Experiment for testing questionnaire translation methods in the European Social Survey (ESS): “Ask the same question” (close) versus more adaptive approaches

Brita Dorer, GESIS (Mannheim – Germany)
Ana Villar (City, University of London)

CSDI, Mannheim, 16 March 2017
Content

- Research questions / goals
- Research design
- Analysis
- CSDI: discussion / thoughts?
Research questions / goals

- One approach **better** than other?
  - Measurement quality
  - Measurement comparability

- Results dependent on languages?
  - Linguistic structure → affect understanding and how question is processed

- Results dependent on country/culture?
  - Factual differences across realities

- Results dependent on type of adaptation (e.g. factual / linguistic / degree of adaptation)?
  - Fail to carry out factual adaptation more problematic than linguistic structure
Independent variable:
- Experimental manipulation
- 2 translation approaches

Dependent variable:
- Measurement quality
- Measurement comparability

Moderating variables:
1. Distance between
   1a. Language (linguistic structure)
   1b. Country/culture (different realities)
2. Type of adaptation
Research design

- Experimental design
- Cross-National Online Survey (CRONOS) panel
- 3 languages: Slovene (Slovenia) – Estonian and Russian (Estonia)
- Each language: 2 translating teams applying both methods
  - Team 1 applies method 1 to Q1-Q20, and method 2 Q21-Q40
  - Team 2 applies method 2 to Q1-Q20, and method 1 Q21-Q40
- 2 parallel Translations + Review / Adjudication
Translating teams + Review session

- Each language: 2 teams of 3 people: 2 translators + 1 reviewer: TRA
- If possible, experience in translating social sciences surveys
  - Not easy to find in all languages
- If possible not having translated ESS / SHARE / ISSP before

- Review session to be recorded
Instructions of translating teams

- Abridged ESS translation guidelines, PLUS

<table>
<thead>
<tr>
<th>Method 1</th>
<th>Method 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>stay as close as possible or sensible</td>
<td>adapt natural language use or survey habit in target country or target language</td>
</tr>
<tr>
<td>not modify semantic or pragmatic meaning of source text</td>
<td>translation should not sound like translation but rather like standalone text</td>
</tr>
<tr>
<td>not word-for-word, but as close as possible</td>
<td></td>
</tr>
</tbody>
</table>

- Trying to re-start from scratch on second approach, even after having applied first method
Organising the translations and review meetings

- Hopefully feasible scenario (avoid contamination of approaches)
  1. Each translator to produce translations following the first assigned method by deadline
  2. Wait 1-2 weeks
  3. Each translator to produce translations following the second assigned method
  4. Each team has a first review meeting (1\textsuperscript{st} assigned method) on an afternoon
  5. Each team has a second review meeting (2\textsuperscript{nd} assigned method) the morning after
Quality assurance, quality monitoring

• QA: via reviewer
  – How much do they need to know about the study?
  – The reviewer should try to ensure that the agreed approach is followed to the extent possible

• QM: via recording of meeting + consulting native speakers
Item selection

- Items selection almost finished: Around 50 questionnaire items (from ESS / SHARE / ISSP)
- Different forms of adaptation:
  a) factual:
     factual information needs to be adapted, e.g. ‘housing complex with services for older people’, types of housing
  b) linguistic-semantic (meaning):
     e.g., difference between ‘blame’ and ‘feel guilty’
Item selection

- c) linguistic-pragmatics (language use):
  e.g.: “would you say that most people can be trusted, or that you can’t be too careful?”
  A: You can’t be too careful
  B: Most people can be trusted
  =>in RU these are not understood as opposites; how will the opposite be handled?
Item selection

- d) **linguistic-structure** (grammar, syntax):
  e.g.: “How likely do you think it is that large numbers of people will actually limit their energy use to try to reduce climate change? Not at all likely – Extremely likely”

1) There is no grammatical future in Estonian: present tense is used instead / other ways to imply future in text
2) RCs: extreme vs. non-extreme translations
Analysis

Analysis ↔ items selected
Depend on each other

Depending on questionnaire items
• Structural equation models / factor analysis
• Comparing to external benchmark
• ‘Simple’ comparison of results
• Test-retest

Content analysis of the review meetings: quality monitoring, but also learn about how instructions work
Discussion / CSDI thoughts?

- Other features / adaptation types than factual and linguistic?
- Instructions? // How much should reviewers (and translators) know about study?
- Scenario: 1-2 weeks between translations resp. 1 night between Review sessions? =>sufficient for avoiding contamination of methods?
THANKS!

Ana & Brita
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654221.
THANKS !

Ana & Brita
Instruction of translating teams

- Documents not finalised
- Trying to re-start from scratch, even after having applied first method
- 1: → stay as close as possible
  → not modify semantic or pragmatic meaning of source text
  → not word-for-word, but as close as possible
- 2: → adapt as would be natural language use or survey habit in target country or target language
  → translation should not sound like translation but rather like standalone text
Organising the translations and review meetings

• Ideal scenario

  – Each team carries out

  – BRITA TO ADD: BASICALLY DO THE ENTIRE PROCEDURE WITH METHOD 1, WAIT 1-2 WEEKS, DO METHOD 2

  – EXPLAIN CHALLENGES: HAVING TWO SEPARATE MEETINGS.
“Ask the same question” (ASQ)

- So far considered as best practice in multilingual / cross-cultural academic social sciences surveys, such as ESS
- ALL languages versions as close as possible to source questionnaire

→ enhance comparability between all versions

- BUT: is this really the best approach?
## Different forms of adaptation

<table>
<thead>
<tr>
<th>Type</th>
<th>Domain</th>
<th>Subtype</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Language</td>
<td>Linguistics-driven</td>
</tr>
<tr>
<td>2</td>
<td>Language</td>
<td>Pragmatics-driven</td>
</tr>
<tr>
<td>3</td>
<td>Language</td>
<td>Comprehension-driven</td>
</tr>
<tr>
<td>4</td>
<td>Culture</td>
<td>Terminological/factual-driven</td>
</tr>
<tr>
<td>5</td>
<td>Culture</td>
<td>Norm-driven</td>
</tr>
<tr>
<td>6</td>
<td>Measurement</td>
<td>Familiarity-/recognizability-/format-driven</td>
</tr>
<tr>
<td>7</td>
<td>Measurement</td>
<td>Other</td>
</tr>
</tbody>
</table>

(Behr 2013)
Examples

“walking 2 blocks”

- ‘block’ not existing in all countries
- if so, may have different dimensions
- allow to adapt?
  (e.g., 300 meters / distance between 2 street lights)

“putting campaign badges or stickers”

- scarfs or other gadgets may be more common in other countries
Examples

“not very vs. not at all satisfied”

• In Slavic languages, this difference cannot be made for linguistic reasons
• Create difference artificially or modify answer scale?