Translation of the Washington Group Tools

Introduction

Tools developed by the Washington Group on Disability Statistics (WG) are intended for use internationally with the goal of producing cross-nationally comparable data on disability. In order for the Washington Group modules (both questions and response options) to be understood in a way that is consistent both within and across countries, it is necessary to have a standardized translation procedure that yields equivalent versions of the questions across a variety of settings and cultures.

The Importance of Translation

The need for accurate translation will serve both national and international purposes. Rare is the country that is unilingual – having a single language in common daily use. Researchers, demographers and survey methodologists are often confronted with the challenge of translating survey questions into multiple languages and often adapting to multiple dialects.

A good translation goes beyond the literal translation of *words* into another language and assures that *concepts* are appropriately reflected and consistent in all versions of the questionnaire.

As an example, the WG Short Set on Functioning question that concerns cognitive difficulty is: "Do you have difficulty remembering or concentrating?" In a few instances it has been observed that 'difficulty remembering' has been translated as 'difficulty memorizing' or 'difficulty with bad memories'. These three constructs are not the same: the first focuses on cognitive abilities (this is what we are interested in), the second introduces learning abilities, and the last one is about affect rather than cognition. For consistency, it is *essential* that the translation from source to target language captures the *same construct*.

> The **Washington Group Implementation Documents** cover the tools developed by the Washington Group on Disability Statistics (WG) to collect internationally comparable disability data on censuses and surveys. The documents address best practices in implementing the Short Set, Extended Set, Short Set – Enhanced, the WG / UNICEF Child Functioning Modules for children 2-4 and 5-17 years of age, and the WG / ILO LFS Disability Module, as well as other WG tools. Topics include translation, question specifications, analytic guidelines, programming code for analyses, the use of the tools for the purposes of disaggregation, and more.
>
> To locate other WG Implementation Documents and more information, visit the Washington Group website: http://www.washingtongroup-disability.com/.

Translation Methods

Different methods have been proposed to standardize translation processes. The two most prominent methods are forward-backward translation and the conceptual or team translation method. Computer-based translation software should be avoided at all costs!

<u>Forward-Backward Translation</u>

The forward-backward method begins with a version of the question set in the language in which it was originally developed, for example, English (source language). This version is given to professional translators who translate the module into another (target) language, for example, French (forward translation). Then, a different professional translates the French version back into English (back translation) and the two English versions are compared. Strict, literal forward/backward translation is not optimal, however, since the method does not necessarily capture the underlying concept being measured nor are nuances of the target language accounted for.

<u>Conceptual or Team Translation</u>

Alternatively, the conceptual or team translation method relies on detailed understanding of the terms used in survey questions as well as the underlying concepts that the questions were intended to measure. This translation model incorporates five functions (TRAPD, described in detail below):

1. Translation
2. Review
3. Adjudication
4. Pretesting
5. Documentation

The Washington Group agreed to adopt the conceptual translation method since disability is a complex concept where the meaning of the term is sensitive to context and culture. In order to have question sets that will yield internationally comparable results it is necessary to focus on conceptual congruence so that respondents in different countries and cultures will have the same understanding of the questions, regardless of whether the exact terms used are the same.

Description of the WG Translation Process

Translation of the Washington Group questions should be based on the final, approved version of the questionnaire in English (source language). If an approved translation exists in a language similar to the translation target language, it would be helpful to consult that translation. However *all* steps in the translation process should be followed.

The translation of each question is based on an English 'translation card'. Translation cards explain to the translator the purpose of the question, why particular wording is used, and describes in detail the main concepts that are being measured.

Below is an example of translation card for one of the Washington Group Short Set questions.

Question1: Do you have difficulty seeing, even if wearing glasses?

> The purpose of this question is to identify persons who have vision difficulties or problems seeing even when wearing glasses (if they wear glasses).
>
> <u>Seeing</u> refers to an individual using his/her eyes and visual capacity in order to perceive or observe what is happening around them.
>
> <u>Even when wearing glasses</u> refers to difficulty seeing with glasses if the respondent has, and uses, them – NOT how vision would be if glasses, or better glasses, were provided to one who needed them.
>
> Included are problems seeing things close up or far away.
>
> Included are problems seeing out of one eye or only seeing directly in front but not to the sides.
>
> Any problem with vision that the respondent considers a problem should be captured

In addition to using professional translators, it is preferable to choose individuals who are working in the field of health/disability and who would have a better understanding of the subject matter and what the questions are trying to capture.

Translators require:
- very good knowledge of the source language (English)
- an excellent command of the target language
- familiarity with the subject matter and the intent of the questions

A competent translator is not only bilingual but bicultural.

In the TRAPD model, two individuals or teams, working separately, *translate* the questions (and responses) into the target language. The translators/teams and at least one reviewer then meet to *review* the translations and make comments on issues they find or changes they recommend – by consensus. 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. This approach differs from that used in the "backward" step of the forward-backward method when, rather than translating the question back into the original language, an adjudicator determines whether or not each question was properly translated such that the intended concepts were actually captured.

Testing of the Translation

When the translation is complete, it should be *tested*. The testing of translated questions is an exercise in validation. Evaluations of the translations focus on issues such as whether the substantive content of a source question is captured in the translation, whether there are changes in pragmatic meaning (what

respondents perceive as the meaning), and whether technical aspects are translated and presented appropriately (e.g. linguistic and survey appropriateness of response scales).

The goal of testing is to assess whether the translation of the survey instrument in the target language accurately reflects all aspects of the source language instrument.

The patterns of interpretation of a question or text in the target language are expected to be comparable to those in the original, source language; and both of these should match the question developer's intent. While it is often the case that fewer resources are devoted to testing survey translations than questions designed in the source language, this testing should not be overlooked.

Focus groups and cognitive interviews can be used to gain insight into the appropriateness of language used in survey translations.

Focus groups can be used to gain target population feedback on item formulation and how questions are perceived. They are generally not suitable for testing questions or the assessment of entire (lengthy) questionnaires. To optimize efficiency, written materials can be used to ask participants to explain terms including answer categories. Oral and aural tasks are more suitable than written when target population literacy levels are low, when oral/aural mode effects are of interest or to assure that the information to evaluate the translation is obtained.

Cognitive interviews allow for problematic issues to be probed in depth, and can identify terms not well understood across all sub-groups of the target population.

Protocols should be developed and documented for all types of tests, with particular care toward designs to investigate potentially concerning survey items.
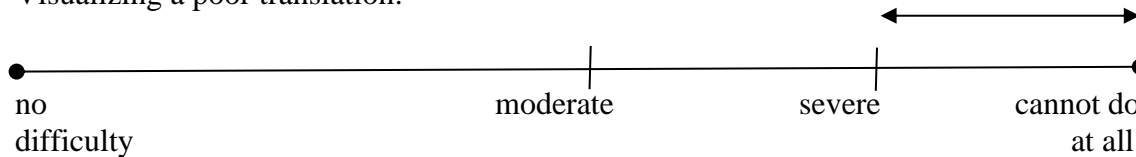
Interviewer and respondent debriefings can be used after all types of pretests, with full documentation of debriefing, to collect feedback and probe comprehension of items or formulations.

Throughout the process, decisions made at every step are *documented* to inform designers and analysts about how the final translation was reached.

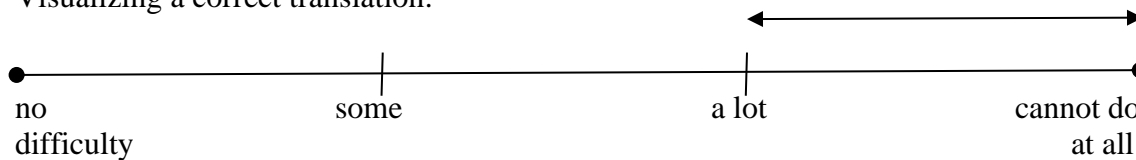## Considering the Translation of Response Options

Translations of questionnaires often focus on translating the question stem. While this is critical, of equal importance is the correct translation of the answer categories. This is particularly important for the Washington Group questions which attempt to capture a continuum of functioning. Improper translation of the answer categories will skew the resulting distribution. 'No difficulty' and 'cannot do at all' are the extremes of the distribution, anchoring its endpoints, and are clear concepts that are unambiguous and straightforward for translation. 'Some difficulty' and 'a lot of difficulty' are less definitive. For example, 'some' and 'a lot' should not be translated using words interpreted as 'moderate' and 'severe' respectively. The result of such a translation would skew the distribution towards the 'cannot do at all' end of the continuum and would miss many of those at risk if the recommended cut-off of 'a lot of difficulty' was utilized. The words used to translate 'some' and 'a lot' should divide the continuum into three relatively equal parts.

Visualizing a poor translation:

```
●————————————————|————————————|——————————————●
no                        moderate        severe        cannot do
difficulty                                                at all
```

Using 'severe' as a translation for 'a lot of difficulty' makes this category closer to 'cannot do at all' so when using the recommended cut-off (a lot of difficulty or cannot do at all) many of those with difficulties of interest will be missed. This translation would also leave those with less than moderate – but some difficulty – in a position of potentially choosing 'no difficulty'. The result is that the proportion of the population with a disability (in the 3rd or 4th categories) will be smaller than it would be had the terms been translated appropriately.

Visualizing a correct translation:

```
●————————————————|————————————|——————————————●
no                        some              a lot         cannot do
difficulty                                                at all
```

The distribution above creates four points equally distributed among the continuum that allows respondents to more easily discriminate among options. The recommended cut-off correctly identifies those with the intended level of difficulties. The translation of 'some' and 'a lot' should identify the population with the same functional status as that identified by the cognitively tested English version. Whether this has been successful can be determined by cognitively testing the translated questions.

Administration of Translations

Prior to beginning the data collection, ensure that there is a version of the questions in the local language. All translation should be done before the survey is fielded. Translation by the interviewer at the time of the interview ["on the fly"] should be avoided since small differences in interpretation of a phrase or even a single word can destroy the reliability and validity of the data. In countries where there are many local languages and resources are not available for 'official' translations into all languages, the WG suggests that teams of interviewers from those specific localities/regions get together before data collection to agree, amongst themselves and with supervision, on a local language version that can be used consistently in the field.

Summary

In addition to being bilingual (with good knowledge of both source and target languages), and bicultural, translators should be familiar with the subject matter and the intent of the questions.

Translations should involve the careful consideration of the constructs of the questions and response options, rather than the words. Literal translation should be avoided in favor of a conceptual, consensus approach. Computer-based translation software should be avoided at all costs!

References

Comparative Survey Design and Implementation: Cross Cultural Survey Guidelines. Online at:
http://ccsg.isr.umich.edu/index.php/chapters/translation-chapter

Miller K. Cognitive Interviewing Methodology to Examine Survey Question Comparability. NCHS 2017 (forthcoming).

Schoua-Glusberg A, Villar A. Assessing Translated Questions via Cognitive Interviewing. In: *Cognitive Interviewing Methodology*, (K Miller, S Willson, V Chepp & JL Padilla (eds). Wiley Series in Survey Methodology, Wiley & Sons, 2014:51-67.

For more information on the Washington Group on Disability Statistics, visit:
http://www.washingtongroup-disability.com/.