Interviewer effects on response latencies in a face-to-face interview survey

Patrick Sturgis, University of Southampton

Paper presented at the NatCen-City Survey Methods seminar October 2018
Joint work with

• Gabi Durrant, University of Southampton

• Olga Maslovskaya, University of Southampton

• Ian Brunton-Smith, University of Surrey
Programme of work on response latencies

• Decomposing variability in response times into respondent, question and interviewer components

• Does satisficing reduce cognitive costs of responding?

• Personality and survey satisficing
Response latency

• Indicator of attitude strength

• Indicator of respondent effort/measurement error (cf. satisficing)

• Interview length -> survey cost

• Cf. influence of interviewers

• Use in survey practice?
Data

- Wave 3: UK Household Longitudinal Survey (2011-2013), n=30,685
- Exclude Scotland and N. Ireland (no MSOA identifiers)
- Exclude ethnic minority boost sample + BHPS
- Exclude interviewers who did not complete interviewer survey
- Analysis sample, n=~13,904 within 288 interviewers
- ~370 questions with auto-generated CAPI time stamps
  - = 3,149,871 records
Interviewer characteristics

• Survey of NatCen interviewers in 2014

• Online and paper self-completion

• 473 interviewers completed the survey of 823 invited (58%)

• Demographic characteristics

• Attitudes to fieldwork, data privacy, personality traits
Respondent characteristics

- Age
- Gender
- Education
- Satisficing response styles
  - Don’t knows
  - Mid-points
  - Straight-lining/non-differentiation
  - Extreme responding
- Personality
- Cognitive ability
Question characteristics

- Team of 4 coders coded questions according to observable characteristics:
  - Number of words, interviewer instructions, position in battery, order in questionnaire, response options, show-card, transition statement, definitions, visual emphasis, self/interviewer-administered
  - Also by question type: demographic, income, behaviour, test, event and whether question addressed potentially sensitive topic
Data Structure

Level 3
Interviewers/
areas

Level 2
Respondents/
Questions

Level 1
Latencies
Cross-classified multi-level model

$$\log(y_i) = x'_i \beta + g_m + f_l + v_k + u_j + e_i,$$

$x'_i = \text{interviewer, area, respondent, question characteristics}$

$\beta = \text{coefficient estimates for } x'_i$

$g_m = \text{area random effect}$

$f_l = \text{interviewer random effect}$

$v_k = \text{respondent random effect}$

$u_j = \text{question random effect}$

$e_i = \text{latency specific residual}$
Latency distributions

a) Neighbourhood cohesion

b) Perceived political influence

c) Self completion
Latency distribution and outliers

- Mean latency = 11 seconds
- Median latency = 7 seconds
- Range = 0-179 seconds
- 18% of latencies were <4 seconds
- Investigation revealed 0 seconds were due to over-writing
- 1-3 seconds were due to interviewer behaviour
- 0 latencies and latencies >100 seconds are excluded
Results
## Variance Components

<table>
<thead>
<tr>
<th>Random effects</th>
<th>Our model</th>
<th>Couper &amp; Kreuter 2013 (females)</th>
<th>Couper &amp; Kreuter 2013 (males)</th>
<th>Olsen &amp; Smyth 2015 (CATI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>interviewer</td>
<td>3.5%</td>
<td>1.5%</td>
<td>1.9%</td>
<td>3.2%</td>
</tr>
<tr>
<td>area</td>
<td>0.6%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>respondent</td>
<td>4.8%</td>
<td>3.8%</td>
<td>6.3%</td>
<td>7.4%</td>
</tr>
<tr>
<td>question</td>
<td>30.5%</td>
<td></td>
<td></td>
<td>53.4%</td>
</tr>
<tr>
<td>residual</td>
<td>60.4%</td>
<td>(94.7%)</td>
<td>(91.8%)</td>
<td>36%</td>
</tr>
</tbody>
</table>
Number of words & use of showcards

Graph showing the predicted RLs in seconds as a function of the number of words.

Bar chart comparing predicted RLs without and with showcards:
- No showcard: 4.47 seconds
- Showcard: 5.19 seconds
Interviewer experience

Predicted RLs in seconds

- 0-4 years: 4.47
- 5-10 years: 4.33
- 10+ years: 4.16
Respondent education & gender

**Predicted RLs in seconds**

<table>
<thead>
<tr>
<th>Qualification</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree</td>
<td>4.47</td>
<td>4.43</td>
</tr>
<tr>
<td>Other higher degree</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCSE</td>
<td>4.60</td>
<td></td>
</tr>
<tr>
<td>Other qualification</td>
<td>4.73</td>
<td></td>
</tr>
<tr>
<td>No qualification</td>
<td>4.63</td>
<td></td>
</tr>
</tbody>
</table>

[Graph showing predicted RLs for different education levels and gender]
Interviewer instructions & visual emphasis

Predicted RLs in seconds

<table>
<thead>
<tr>
<th>Condition</th>
<th>Predicted RLs in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>No interviewer instructions</td>
<td>4.47</td>
</tr>
<tr>
<td>Interviewer instructions</td>
<td>5.43</td>
</tr>
<tr>
<td>No visual emphasis</td>
<td>4.47</td>
</tr>
<tr>
<td>Visual emphasis</td>
<td>5.27</td>
</tr>
</tbody>
</table>
Position in battery & help screens

![Bar chart showing predicted RLs in seconds for different conditions: Not later in a battery (4.47 seconds), Later in a battery (3.34 seconds), No help (4.47 seconds), Help (5.00 seconds).]
Response options

Predicted RLs in seconds

- Single response option: 4.47 seconds
- Multiple responses: 6.45 seconds
- Open textual or numeric answer: 9.03 seconds
Respondent & interviewer age

- Age of respondents:
  - 10-20: 3.96
  - 21-30: 4.03
  - 31-40: 4.17
  - 41-50: 4.28
  - 51-60: 4.47

- Age of interviewers:
  - 10-20: 3.96
  - 21-30: 4.03
  - 31-40: 4.17
  - 41-50: 4.28
  - 51-60: 4.47
Interviewer attitudes:
‘Wave 1 UKHLS was too long’

Predicted RL in seconds

- Strongly agree: 4.76
- Agree: 4.59
- Disagree: 4.44
- Strongly disagree: 4.27
No effects for most interviewer characteristics

- sex
- Personality & trust
- Satisfaction with work as interviewer
- Attitudes to persuasion
- Attitudes to value of surveys
- Attitudes to privacy
Response-styles

No straightlining tendency

Straightlining tendency

Predicted RLs in seconds

Number of extreme responses

Predicted RLs in seconds

4.47

4.37

0 5 10 15 20 25 30
Interactions
interviewer gender and question type

Type of question

<table>
<thead>
<tr>
<th>Type of question</th>
<th>male</th>
<th>female</th>
</tr>
</thead>
<tbody>
<tr>
<td>demographic</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>attitude and perception</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>friends and family</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>health</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>income</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>behaviour</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>test</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>event</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
age of respondents and straight-lining

![Graph showing the relationship between age of respondents and response latencies in seconds. Two lines are shown: one for low straightlining tendency and one for high straightlining tendency. The x-axis represents the age of respondents, ranging from 16 to 96, while the y-axis represents response latencies in seconds, ranging from 1 to 8. The graph illustrates a positive correlation between age and response latencies.]
Response latencies in seconds

Age of respondents

- 31-40 year old
- 41-50 year old
- 51-60 year old
- 61-70 year old
- 71-79 year old

Age of respondents and age of interviewers
Interviewer & question specific Intra-class correlations (ICC)
The location scale model

\[ y_{i(jk)} = x'_{i(jk)} \beta + u_j + v_k + e_{i(jk)} \]

\[ \ln \left( \sigma^2_{e_{i(jk)}} \right) = w'_{i(jk)} \alpha + u^{[2]}_j \]
Interviewer ICCs over all questions
Interviewer ICCS for specific questions

Q1

Q3
Question specific ICCs
Discussion

• Small but significant interviewer component (3.5%)
• Similar size to previous studies using quite different designs
• Age important for respondents and interviewers
• Pattern for questions intuitive, similar to previous studies
• Little effect of interviewer attitudes/personality/experience
• ICC plots for detecting ‘problem’ interviewers & questions
• Does satisficing reduce cognitive costs?