HOW DO SURVEY INTERVIEWERS IMPACT RESPONDENTS’ SATISFICING TENDENCIES?

An Analysis Based on Audio-Recordings of Face-to-Face Interviews

Bettina Müller
Silvia Schwanhäuser

ESRA 2021 Conference
July 9, 2021
BACKGROUND

• Response styles such as non-differentiation and item-nonresponse are commonly attributed to respondent satisficing

• Effect of interviewer behavior on such response styles has rarely been addressed
  – Assistance and standardized interviewing may counteract satisficing
  – Interviewers may attune to individual response styles, promoting satisficing
  – Interviewers may provoke response styles
• Response styles such as non-differentiation and item-nonresponse are commonly attributed to respondent satisficing

• Effect of interviewer behavior on such response styles has rarely been addressed
  – Assistance and standardized interviewing may counteract satisficing
  – Interviewers may attune to individual response styles, promoting satisficing
  – Interviewers may provoke response styles

• Evidence of interviewer influence, but lack of direct measures to study mechanisms
BACKGROUND

• Response styles such as non-differentiation and item-nonresponse are commonly attributed to respondent satisficing

• Effect of interviewer behavior on such response styles has rarely been addressed
  – Assistance and standardized interviewing may counteract satisficing
  – Interviewers may attune to individual response styles, promoting satisficing
  – Interviewers may provoke response styles

• Evidence of interviewer influence, but lack of direct measures to study mechanisms

• Drawing on audio-recordings of F2F interviews in the German panel study “Labour Market and Social Security” (PASS), we study the interaction of interviewer and respondent effects on non-differentiation, extreme responding, and item-nonresponse
RESPONDENT-INTERVIEWER INTERACTION

• Standardized interviewing
  – Ideal question-answer process: Respondents interpret question meaning, retrieve requested information, integrate recalled information and question meaning, edit and communicate answer (Tourangeau et al. 2000)
  – Interviewers ensure that respondents thoroughly engage with all of these steps
RESPONDENT-INTERVIEWER INTERACTION

• Standardized interviewing
  – Ideal question-answer process: Respondents interpret question meaning, retrieve requested information, integrate recalled information and question meaning, edit and communicate answer (Tourangeau et al. 2000)
  – Interviewers ensure that respondents thoroughly engage with all of these steps

• Interactional component of survey interviews interferes with ideal question-answer process (Schaeffer & Maynard 2008)
  – Conversational principles may result in suggestive behavior (Ongena & Dijkstra 2006)
  – Rapport, i.e., mutual attentiveness, positivity, and coordination in talk (Lavin & Maynard 2001) may lead interviewers to adjust to individual respondents
• Both standardized and flexible interviewing may lead to high data quality if the question-answer process proceeds without problems (Schober & Conrad, 1997, 2002)
• Both standardized and flexible interviewing may lead to high data quality if the question-answer process proceeds without problems (Schober & Conrad, 1997, 2002)

• Difficult-to-survey populations
  – Individuals with migration background or low education face greater difficulties in answering survey questions
  – Higher level of problematic deviations from an ideal question-answer process (e.g., requests for clarification, answers not in line with answer scales)
Both standardized and flexible interviewing may lead to high data quality if the question-answer process proceeds without problems (Schober & Conrad, 1997, 2002).

Difficult-to-survey populations
- Individuals with migration background or low education face greater difficulties in answering survey questions.
- Higher level of problematic deviations from an ideal question-answer process (e.g., requests for clarification, answers not in line with answer scales).
  - Respondents may be more prone to satisficing and interviewer influence.
  - Interviewers may be more inclined to deviate from standardized interviewing.
MECHANISMS OF INTERVIEWER INFLUENCE

- Moderating influence of interviewer probing
  - Assistance and standardized interviewing may counteract satisficing
  - Interviewers may attune to individual response styles, promoting satisficing
MECHANISMS OF INTERVIEWER INFLUENCE

• Moderating influence of interviewer probing
  – Assistance and standardized interviewing may counteract satisficing
  – Interviewers may attune to individual response styles, promoting satisficing

• Interviewers may provoke response styles

![Diagram showing the relationship between interviewer behavior, respondent satisficing, and non-differentiation behavior. The diagram illustrates how interviewer probing influences respondent behavior, leading to non-differentiation and extreme responding item nonresponse.]
PREVIOUS RESEARCH

• Evidence of interviewer influence on response behavior regarding
  – acquiescence and non-differentiation in attitudinal questions
    (Hox et al. 1991; Loosveldt & Beullens 2017; Olsen & Bingen 2011)
  – straightlining (Vandenplas et al. 2017)
  – “don’t know” and “no opinion” answers (Pickery & Loosveldt 1998, 2004)
PREVIOUS RESEARCH

• Evidence of interviewer influence on response behavior regarding
  – acquiescence and non-differentiation in attitudinal questions
    (Hox et al. 1991; Loosveldt & Beullens 2017; Olsen & Bingen 2011)
  – straightlining (Vandenplas et al. 2017)
  – “don’t know” and “no opinion” answers (Pickery & Loosveldt 1998, 2004)

• Inconclusive findings regarding mechanisms of interviewer influence

• Common conclusion: Direct measures from interview recordings to further investigate interviewer influences
ANALYSIS APPROACH

1. **Multilevel Modeling**
   Identification of exceptional interviewers, with highly positive/negative effects on non-differentiation, extreme responding, and item-nonresponse
1. Multilevel Modeling
Identification of exceptional interviewers, with highly positive/negative effects on non-differentiation, extreme responding, and item-nonresponse

2. Behavior Coding
Coding of audio recorded interviewer-respondent interactions using behavior coding
ANALYSIS APPROACH

1. Multilevel Modeling
   Identification of exceptional interviewers, with highly positive/negative effects on non-differentiation, extreme responding, and item-nonresponse

2. Behavior Coding
   Coding of audio recorded interviewer-respondent interactions using behavior coding

3. Analysis of Behavior Coding Data
   Determine influence of observed behavior on quality indicators
DATA AND ANALYSIS SAMPLE

• The German panel study “Labour Market and Social Security” (PASS)
  – Initially more than 12,000 randomly sampled households (start: 2006)
  – Research on labor market, welfare state, and poverty in Germany
  – Oversampling low-income households
  – Mix of computer-assisted personal (CAPI) and telephone interviews (CATI)

• Analysis sample
  – Data from Wave 13 and CAPI interviews, excluding interviews in foreign language (N=71)
  – Cases with valid answers on dependent and independent variables
  – N=7,427 cases conducted by 251 interviewers
## DATA QUALITY INDICATORS

<table>
<thead>
<tr>
<th>Non-Differentiation</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>STRAIGHT</td>
<td>Identical responses to all items</td>
</tr>
<tr>
<td>MAX STRA</td>
<td>Maximum sequence of identical responses</td>
</tr>
<tr>
<td>REP</td>
<td>Fraction of responses identical to previous one</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation of responses in one item block</td>
</tr>
<tr>
<td>MULL</td>
<td>Average square root of absolute difference between any two items</td>
</tr>
<tr>
<td>AV DEV</td>
<td>Average distance between two subsequent answers</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extreme Responding</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ERS</td>
<td>Fraction of extreme responses within an item block</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item-Nonresponse</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>INR</td>
<td>Fraction of item-nonresponse</td>
</tr>
<tr>
<td>INR DK</td>
<td>Fraction of “Don’t know“</td>
</tr>
<tr>
<td>INR NA</td>
<td>Fraction of “No Answer“</td>
</tr>
</tbody>
</table>
# DATA QUALITY INDICATORS

<table>
<thead>
<tr>
<th>Question</th>
<th>Response Scale / Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain-specific satisfaction:</strong> Health, dwelling, standard of living in general</td>
<td>very dissatisfied (0) – very satisfied (10)</td>
</tr>
<tr>
<td><strong>Big 5</strong></td>
<td>not at all (1) – absolutely (5)</td>
</tr>
<tr>
<td><strong>Trust in institutions:</strong> Political parties, the German Government, the German constitutional court, the police, the newspaper industry</td>
<td>no trust at all (0) – complete trust (10)</td>
</tr>
<tr>
<td><strong>Frequency of leisure activities:</strong> Going out, visits, attending sporting events, cultural events, going on trips or short journeys</td>
<td>daily (1) - once a week (2) - once a month (3) - less often (4) - never (5)</td>
</tr>
</tbody>
</table>
## DATA QUALITY INDICATORS

### Question | Response Scale / Categories
---|---
Interest in politics | very much (1) – not at all (10)

**Trust in institutions**: Political parties, the German Government, the German constitutional court, the police, the newspaper industry | no trust at all (0) – complete trust (10)

Satisfaction regarding democracy in Germany | entirely dissatisfied (0) – entirely satisfied (10)

Political orientation | far left (0) – far right (10)
MULTILEVEL MODELS

Separate model for each indicator

Three-Level Logistic Model (STRAIGHT)

<table>
<thead>
<tr>
<th>Level 1: Item Block</th>
<th>Level 2: Respondent</th>
<th>Level 3: Interviewer</th>
</tr>
</thead>
</table>

Three-Level Linear Model (MAX STRA, REP, ND, MULL, AV DEV, ERS)

<table>
<thead>
<tr>
<th>Level 1: Item Block</th>
<th>Level 2: Respondent</th>
<th>Level 3: Interviewer</th>
</tr>
</thead>
</table>

Two-Level Linear Model (INR, INR DK, INR NA)

<table>
<thead>
<tr>
<th>Level 1: Respondent</th>
<th>Level 2: Interviewer</th>
</tr>
</thead>
</table>
Linear three-level random intercept model

\[ INDICATOR_{bji} = \gamma_{000} + \sum_c \gamma_{0co} Control_{cji} + \theta_0i + \mu_0ji + \varepsilon_{bji} \]
MULTILEVEL MODEL

Linear three-level random intercept model

\[ \text{INDICATOR}_{bji} = \gamma_{000} + \sum_c \gamma_{0co} \text{Control}_{cji} + \theta_{0i} + \mu_{0ji} + \varepsilon_{bji} \]

<table>
<thead>
<tr>
<th>( \gamma_{000} )</th>
<th>Regression intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_{0i} )</td>
<td>Residuals interviewer level</td>
</tr>
<tr>
<td>( \mu_{0ji} )</td>
<td>Residuals respondent level</td>
</tr>
<tr>
<td>( \varepsilon_{bji} )</td>
<td>Residuals block level</td>
</tr>
</tbody>
</table>
**MULTILEVEL MODEL**

Linear three-level random intercept model

\[ \text{INDICATOR}_{bji} = \gamma_{000} + \sum_c \gamma_{0c0} \text{Control}_{cji} + \theta_{0i} + \mu_{0ji} + \varepsilon_{bji} \]

<table>
<thead>
<tr>
<th>Respondent Covariates</th>
<th>Demographic</th>
<th>Gender, age, migration background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social background</td>
<td>Occupational status, employment status, CASMIN classification</td>
<td></td>
</tr>
<tr>
<td>Household characteristics</td>
<td>Household size, number of underage children</td>
<td></td>
</tr>
<tr>
<td>Panel experience</td>
<td>Number of waves</td>
<td></td>
</tr>
<tr>
<td>Area Covariates</td>
<td>Regional size category, federal state</td>
<td></td>
</tr>
</tbody>
</table>
# INTRACLASS CORRELATION (ICC)

<table>
<thead>
<tr>
<th>Model</th>
<th>ICC (Interviewer-Level)</th>
<th>ICC (Respondent-Level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STRAIGHT</td>
<td>4.3% (0.043)</td>
<td>4.7% (0.047)</td>
</tr>
<tr>
<td>MAX STRA</td>
<td>1.3% (0.013)</td>
<td>4.4% (0.044)</td>
</tr>
<tr>
<td>REP</td>
<td>1.2% (0.012)</td>
<td>4.7% (0.047)</td>
</tr>
<tr>
<td>SD</td>
<td>2.3% (0.023)</td>
<td>10.5% (0.105)</td>
</tr>
<tr>
<td>MULL</td>
<td>2.1% (0.021)</td>
<td>6.1% (0.061)</td>
</tr>
<tr>
<td>AV DEV</td>
<td>1.5% (0.015)</td>
<td>7.4% (0.074)</td>
</tr>
<tr>
<td>ERS</td>
<td>3.0% (0.030)</td>
<td>16.5% (0.165)</td>
</tr>
<tr>
<td>INR</td>
<td>9.3% (0.093)</td>
<td>-</td>
</tr>
<tr>
<td>INR DK</td>
<td>9.7% (0.097)</td>
<td>-</td>
</tr>
<tr>
<td>INR NA</td>
<td>8.4% (0.084)</td>
<td>-</td>
</tr>
</tbody>
</table>

How Do Survey Interviewers Impact Respondents' Satisficing Tendencies?
RESULTS MULTILEVEL MODELS

1. Multilevel Modeling

Estimated Random Intercept per Interviewer

Max. sequence of identical responses

Standard deviation of responses to items
Estimated Random Intercept per Interviewer

Max. sequence of identical responses

Standard deviation of responses to items

Sample for behavior coding: “Exceptional” interviewers with low or high estimates
BEHAVIOR CODING

• Sample of audio recordings
  … from interviews with respondent consent to recording (33.5% of F2F interviews W13)
  … of exceptional interviewers according to intercepts from multilevel regression analyses (N=99)
  … for interviewers with ≤ 5 recordings, all recordings
  … for interviewers with > 5 recordings, staggered according to number of recordings (N=558)
BEHAVIOR CODING

• Sample of audio recordings
  … from interviews with respondent consent to recording (33.5% of F2F interviews W13)
  … of exceptional interviewers according to intercepts from multilevel regression analyses (N=99)
  … for interviewers with ≤ 5 recordings, all recordings
  … for interviewers with > 5 recordings, staggered according to number of recordings (N=558)

• Coding of sequential information at the question-answer level (Ongena & Dijkstra 2006)
  – First two exchanges per item/question
    – Interviewer: Presenting answer categories as scripted, with minor, or major changes
    – Respondent: Requests for clarification, inadequate answers, remarks that point to uncertainty
      → Stimulus for interviewer probing and second sequence
  – Control variables: Difficulties regarding question and item presentation, interview distortions, language skills, more than two sequences
I: Our everyday actions are influenced by our beliefs and personal attitudes. I will now name some attitudes a person can have. Presumably some of these attitudes will apply to you, some will not. Please tell me for every attitude if it applies to you or not.

I: You can level your answer as follows: “does not apply at all”, “rather does not apply”, “neither-nor”, “rather applies” or “applies a lot”.

I: I tend to criticize people.

R: Could you repeat the options?

I: “does not apply at all” or “applies a lot”

R: Okay, this doesn’t apply at all
Our everyday actions are influenced by our beliefs and personal attitudes. I will now name some attitudes a person can have. Presumably some of these attitudes will apply to you, some will not. Please tell me for every attitude if it applies to you or not.

You can level your answer as follows: “does not apply at all”, “rather does not apply”, “neither-nor”, “rather applies” or “applies a lot”.

I: I tend to criticize people.

R: Could you repeat the options?

I: “does not apply at all” or “applies a lot”

R: Okay, this doesn’t apply at all
I: Our everyday actions are influenced by our beliefs and personal attitudes. I will now name some attitudes a person can have. Presumably some of these attitudes will apply to you, some will not. Please tell me for every attitude if it applies to you or not.

I: You can level your answer as follows: “does not apply at all”, “rather does not apply”, “neither-nor”, “rather applies” or “applies a lot”.

I: I tend to criticize people.

R: Could you repeat the options?

I: “does not apply at all” or “applies a lot”

R: Okay, this doesn’t apply at all
I: Our everyday actions are influenced by our beliefs and personal attitudes. I will now name some attitudes a person can have. Presumably some of these attitudes will apply to you, some will not. Please tell me for every attitude if it applies to you or not.

I: You can level your answer as follows: “does not apply at all”, “rather does not apply”, “neither-nor”, “rather applies” or “applies a lot”.

I: I tend to criticize people.

R: Could you repeat the options?

I: “does not apply at all” or “applies a lot”

R: Okay, this doesn’t apply at all
ANALYSIS OF BEHAVIOR CODING DATA

1. Effect of initial interviewer behavior

   Initial Interviewer Behavior → Non-differentiation extreme responding item nonresponse

   Sequence 1 → Sequence 2

2. Effect of interviewer probing, based on subsample with second sequence

   Interviewer Probing → Non-differentiation extreme responding item nonresponse

   Sequence 1 → Sequence 2
OUTLOOK AND DISCUSSION

• Next steps
  – Coding interviews and test for reliability of coding
  – Analysis of behavior coding data

• Possible methodological problems
  – Interviewers know whether the interview is recorded and may optimize their behavior
  – Sample of recordings: Fewer recordings among interviewers who deviate from standardization?
CONTACT

Silvia Schwanhäuser
✉️ Silvia.Schwanhaeuser2@iab.de
📞 +49 911 179 2770

Bettina Müller
✉️ Bettina.Mueller2@iab.de
📞 +49 911 179 8485
REFERENCES


APPENDIX
APPENDIX | QUALITY INDICATORS

Boxplot for different measures of non-differentiation
APPENDIX | QUALITY INDICATORS

Mean, standard deviation, and number of observations of unstandardized quality indicators separate for different item blocks

<table>
<thead>
<tr>
<th>Non-Differentiation</th>
<th>Attitudes</th>
<th>Pol. Participation</th>
<th>Leisure activities</th>
<th>Big5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>STRAIGHT</td>
<td>0.10</td>
<td>0.30</td>
<td>7,427</td>
<td>0.03</td>
</tr>
<tr>
<td>MAX STRA</td>
<td>0.57</td>
<td>0.67</td>
<td>7,427</td>
<td>2.28</td>
</tr>
<tr>
<td>REP</td>
<td>0.28</td>
<td>0.34</td>
<td>7,427</td>
<td>0.27</td>
</tr>
<tr>
<td>ND</td>
<td>1.35</td>
<td>1.03</td>
<td>7,427</td>
<td>1.21</td>
</tr>
<tr>
<td>MULL</td>
<td>1.05</td>
<td>0.55</td>
<td>7,427</td>
<td>0.94</td>
</tr>
<tr>
<td>AV DEV</td>
<td>1.70</td>
<td>1.47</td>
<td>7,427</td>
<td>1.42</td>
</tr>
<tr>
<td>INR</td>
<td></td>
<td>-</td>
<td>-</td>
<td>1.52</td>
</tr>
<tr>
<td>INR DK</td>
<td></td>
<td>-</td>
<td>-</td>
<td>1.02</td>
</tr>
<tr>
<td>INR NA</td>
<td></td>
<td>-</td>
<td>-</td>
<td>0.49</td>
</tr>
<tr>
<td>ERS</td>
<td>12.87</td>
<td>23.75</td>
<td>7,427</td>
<td>25.88</td>
</tr>
</tbody>
</table>

How Do Survey Interviewers Impact Respondents’ Satisficing Tendencies? // Page 25
APPENDIX | MULTILEVEL MODEL

Linear two-level random intercept model

\[
INDICATOR_{ji} = \gamma_{00} + \sum_c \gamma_{c0} Control_{cji} + \mu_{0i} + \epsilon_{ji}
\]

Logistic three-level random intercept model

\[
\text{Logit}(INDICATOR_{bji}) = \gamma_{000} + \sum_c \gamma_{0c0} Control_{cji} + \vartheta_{0i} + \mu_{0ji} + \epsilon_{bji}
\]