FORS

explore.understand.share.
When does adding a survey language make sense?
Representation bias, response rates, and strategic considerations

3MC 2016
28.07.2016, Chicago

Michael Ochsner & Sven E. Hug
Outline

- The issue: Language and representation bias
- Two surveys adding/reducing survey languages
  - Real-life data: not a clean experimental design
- Results
- Conclusions
The Issue: Language and Representation Bias

- Minorities are underrepresented
  - Adding languages might help get them in
- Political reasons
  - People might not like the survey when a language is missing
- Adding languages is costly
  - Financial pressures → Change of mode
Two surveys

- Survey of humanities scholars about research quality
  - Focus on three disciplines
  - First survey two languages (German & English)
  - Second survey four languages (German, English, French, Italian)

- Mixed-Mode Experiment of a population survey in a Swiss city
  - Multilingual survey (6 languages)
  - CATI (severe drops in response rates)
  - Web/Paper Experiment (only one language: German)
Survey of Scholars: Adding Languages

- **Response rate**
  - First survey 44%, second survey 39% (usually 15-25%)

- **Language Region:**
  - Deviation in first survey: DCH -2%; WCH: 0%
  - Deviation in second survey: DCH: -2%; WCH: 10%

- **Language x Discipline**

<table>
<thead>
<tr>
<th>Language x Discipline</th>
<th>1st Survey (2 Lang)</th>
<th>2nd Survey (4 Lang)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DCH</td>
<td>WCH</td>
</tr>
<tr>
<td>German Lit</td>
<td>-0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>English Lit</td>
<td>-0.09</td>
<td>0.21</td>
</tr>
<tr>
<td>Art History</td>
<td>0.10</td>
<td>-0.26</td>
</tr>
</tbody>
</table>
Survey of Scholars: Adding Languages

- Mother tongue of the participants

<table>
<thead>
<tr>
<th></th>
<th>1st Survey (2 Lang)</th>
<th>2nd Survey (4 Lang)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>German</td>
<td>English</td>
</tr>
<tr>
<td>All</td>
<td>71</td>
<td>13</td>
</tr>
<tr>
<td>German Lit</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>English Lit</td>
<td>35</td>
<td>53</td>
</tr>
<tr>
<td>Art History</td>
<td>61</td>
<td>0</td>
</tr>
</tbody>
</table>
General Population Survey: Reducing Languages

- **Response rates**
  - Compensation with mode: Web/Paper vs. CATI
  - 33% RR CATI vs 56% RR Web/Paper
    - Mode stronger than language
- **Response bias sample (% deviation of rr)**
  - Permit: Deviation CATI: 29%, Web/Paper: 10.5%
  - Total (6 Variables): CATI 100%, Web/Paper: 43%
- **Response bias population (incl sample bias)**
  - Migration: Swiss: CATI: -1%, Web/Paper: 7%
  - Migration: other: CATI: 16%, Web/Paper: 13%
  - Nation, Age, Education: CATI: 27%, Web/Paper: 26%
General Population Survey: Reducing Languages

- Language
  - No information on population
  - Web/Paper: higher share of German
  - Web/Paper: more equilibrated share of other languages
Conclusions

- Adding Language
  - Expensive
  - Translation, Comparability
  - Additional language can reduce bias

- Mode can compensate for language
  - Mode can reduce more bias than language

- When is it worth adding a language?
  - Ethical/political reasons
  - Small or big N (additional cases make a difference)
  - Interest in members of group speaking a language
  - Language is related to concept in question
Thank You for Your Attention!

- **Links:**
  - michael.ochsner@fors.unil.ch
  - sven.hug@gess.ethz.ch

- **Contact:**
  - FORS
  - Géopolis
  - CH-1015 Lausanne
  - Phone: +41 21 692 37 30
  - http://www.forscenter.ch