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)  

ESRA, Lisbon, 18 July 2017
1. Background
Dilemma

• Most surveys: “Ask-the-same-question (ASQ)”
• No definition of what ‘same question’ means
• Researchers make different interpretations
  – ‘As close as possible’ at the word level
  – Others focus on making stimulus functionally equivalent and believe this requires adaptation
• No empirical evidence yet on which interpretation leads to more comparable data
Deciding level of closeness: example

How much of the electricity used in [country] should be generated from each energy source?

Energy sources: Coal, Natural gas, Hydroelectric power, Nuclear power, Solar power, Wind power, Biomass

- A very large amount
- A large amount
- A medium amount
- A small amount
- None at all
Deciding level of closeness: example (cont'd)

• Advance translation and translation verification suggested in most languages 'amount'
  → share / portion / part

• BUT portion posed limitations on the answers: you can't have a very large portion of all energy sources

• TEP: Natural language use vs. safe translation?

• Which version yields most comparable data?
## 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)
2. Research goals
Research goals / questions

• What to do when "dilemmas" come up?
• Which approach is better: 'as close as possible' vs. 'encouraging adaptation'
  – Better = yielding more comparable data
• Under which conditions is that approach better? Does the effect differ by...
  – Language: are some languages more resilient to close translations than others?
  – Type of adaptation: factual, linguistic...
  – Degree of adaptation: are some adaptations "too much"?
Overall model

Independent variable:
Experimental manipulation
2 translation approaches

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

Dependent variable:
Measurement quality
Measurement comparability
Research design

• 2 sets of instructions produced: close vs adapted
• 2 languages: Slovene (Slovenia) & Estonian (Estonia)
• For each language 3 teams and 3 sets of items:
  – 2 parallel Translations following experimental instructions
  – Review / Adjudication meeting with instructions present
• Manipulation checks
  – Audio taping
  – Follow up interviews with translators
• Data collection: Cross-National Online Survey (CRONOS) panel
  – 60 items, 500 respondents per language (including source language)
Instructions excerpt

Encouraging close TR:
• As close as possible and sensible, not literal
• Don’t modify semantic and pragmatic meaning
• Don't modify examples, idioms, symbols
• Survey speak allowed

Encouraging adaptation:
• To make suitable to the socio-cultural context
• Idioms, colloquial language
• Adapt symbols
• Should not sound like translation
Close = 'As close as possible'
Adapt = 'Encouraging adaptation'

<table>
<thead>
<tr>
<th>Items</th>
<th>Team 1</th>
<th>Team 2</th>
<th>Team 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-20</td>
<td>Close</td>
<td>Adapt</td>
<td>Adapt</td>
</tr>
<tr>
<td>21-40</td>
<td>Adapt</td>
<td>Close</td>
<td>Adapt</td>
</tr>
<tr>
<td>41-60</td>
<td>Close</td>
<td>Adapt</td>
<td>Close</td>
</tr>
</tbody>
</table>
Organising the translations and review meetings

• Challenge: avoid contamination of approaches
  1. Each translator to produce translations following the first assigned method by deadline
  2. Each team has a first review meeting (1\textsuperscript{st} assigned method)
  3. Wait 1-2 weeks
  4. Each translator to produce translations following the second assigned method
  5. Each team has a second review meeting (2\textsuperscript{nd} assigned method)
  6. Another 1-2 weeks
  7. Etc.
Item selection

- 60 questionnaire items (from ESS / SHARE / ISSP / EVS)
- Guiding principles:
  - Include a variety of forms of adaptation
  - Chose questions that allow invariance testing
  - Questions that had been shown problematic in translation, especially in Estonian and Slovenian
Analysis

• Analysis ↔ items selected
• Depending on questionnaire items
  – Multigroup Confirmatory Factor Analysis
  – Item response theory
  – Web probing
  – Questions from benchmark surveys
  – Content analysis of the review meetings: also learn about how instructions worked
Challenges setting up the experiment

• Define the level of closeness and adaptation for each experimental condition
• Translate that definition into a manipulation: instructions
• Finding suitable translators and reviewers
• Ensure that participants adhere to the instructions
• Select questions that may be affected in all languages
• Defining what makes an approach better
  – Control for translator quality effects
  – Control for type of adaptation effects
  – Control for language effects
Special thanks to

- Natalja Menold, GESIS
- ESS Translation Expert Panel
- ESS Methods Advisory Board
- Frederic Mahler, SHARE
THANKS !

Ana & Brita

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 654221.