Translation is not adaptation

A checklist for adapting a questionnaire into another language

Dr Milos Kankaras | Center for Psychology | Podgorica, November 2026

Who the checklist is for. Master's and doctoral students and first-time authors who use a questionnaire or scale written in another language. It is free and may be shared in full.

Why translation is not enough

A translated scale is text. An adapted scale is text plus evidence that its items mean the same thing in the new language, that people answer in the same way and that the scale measures the same construct. The International Test Commission (International Test Commission [ITC], 2018) therefore asks, beyond translation, for evidence that instructions and items have similar meaning, that response formats are suitable and that reliability and structure were checked on data from the new population. A review of 31 guidelines for cross-cultural adaptation of questionnaires found no consensus on methods and treated adaptation and validation as two different processes (Epstein et al., 2015). It is therefore reasonable to rely on one openly available source, such as the ITC guidelines, and to cite it.

The checklist has five stages. Each says what to do and what the paper must state. If a stage stays empty, say so in the paper instead of leaving it out.

The five stages an adaptation must show

Stage 1. Before translation: rights and construct

- Obtain permission from the author or rights holder before any adaptation (ITC, PC-1). That an instrument is easy to find online does not mean it may be translated.
- Check that the construct has the same meaning in your population and is close enough to the purpose for which you will use the scores (ITC, PC-2).
- **State in the paper:** the source and version of the instrument, the author, to whom the request was addressed and what was answered.

Stage 2. Translation

- Use at least two people whose first language is the target language, who know the subject and, ideally, live where the scale will be used (ITC, TD-1). ITC does not regard translation by a single person as acceptable practice.

- Have two translations made independently, then have a third person or an expert group reconcile them into one version (ITC, TD-2).
- Use back translation as an additional check, not as the only evidence. ITC warns that in its narrowest form the procedure never reviews the target version, so the text can be easy to translate back and still read unnaturally.
- Decide in advance which variety of the language the scale is adapted for and record where it differs from the others. The same item may not read equally clearly in every country or community that shares the language, and a pretest, not an assumption, tests that.
- **State in the paper:** who translated and with what qualifications, how reconciliation proceeded and which decisions were disputed.

Stage 3. Evidence of meaning: expert review and pretest

- Have a language expert and a subject expert review every item, with questions such as whether the translation omitted, substituted or added anything (ITC, TD-2 and TD-3).
- Run a small pretest with respondents from the target population, with a conversation afterwards: what they understood by each item and what struck them as odd. ITC warns that comments from respondents and administrators are often more valuable than the answers to the items (TD-3).
- Check the instructions, response options and mode of administration, not only the items (ITC, TD-4).
- **State in the paper:** how many respondents took part, how they were chosen, what the conversations found and which items were changed as a result, with two or three examples.

Stage 4. Pilot and numbers

- Collect pilot data before the main study. For item analysis ITC considers even a modest sample useful, for example one hundred respondents, while checking factor structure needs fairly large samples, perhaps three hundred respondents or more (ITC, TD-5; C-1).
- Inspect the items: items that are too easy, too hard or have a low or negative relation to the scale need a look before you decide that they stay (ITC, TD-5).
- Report reliability with both coefficient alpha and coefficient omega. Alpha assumes equal factor loadings and, where they differ, is only a lower bound of reliability, while omega is more generally applicable (ITC, 2018). McNeish (2018) therefore described omega and other indices with less strict assumptions as alternatives that have advantages over alpha.
- Check the factor structure on your own data, by exploratory or confirmatory analysis, and state the criteria by which you judged model fit. Do not change the model until it fits without saying so.
- Look for evidence from relations with related measures: expected relations should be present and unexpected ones absent (ITC, C-2).
- **State in the paper:** the size and description of the sample, how it was collected, all results, and why the sample is a limitation.

Stage 5. Reporting and the limits of the conclusion

- Document every change from the original and the evidence that justifies it (ITC, Doc-1).

- Norms and cut-offs from the original do not transfer by themselves. An adapted version needs evidence for norms, reliability and validity in the new population (ITC, C-3). Where there is none, do not use the original norms and cut-offs to classify respondents.
- Compare scores of different language groups only once measurement invariance has been checked at the level at which the scores are compared (ITC, SSI-2). Conventions for reporting invariance are still in flux (Putnick & Bornstein, 2016), so state the procedure and criteria you used.
- Write "adapted version", and "valid" only for what has been shown, for a stated purpose and population.

What each claim in the paper must have behind it

Faculties and committees ask for different things, so treat this as a frame for the conversation with your supervisor, not as a rule. Every claim in the methods section needs evidence that can be shown.

Claim in the paper	Evidence that supports it
The scale was translated	Two independent translations, reconciliation, translators' qualifications, back translation as a supplement
The items are understandable	Pretest with a conversation, list of changes and reasons
The scale is reliable	Alpha and omega on your sample, with the sample size
The structure of the scale is confirmed	Analysis on your data, criteria, every decision about the model
Scores can be compared with another group	Measurement invariance at the appropriate level
Norms or cut-offs from the original are used	Evidence that they hold in the new population, or leave them out

Three failures that recur

A literal translation reported as valid. Translation is the first step, but the evidence from stages 3 and 4 speaks to the validity of scores. Where it is missing, the paper says "translated scale" and names the limitation.

Back translation as the only evidence. If the back translation matches the original, that shows only that the text can be turned back. It does not show that a respondent understands it. Add expert review and a pretest.

Numbers without limits. Alpha and factor analysis on a small convenience group, followed by a comparison with other countries or classification by foreign norms. The numbers are correct and the conclusion drawn from them is not shown. State the sample size, the limitation and what the result does not allow.

Before you submit

Your supervisor should read the methods section. If you also need an independent expert reading of the adaptation method, that is a separate service, and this checklist does not replace it.

About the author

Dr Milos Kankaras is a psychometrician with a PhD in social sciences (Tilburg University) and more than 25 years of research experience in developing and checking measurement instruments, including large international studies. The Center for Psychology also offers measurement-quality services, so the reader should know about that interest. The checklist does not recommend or reject any individual instrument.

Questions and comments: info@centerforpsychology.me

References

Epstein, J., Santo, R. M., & Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of Clinical Epidemiology, 68*(4), 435-441. https://doi.org/10.1016/j.jclinepi.2014.11.021

International Test Commission. (2018). ITC guidelines for translating and adapting tests (second edition). *International Journal of Testing, 18*(2), 101-134. https://doi.org/10.1080/15305058.2017.1398166

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods, 23*(3), 412-433. https://doi.org/10.1037/met0000144

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71-90. https://doi.org/10.1016/j.dr.2016.06.004