software itself (EFS) catered for multilingual implementation in the sense that once a source instrument was programmed it could easily be replicated in different languages.

5.4.4 Linguistic Differences with Technical Implications

Computerized surveys have certain features that make them unique compared to paper-and-pencil surveys. One of these features is the possibility to use responses given earlier in the survey to adapt survey text in later questions (*fills*). For instance, questions can be tailored to refer to a previously mentioned male or female partner; or questions can be asked in present or past tense depending on whether a situation currently applies or whether it applied in the past. With survey software taking on such adjustments, interviewers in interviewer-administered surveys can focus on the interviewing task itself and do not have to adapt text to a given respondent (Latour et al., 2013). In self-administered surveys, the respondents can focus on relevant text for their situation without being distracted by irrelevant information.

The multi-country *Survey of Health, Aging, and Retirement in Europe* (SHARE) (Das et al., 2005) used fills, such as the automatic insertion of 'he' or 'she' depending on the gender of a partner as indicated earlier. The multilingual implementation of the English source version proved challenging, however:

> At first sight this seemed to be straightforward, but because of country specific [sic!] grammar and syntax it became complicated. In later versions of the CAPI instrument generic fill texts used in multiple question texts were no longer used. Instead, each question had its own fills, using question-specific fill names. (p. 17)

Also in the Programme for the *International Assessment of Adult Competencies* (PIAAC), fills – or dynamic text, as they called it – was used to accommodate different respondent situations.[7] In this study, too, the researchers experienced challenges across languages. For instance, for the source question "In your ^JobLastjob, how often ^DoDid you usually . . . read directions or instructions?" it is sufficient to insert the words "current job" and "do" for the text indicated through ^ and the result is a perfect sentence in present tense in the English language. When "last job" and "did" are inserted, the resulting sentence successfully captures past respondent activities. However, in many languages, a close translation of the question, including a literal translation of fills, did not work, because past and present tense are not formed in the same manner as in the English language. Oftentimes, other solutions, including translating the entire question (or larger parts thereof) for all respondent conditions, had to be resorted to (Latour et al., 2013).

Fills are not only difficult for these linguistic-technical reasons, however. It can also be difficult to convey to translators unfamiliar with questionnaires how they are supposed to understand and translate these fills. On the other hand, if questions are completely written out for different respondent conditions, special attention needs to be directed to the briefing of translators so that they understand the difference between similarly worded sentences, their respective role in the survey, and consistency needs. For instance, in a GESIS study with tuberculosis patients from Somalia

---

[7]The source (background) questionnaire of PIAAC Cycle 1 can be downloaded here: http://www.oecd.org/skills/piaac/data/ (accessed 4 July 2021).

and Ethiopia, many sentences were partially replicated. One sentence, for instance, asked: 'Did the doctor voice suspicion that you may have tuberculosis?' The subsequent sentence read: 'Did one of these doctors voice suspicions that you may have tuberculosis?' The second question applied to situations where the respondent had several doctors taking care of him/her. The translations were supposed to be identical, except for the difference between 'the doctor' vs. 'one of the doctors' (and any language adjustments needed in Somali and Tigrinya because of this difference). The translators did not always translate sentences such as these consistently, which may have been due to the fact that the set-up of the questionnaire (e.g., who gets which question) was difficult to understand and general survey principles (e.g., standardization in surveys) not known. CAT tools with translation memories would have helped the translators to translate consistently in any case.

The software localization field, too, knows the challenges that come with fills, in particular if these are based on a rather simple source language, at least on the structural level. English, for instance, has a simple morphology, with word endings that do not undergo many changes from one sentence to another. This characteristic does not necessarily apply to other languages (Valli, 2019). De la Cova (2016) observes that English word order and lack of gender may not replicate well in other languages. Valli (2019) recommends that the number and nature of fills should be well considered in advance. The same applies to surveys. Moreover, those knowledgeable of translation and the different linguistic needs of the survey languages should have a say in the set-up of fills in a source questionnaire so that problems can be prevented. "Writing for translation" (De la Cova, 2016, p. 253) or even programming for translation could be the main message here. If ease of programming in the source language prevails, the resulting translations may be suboptimal, possibly even artificial, with detrimental effects on the validity and comparability of data.

Another known challenge with technical implication is that of text expansion. Compared to English, other languages are often longer (Dept et al., 2017). Microsoft (2018) states that text strings, when translated into German or Dutch, often expand by 40% (2018; see also Valli, 2019). This needs to be taken into account when designing buttons, menus or dialogue boxes in software. Transferred to the survey context, developers should ensure that buttons (e.g., 'forward' or 'backward') are sufficiently large to cater for different languages, or that pop-up windows contain all relevant text – without incorrect hyphenation or truncation. If the survey software offers default sizes for certain elements, the various testing scenarios should establish whether this is sufficient (see Sect. 5.7). The needed or required text length also plays a role when designing open-ended text boxes where the size of the text box is known to influence the response length (Dillman et al., 2014). Thus, these text boxes should fit the expected or desired response length – and they might possibly even be enlarged in general to cater to the response length in different languages (see also Meitinger et al., 2019, on response patterns for open-ended questions in different languages).

Directly related to open-ended text boxes, respondents should be able to type in characters from different scripts into open-ended text boxes. Thus, there should be no system restriction on the type of data that can be entered into these boxes. This

challenge is exacerbated when migrants are supposed to enter text into open-ended text boxes in self-administered surveys where laptops, tablets, etc. are handed out to respondents by interviewers. It needs to be ensured that the software and the keyboard support text entries in different scripts.

5.4.5   Content Differences with Technical Implications

Often, the default situation in multilingual computerized surveys is that a generic source questionnaire serves as a blueprint for all other language versions. Additional design solutions, however, should be possible to allow a cultural group to adapt content, such as adding relevant questions or response categories or changing routing based on culture-specific needs. How this can be achieved depends on the survey software and overall design decisions. However, in a purposefully designed comparative study, adaptations to the source questionnaire should all be signed off from a central organization and documented to ensure comparability. Asking about highest educational attainment based on country-specific response categories may serve as an example of an adaptation.

5.4.6   Preparing Source Questionnaires for Computer-Aided Translation

Sometimes, the output format of the survey software requires the use of dedicated translation tools (CAT tools) to read the file, but also the normal text processing formats Word or Excel are supported by CAT tools. CAT tools can only show their strength if the source text – here: the source questionnaire – is optimally prepared. This includes avoidance of manual hyphenation, of manual hard returns or of multiple blank spaces (instead of tabs); also key terminology or repetitive elements should be worded and spelled consistently. These simple style and formatting rules allow translation memories to correctly display identical or similar translations as stored in the translation memory, or term databases to reliably show pre-defined terminologies (Esselink, 2000; Valli, 2019).

5.4.7   Internationalization Testing

The localization industry calls for internationalization testing before a software product can be localized and passed on to the next step in the workflow (Esselink, 2000). This includes checking whether the software is ready for localizability. Essentially, the testing involves the issues and challenges addressed above. Key

questions to ask are: Does the software support all needed characters and scripts? Does it support different regional (date, time, etc.) formats? Does it allow for text expansion? Does it run on required operating systems? Is all text that needs to be translated or adapted accessible – likewise for icons, images or graphics?

To help internationalization testing, so-called pseudo translation (e.g., replacing text with more characters or with accented characters) can be carried out in an easy, low-cost way to identify issues in other languages, such as spacing issues, truncated text or issues with scripts. Based on this exercise and its analysis, recommendations can be made on how to proceed (Esselink, 2000; Lerum et al., 2014; Schäler, 2010).

These testing procedures from localization can equally be implemented for survey translation. Moreover, I want to stress that a source questionnaire itself should have been thoroughly tested in terms of wording, routing, and overall design, before proceeding to internationalization testing and then translation. Implementing changes on the source version – and consequently in all language versions – once the translation has started are cost-, time-, and work-intensive and there is the risk that not all source improvements are consistently implemented in all language versions.

5.5   Prior to Translation: Preparing Translation Teams

Once the source material is ready, translation can begin. In the localization industry, translators receive a so-called localization kit that does not only contain the source material to be translated but also reference materials, including translation memories, terminology databases, style guides, milestones, etc. (Esselink, 2000; Schäler, 2010; Valli, 2019). The importance of additional project information has already been discussed in Sect. 5.2, under the notion of briefing. For questionnaires, especially if the normal flow of text is interrupted through fills, or if text strings (e.g., for buttons or error messages) are not understandable without context information, annotations for translators will be helpful.

5.6   Translation, Including Adaptation

Having received the localization kit, translators in the localization industry start translating. Their task will always involve the use of specialized localization software. For survey translation, depending on output formats of source questionnaires from the survey software, translation files could be XLIFF files, which require the use of dedicated CAT tools for the translation, or Excel files, which can be translated with CAT tools but also with normal text processing programs. Translation may also take place within the survey software itself in specifically designed language editors. CAT tools can aid translation by supporting consistency of terms or of reoccurring text elements through the use of term databases and translation memories, or by

allowing the use of several automated checking routines (e.g., on spelling, punctuation, figures or formatting).

A number of translation decisions must be taken in view of the survey interface as well as with respondent activities in mind (see also Pan et al., 2020). If interviewer or respondent instructions refer to buttons on the screen (e.g., to the 'Help' button or the 'Next' button), the same translation of key terms should be used for the buttons themselves so that ease of navigation on the screen is ensured. This essentially means that the translation of the survey interface and the translation of the questionnaire should be coordinated in one way or the other.

Interviewer or respondent instructions such as 'Mark all that apply' or 'Tick only one box' should be translated with the ultimate layout and the concrete interviewer or respondent activity on the screen in mind. The translation of 'mark', 'tick' or 'box' could vary depending on these features or activities.

For questions that are asked in an open-ended fashion (e.g., the number of hours that a respondent spends on a given activity), it is important to consider the design of the survey and the position of the open-ended text box. Depending on whether it comes before the unit (here: 'hours') or after, the translation may need to be linguistically adapted to this position.

Words or phrases helping to structure a questionnaire or interview (such as: 'In the *following* . . .' or 'To what extent do you agree or disagree with the *following* statements?') always need to be translated in view of the visual survey design. The "following" could be translated in the sense of 'as follows below' if, and only if, the respondents themselves see the questions below. Otherwise, 'following' will need to be translated in a temporal sense.

If fills are used in a questionnaire, translators should be trained on how to understand these features and what they need to consider during translation. Wang et al. (2017) provide an example based on Chinese where strict adherence to the English fill structure resulted in a defective text in Chinese. Possibly, the translator (s) was not sufficiently informed on the use of fills. If fills do not work in a target language, this should be openly communicated to research teams.

5.7   Technical Pretesting

In the localization industry, testing is an integral part of software localization. It can start once the software is compiled in the target language. Testing is always based on the real application. The localization field differentiates between (a) linguistic testing, (b) interface testing,[8] and (c) functionality testing (Esselink, 2000).

The linguistic test targets all aspects related to language. Key questions that also apply to survey translation are: Has all text been translated, including error messages? Do the different scripts display correctly? Does the text hyphenate correctly?

---

[8]This is called 'cosmetic' testing in Esselink (2000).

Do fills display correctly (e.g., do they appear in the correct position in the sentence, or is capitalization or a small letter of a fill appropriate at the given position in the sentence)? Is all text translated in the intended sense, including interface elements, button labels, etc. (Pan et al., 2020)?

The interface test focuses on visual aspects. Questions that should be addressed here are: Is the text in dialog boxes or error messages displayed completely, i.e. without truncation? Are dialog boxes or error messages adequately (re)sized? Does the text fit on the screen in different resolutions? Is the localized version aesthetically acceptable? Do drop box designs display all response options? Are the different format conventions (e.g., date or time formats) correctly implemented?

Eventually, the actual functionality of the software is focused on. Esselink (2000) – speaking for the software localization sector – holds that functionality testing usually mirrors the processes that have been implemented on the source product. Moreover, the more thoroughly the source product has been prepared and tested prior to localization, the fewer problems will be found during testing of the localized product. Transferred to the survey context, the key question to ask during this final testing is whether the entire questionnaire works as intended, or whether problems were introduced through translation. For survey translation, functionality testing should involve the use of various testing scenarios that cover respondent groups receiving different parts of the questionnaire. Such a full-blown test can also identify whether the translation works in the context of more extensive routing and in different types of "paths". During the actual translation, questions are translated in a linear fashion one after the other. In a concrete survey context, however, this linear fashion may not be applicable since routing based on a given answer may send a respondent to questions much later in the questionnaire. Hence, this real-life testing is extremely important for ensuring that the questionnaire is intelligible in the different "paths" that a respondent could take through the survey.

In the localization industry, localized software often undergoes compatibility testing that checks how compatible a new product is with other products that are available in the target language (e.g., platforms, devices, web browsers). This type of testing can be crucial for translated surveys as well. For instance, Dillman et al. (2014, p. 345) discovered in one of their web studies that toggling back and forth between an English and Spanish questionnaire only worked in some versions of browsers, but not in others.

Last but not least, a remark seems appropriate on how this technical pretesting step relates to usability testing, which has gained momentum in survey research due to the rise of computer-assisted surveys. The main goals of usability testing are to improve data quality by the reduction of errors, and to prevent item or unit non-response by the reduction of respondent burden. Usability testing should not be confused with best practice questionnaire design; rather it should build upon best practice design and provide the ultimate test that ensures that interviewers and respondents can record answers easily and accurately (Geisen & Romano Bergstrom, 2017). Transferred to multilingual surveys, usability testing should first and foremost target the questionnaire in general, and as such it should be part of the aforementioned testing of the source questionnaire. However, surveys that undergo

larger cultural adjustments (e.g., a change in the writing direction) may be in need of additional usability testing once the translated version is available. This testing may still be similar to the interface and functionality testing described above.

5.8   Discussion and Recommendations

The aforementioned observations have hopefully shed some light on the complexities that can come with implementing a multilingual and multicultural survey questionnaire in a computerized survey environment. The good news is that the complexities can be mastered. What is important, however, is that technical aspects are considered early on in the development process in a multilingual survey. Just as early cross-cultural collaboration, translatability assessments or advance translations are important to ensure ease of translation and cultural relevance, so are early checks in ensuring that the software and programming work in a multilingual context. That is: When developing and programming a source questionnaire for a multilingual study, the diversity of study languages and their respective needs should always be considered. Ultimately, a well-designed survey on the linguistic, cultural as well as technical level is the pre-condition for sound data.

We require respondents to invest time and effort into replying to our questions, even though their (intrinsic) motivation may be low. We should, on our side, invest time and effort into providing questionnaires that are linguistically and culturally appropriate and function as intended. The best possible way to achieve this is early cooperation between survey developers, linguists, and translation technologists (see Lupsa, 2018, cited in Behr & Zabal, 2019). Prior to finalizing the design in the source language, issues or challenges for other languages can thus be identified and remedies found. In the same vein, Valli (2019) states for the field of software localization: "At minimum, the localization team should be *involved* in those product development phases in order to raise awareness about the future linguistic pitfalls." Checklists as the one provided in the appendix – or the work by Esselink (2000) or Microsoft (2018) – may additionally help to inform multilingual technical survey design. Furthermore, planning and workflows will need to cater to these additional layers of cooperation.

Finally, testing of translations in the computerized survey environment will need to be factored in both time- and budget-wise, and so does another loop of adjustments following the feedback from a first round of multilingual linguistic, interface, and functionality testing. In particular in languages that a research team does not speak (e.g., languages of refugees in a country) additional resources are required for this testing.[9]

---

[9] Some general information on scheduling and budgeting translation activities can be found here: https://ccsg.isr.umich.edu/chapters/translation/ (accessed 4 July 2021).

To conclude, the translation approach is rarely described (in detail) in multilingual migration research. It would be helpful if researchers document their procedures and include challenges encountered or lessons learned both in translation and in technical survey implementation. This way, surveys can raise awareness on challenges, they can learn from each other, and build on each other's experiences. We already see this transfer of lessons learned with other types of study challenges in migration research, for instance on sampling, field work or item understanding (Formea et al., 2014; Haug et al., 2019; Röder, 2018).

Appendix: Checklist – With a Focus on the Interplay Between Translation and Survey Technology

In General
- Specify the language(s) for a study and check whether it/they can be implemented in the chosen or planned survey software (see Sect. 5.4.1).
  - This includes checking to what extent text in different scripts can be entered, particularly in bilingual response situations (see Sect. 5.4.4).
- Check if multilingual recording can be added, if desired.

During Questionnaire Development and Implementation
- Identify textual elements that may require cultural adjustments (e.g., response formats in terms of date, time, name, address, telephone number or decimal separators) and decide on ways how this can be handled in the software (see Sect. 5.4.2).
- Identify non-textual elements that may require cultural adjustments (e.g., icons, images, links or colors) and decide on ways how this can be handled in the software (see Sect. 5.4.3).
- Identify fills in the questionnaire (if any are planned) and check to what extent these can be replicated in other languages (see Sect. 5.4.4).
- Consider text expansion in translation compared to the source questionnaire and check to what extent buttons, dialog boxes, text boxes, etc. cater to this (see Sect. 5.4.4).
- Plan for ways to allow cultural groups to adapt content, e.g. to suppress an irrelevant response category or to add a question (see Sect. 5.4.5).
- Prepare the source questionnaire in line with format and style requirements that pave the way for efficient use of computer-aided translation (CAT) tools (see Sect. 5.4.6).
- Once implemented, test the source questionnaire thoroughly.

Prior to Translation
- Provide instructions and further pertinent information to translation teams (see Sect. 5.5).

  – This should include information on particular features of the interface or technical questionnaire design (e.g., on the use of fills) (see Sect. 5.6).

Translation, Including Adaptation
- Consider having translators translate with CAT tools to render the translation process more efficient and less error-prone (see Sect. 5.6).
- Translate with the interface and the actual survey implementation in mind.
- Conduct linguistic, interface, and functionality testing of the translated versions (see Sect. 5.7).

References

Acquadro, C., Conway, K., Hareendran, A., Aaronson, N., & European Regulatory Issues and Quality of Life Assessment (ERIQA) Group. (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health, 11*(3), 509–521. https://doi.org/10.1111/j.1524-4733.2007.00292.x

Acquadro, C., Patrick, D. L., Eremenco, S., Martin, M. L., Kuliś, D., Correia, H., Conway, K., & International Society for Quality of Life Research. (2018). Emerging good practices for translatability assessment (TA) of patient-reported outcome (PRO) measures. *Journal of Patient-Reported Outcomes, 2*(8), 1–11. https://doi.org/10.1186/s41687-018-0035-8

AIFS – Australian Government, Department of Social Services, & Australian Institute of Family Studies. (2018). Building a new life in Australia: The longitudinal study of humanitarian migrants. *Data Users Guide: Release 4.0.* https://aifs.gov.au/sites/default/files/bnla_data_users_guide_release_4.pdf. Accessed 6 October 2019. Accessed 18 Oct 2019

Behr, D. (2018a). Translating questionnaires for cross-national surveys: A description of a genre and its particularities based on the ISO 17100 categorization of translator competences. *Translation & Interpreting, 10*(2), 5–20. https://doi.org/10.12807/ti.110202.2018.a02

Behr, D. (Ed.) (2018b). Surveying the migrant population: Consideration of cultural and linguistic issues. GESIS-Schriftenreihe 19. http://nbn-resolving.de/urn:nbn:de:0168-ssoar-58074-2

Behr, D., & Scholz, E. (2011). Questionnaire translation in cross-national survey research: On the types and value of annotations. *mda: Methoden, Daten, Analysen, 5*(2), 157–179.

Behr, D., & Shishido, K. (2016). The translation of measurement instruments for cross-cultural surveys. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 269–287). Sage.

Behr, D., & Zabal, A. (2019). A meeting report: OECD-GESIS seminar on translating and adapting instruments in large-scale assessments (2018). *Measurement Instruments for the Social Sciences, 1*, 10. https://doi.org/10.1186/s42409-019-0011-y

Behr, D., Dept, S., & Krajĉeva, E. (2018). Documenting the survey translation and monitoring process. In T. P. Johnson, B.-E. Pennell, I. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional and multicultural contexts (3MC)* (pp. 341–356). John Wiley & Sons.

Behr, D., Meitinger, K., Braun, K., & Kaczmirek, L. (2019). Cross-national web probing: An overview of its methodology and its use in cross-national studies. In P. C. Beatty, D. Collins, L. Kaye, J. L. Padilla, G. Willis, & A. Wilmot (Eds.), *Advances in questionnaire design, development, evaluation and testing* (pp. 521–544). Wiley.

Das, M., Vis, C., & Weerman, B. (2005). Developing the survey instruments for SHARE. In A. Börsch-Supan & H. Jürges (Eds.), *The survey of health, aging, and retirement in Europe–methodology*. MEA. http://www.share-project.org/uploads/tx_sharepublications/SHARE_BOOK_METHODOLOGY_Wave1.pdf. Accessed 6 Oct 2019

De la Cova, E. (2016). Translation challenges in the localization of web applications. *Sendebar: Revista de la Facultad de Traducción e Interpretación, 27*, 235–266.

Dept, S., Ferrari, A., & Halleux, B. (2017). Translation and cultural appropriateness of survey material in large-scale assessments. In P. Lietz, J. C. Cresswell, K. F. Rust, & R. J. Adams (Eds.), *Implementation of large-scale education assessments* (pp. 168–192). John Wiley & Sons.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.

Dorer, B. (2011). *Advance translation in the 5th round of the European Social urvey (ESS). FORS working paper series 2011-4*. FORS. https://doi.org/10.24440/FWP-2011-00004

Dorer, B. (2020). *Advance translation as a means of improving source questionnaire translatability?: Findings from a think-aloud study for French and German*. Frank & Timme.

Eremenco, S., Pease, S., Mann, S., & Berry, P. (2018). Patient-reported outcome (PRO) consortium translation process: Consensus development of updated best practices. *Journal of Patient-Reported Outcomes, 2*(12), 1–12. https://doi.org/10.1186/s41687-018-0037-6

Esselink, B. (2000). *A practical guide to localization*. John Benjamins Publishing.

Formea, C. M., Mohamed, A. A., Hassan, A., Osman, A., Weis, J. A., Sia, I. G., & Wieland, M. L. (2014). Lessons learned: Cultural and linguistic enhancement of surveys through community-based participatory research. *Progress in Community Health Partnerships: Research, Education, and Action, 8*(3), 331–336.

Gambier, Y. (2016). Rapid and radical changes in translation and translation studies. *International Journal of Communication, 10*, 887–906.

Geisen, E., & Bergstrom, J. R. (2017). *Usability testing for survey research*. Morgan Kaufmann.

Goerman, P. L., & Caspar, R. A. (2010). Managing the cognitive pretesting of multilingual survey instruments: A case study of pretesting of the US Census Bureau bilingual Spanish/English questionnaire. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, B.-E. Pennell, & T. W. Smith (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 75–90). Wiley.

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 cultural and linguistic issues* (pp. 29–41). GESIS. http://nbn-resolving.de/urn:nbn:de:0168-ssoar-58074-2

Hansen, S. E., Jung Lee, H., Lin, Y., & McMillan, A. (2016). Instrument technical design. In *Guidelines for best practice in cross-cultural surveys*. Survey Research Center, Institute for Social Research, University of Michigan. https://ccsg.isr.umich.edu/chapters/instrument-technical-design/. Accessed 4 July 2021

Harkness, J. (2003). Questionnaire translation. In J. Harkness, F. J. R. van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 35–56). Wiley.

Harkness, J. A., van de Vijver, F. J. R., & Mohler, P. P. (Eds.). (2003). *Cross-cultural survey methods*. Wiley.

Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P. P., et al. (Eds.). (2010a). *Survey methods in multinational, multiregional, and multicultural contexts*. Wiley.

Harkness, J. A., Villar, A., & Edwards, B. (2010b). Translation, adaptation, and design. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, et al. (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 117–140). Wiley.

Haug, S., Lochner, S., & Huber, D. (2019). Methodological aspects of a quantitative and qualitative survey of asylum seekers in Germany – A field report. *Methods, Data, Analyses, 13*(2), 321–340.

International Test Commission. (2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). https://www.intestcom.org/page/14. Accessed 4 July 2021.

Jacobsen, J. (2018). Language barriers during the fieldwork of the IAB-BAMF-SOEP survey of refugees in Germany. In D. Behr (Ed.), *Surveying the migrant population: Consideration of cultural and linguistic issues* (pp. 75–84). GESIS. http://nbn-resolving.de/urn:nbn:de:0168-ssoar-58074-2

Johnson, T. P., Pennell, B. E., Stoop, I. A., & Dorer, B. (Eds.). (2018). *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts (3MC)*. John Wiley & Sons.

Latour, T., Jadoul, R., & Wagner, M. (2013). *Development of the CAPI questionnaire system*. Technical report of the Survey for Adult Skills (PIAAC). https://www.oecd.org/skills/piaac/_Technical%20Report_17OCT13.pdf. Accessed 18 Oct 2019

Lerum, C. B., Nelson, J. A., de Matos Capistrano, A. Integrated application localization. U.S. Patent No. 8,789,015. 22 Jul. 2014. https://patentimages.storage.googleapis.com/cd/4d/51/bb3e2a7a524a78/US8789015.pdf. Accessed 18 Oct 2019.

Maroto, J., & De Bortoli, M. (2001). *Web site localization*. http://pure.au.dk/portal/files/11487/Appendix_3.pdf. Accessed 18 Oct 2019.

Martinez, M., Marin, V., & Schoua-Glusberg, A. (2006). Translating from English to Spanish: The 2002 National Survey of Family Growth. *Hispanic Journal of Behavioral Sciences, 28*(4), 531–545.

Meitinger, K., Behr, D., & Braun, M. (2019). Using apples and oranges to judge quality? Selection of appropriate cross-national indicators of response quality in open-ended questions. *Social Science Computer Review*. https://doi.org/10.1177/0894439319859848

Microsoft. (2018). Internationalization checklist. https://docs.microsoft.com/en-us/windows/win32/intl/internationalization-checklist. Accessed 18 Oct 2019.

Mohler, P., Dorer, B., de Jong, J., & Hu, M. (2016). Translation: Overview. In *Guidelines for best practice in cross-cultural surveys*. Survey Research Center, Institute for Social Research, University of Michigan. http://www.ccsg.isr.umich.edu/

Pan, Y., Landreth, A., Park, H., Hinsdale-Shouse, M., & Schoua-Glusberg, A. (2010). Cognitive interviewing in non-English languages: A cross-cultural perspective. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, et al. (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 91–113). Wiley.

Pan, Y., Sha, M., & Park, H. (2020). *The sociolinguistics of survey translation*. Routledge.

Pötzschke, S., & Braun, M. (2017). Migrant sampling using Facebook advertisements: A case study of polish migrants in four European countries. *Social Science Computer Review, 35*(5), 633–653. https://doi.org/10.1177/0894439316666262

Pym, A. (2011). Website localization. In K. Malmkjaer & K. Windle (Eds.), *The Oxford handbook of translation studies* (pp. 410–423). Oxford University Press.

Röder, A. (2018). Methodische Herausforderungen quantitativer Befragungen von Geflüchteten am Beispiel einer Vorstudie in Sachsen. *Z'Flucht Zeitschrift für Flüchtlingsforschung, 2*(2), 313–329.

Sandrini, P. (2008). Localization and translation. *MuTra Journal, 2*, 167–191. http://translationconcepts.org/pdf/MuTra_Journal2_2008.pdf#page=167. Accessed 18 Oct 2019

Schäler, R. (2010). Localization and translation. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of translation studies* (Vol. 1, pp. 209–214). John Benjamins B.V.

Sha, M., Hsieh, Y., & Goerman, P. (2018). Translation and visual cues: Towards creating road map for limited English speakers to access translated internet surveys in the United States. *Translation & Interpreting, 10*(2), 142–158. https://doi.org/10.12807/ti.110202.2018.a10

Sin-wai, C. (2016). *The future of translation technology: Towards a world without Babel*. Routledge.

Smith, T. (2004). Developing and evaluating cross-national survey instruments. 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. 431–452). Wiley.

Upsing, B., Gissler, G., Goldhammer, F., Rölke, H., & Ferrari, A. (2011). Localisation in international large-scale assessments of competencies: Challenges and solutions. *Localisation Focus, 10*(1), 44–57.

Valli, P. (2019). Fundamentals of localization for non-localizers. In B. Maylath & K. St. Amant (Eds.). Translation and localization: A guide for technical and professional communicators. : Routledge.

van de Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17–45). Cambridge University Press.

Wang, L., Sha, M., & Yuan, M. (2017). Cultural fitness in the usability of US census internet survey in Chinese language. *Survey Practice, 10*(3). https://doi.org/10.29115/SP-2017-0018

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health, 8*(2), 94–104.

Wild, D., Eremenco, S., Mear, I., Martin, M., Houchin, C., Gawlicki, M., et al. (2009). Multinational trials—Recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: The ISPOR patient-reported outcomes translation and linguistic validation good research practices task force report. *Value in Health, 12*(4), 430–440.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Sage.

Willis, G. B. (2015). Research synthesis. The practice of cross-cultural cognitive interviewing. *Public Opinion Quarterly, 79*, 359–395.

Willis, G. B. (2016). Questionnaire pretesting. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 359–381). Sage.