Smart surveys: How active should respondents be in passive sensor data collection?

Casestudy: Household Budget Survey App

Anne Elevelt, B. Schouten, E. Rodenburg, J. Akkermans, J. de Groot & L. Hollanders
UNECE, Rome– 27 October 2022
Smart surveys

Smart surveys have at least one of the following smart features:
1. Device intelligence
2. Internal sensors
3. External sensors
4. Public online data
5. Personal online data
6. Linkage consent
Why use smart surveys?

- Ease the response task
- Decrease respondent burden
- Reduce respondent inability
- Improve data quality
Why use smart surveys?

- Ease the response task
- Decrease respondent burden
- Reduce respondent inability
- Improve data quality

--> Active – passive data trade-off
Why active data collection?

- Respondent engagement
- Sensor error adjustment
- Legal (ethical)
How active or passive should we make our respondents?
Case study: Household budget survey app
Smart feature: OCR scanning
Results
Motivation and involvement

- Time spent in app

![Graph showing total in-app time (in seconds) per day. The graph compares two modes: Interviewer and Letter. The time decreases over days, with the Interviewer mode generally spending more time in the app compared to the Letter mode.]
Motivation and involvement

- Pages per day
Motivation and involvement

- Entries per day
## Data quality – Activity in app

<table>
<thead>
<tr>
<th></th>
<th>In-app editing</th>
<th>No in-app editing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td></td>
<td>Active</td>
<td>Scanned</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>20.8 (11.8)</td>
<td>20.4 (12.9)</td>
</tr>
<tr>
<td><strong>Scans</strong></td>
<td>4.8 (6.3)</td>
<td>7.6 (6.4)</td>
</tr>
<tr>
<td><strong>Manual</strong></td>
<td>16.0 (13.1)</td>
<td>12.8 (11.0)</td>
</tr>
</tbody>
</table>
# Data quality – In-app editing vs ‘the truth’

<table>
<thead>
<tr>
<th></th>
<th>Submitted by respondent</th>
<th>OCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct amount</td>
<td>7.8%</td>
<td>2.7%</td>
</tr>
<tr>
<td>Difference amount &lt;= 1 Euro</td>
<td>63.9%</td>
<td>11.1%</td>
</tr>
<tr>
<td>Correct number of products</td>
<td>27.6%</td>
<td>16.1%</td>
</tr>
<tr>
<td>Difference number of products &lt;=1</td>
<td>43.5%</td>
<td>45.5%</td>
</tr>
<tr>
<td>No zero amounts</td>
<td>38.0%</td>
<td>18.5%</td>
</tr>
</tbody>
</table>
Conclusion
Conclusion

• Give respondents the option to edit: Enabling in-app editing partially improves data quality
  • No effect on drop-out rates
• Challenge them to do more editing tasks
  • While keeping an eye on whether respondents are doing this reliably
Discussion

• Any experiences with active involvement in receipt scanning?
• Any experiences more generally with active involvement in text extraction?
• Where to set the boundary in active involvement (given the desire to ease tasks)?
Thank you!

a.eleveit@cbs.nl