INSTITUTE FOR EMPLOYMENT RESEARCH
The Research Institute of the Federal Employment Agency

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

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

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

RESPONDENT-INTERVIEWER INTERACTION

- Both standardized and flexible interviewing may lead to high data quality if the question-answer process proceeds without problems (Schober & Conrad, 1997, 2002)

RESPONDENT-INTERVIEWER INTERACTION

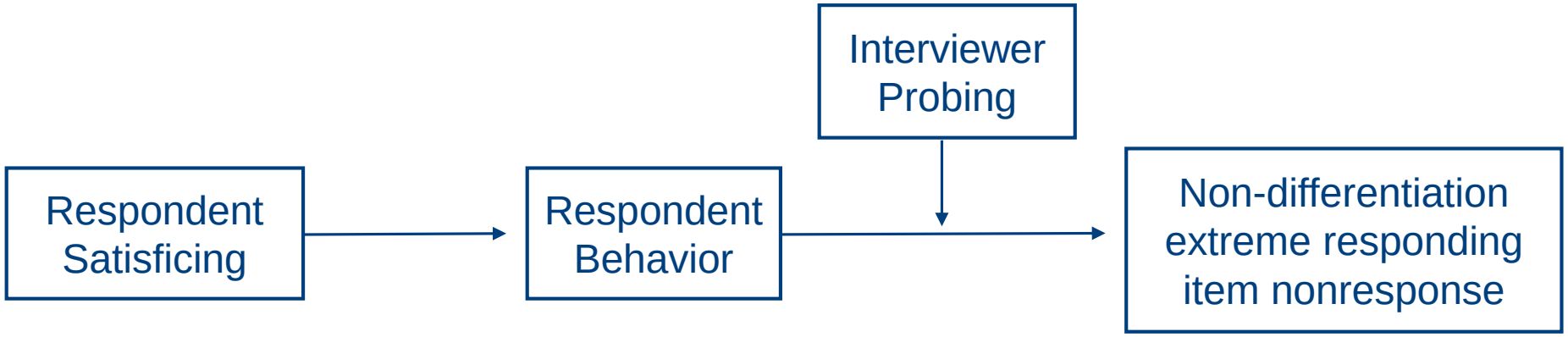
- Both standardized and flexible interviewing may lead to high data quality if the question-answer process proceeds without problems (Schober & Conrad, 1997, 2002)
- Difficult-to-survey populations
  - Individuals with migration background or low education face greater difficulties in answering survey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

RESPONDENT-INTERVIEWER INTERACTION

- 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

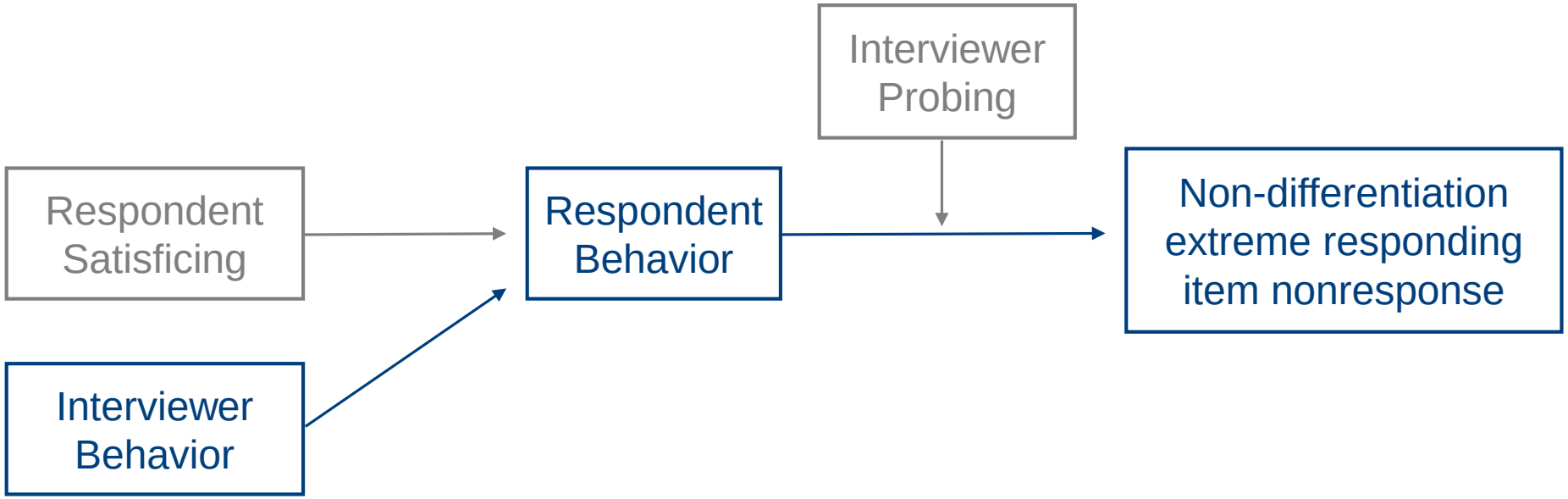
MECHANISMS OF INTERVIEWER INFLUENCE

- 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



PREVIOUS RESEARCH

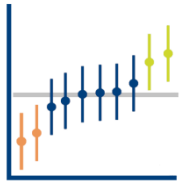
- Evidence of interviewer influence on response behavior regarding
  - acquiescence and non-differentiation in attitudinal questions (Hox et al. 1991; Loosveldt & Beullens 2017; Olsen & Bingen 2011)
  - straightlining (Vandenplas et al. 2017)
  - “don’t know“ and “no opinion” answers (Pickery & Loosveldt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

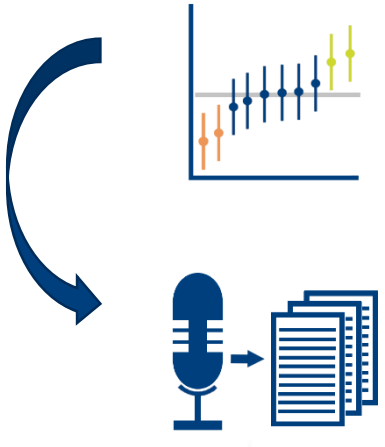
- Inconclusive findings regarding mechanisms of interviewer influence
- Common conclusion: Direct measures from interview recordings to further investigate interviewer influences

ANALYSIS APPROACH



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

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

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

Non-Differentiation		Values
STRAIGHT	Identical responses to all items	0, 1
MAX STRA	Maximum sequence of identical responses	0 – Max
REP	Fraction of responses identical to previous one	0 – 1
SD	Standard deviation of responses in one item block	0 – Max
MULL	Average square root of absolute difference between any two items	0 – Max
AV DEV	Average distance between two subsequent answers	0 – Max
Extreme Responding		
ERS	Fraction of extreme responses within an item block	0 – 100
Item-Nonresponse		
INR	Fraction of item-nonresponse	0 – 100
INR DK	Fraction of „Don‘t know“	0 – 100
INR NA	Fraction of “No Answer”	0 – 100

DATA QUALITY INDICATORS

	Question	Response Scale / Categories
Non-Differentiation Extreme Responding	**Domain-specific satisfaction**: Health, dwelling, standard of living in general	very dissatisfied (0) – very satisfied (10)
	Big 5	not at all (1) – absolutely (5)
	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)
	Frequency of leisure activities: Going out, visits, attending sporting events, cultural events, going on trips or short journeys	daily (1) - once a week (2) - once a month (3) - less often (4) - never (5)

DATA QUALITY INDICATORS

	Question	Response Scale / Categories
Item-Nonresponse	**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)

Level 1: Item Block	Level 2: Respondent	Level 3: Interviewer

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

Level 1: Item Block	Level 2: Respondent	Level 3: Interviewer

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

Level 1: Respondent	Level 2: Interviewer

MULTILEVEL MODEL

Linear three-level random intercept model

$$INDICATOR_{bji} = \gamma_{\ 000} + \sum_c \gamma_{0c0} Control_{cji} + \vartheta_{0i} + \mu_{0ji} + \varepsilon_{bji}$$

MULTILEVEL MODEL

Linear three-level random intercept model

$$INDICATOR_{bji} = \gamma_{000} + \sum_c \gamma_{0c0} Control_{cji} + \vartheta_{0i} + \mu_{0ji} + \varepsilon_{bji}$$

γ_{000}	Regression intercept
ϑ_{0i}	Residuals interviewer level
μ_{0ji}	Residuals respondent level
ε_{bji}	Residuals block level

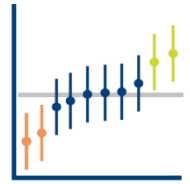
MULTILEVEL MODEL

Linear three-level random intercept model

$$INDICATOR_{bji} = \gamma_{000} + \sum_c \gamma_{0c0} Control_{cji} + \vartheta_{0i} + \mu_{0ji} + \varepsilon_{bji}$$

Respondent Covariates	Demographic	Gender, age, migration background
	Social background	Occupational status, employment status, CASMIN classification
	Household characteristics	Household size, number of underage children
	Panel experience	Number of waves
Area Covariates		Regional size category, federal state

INTRACLASS CORRELATION (ICC)

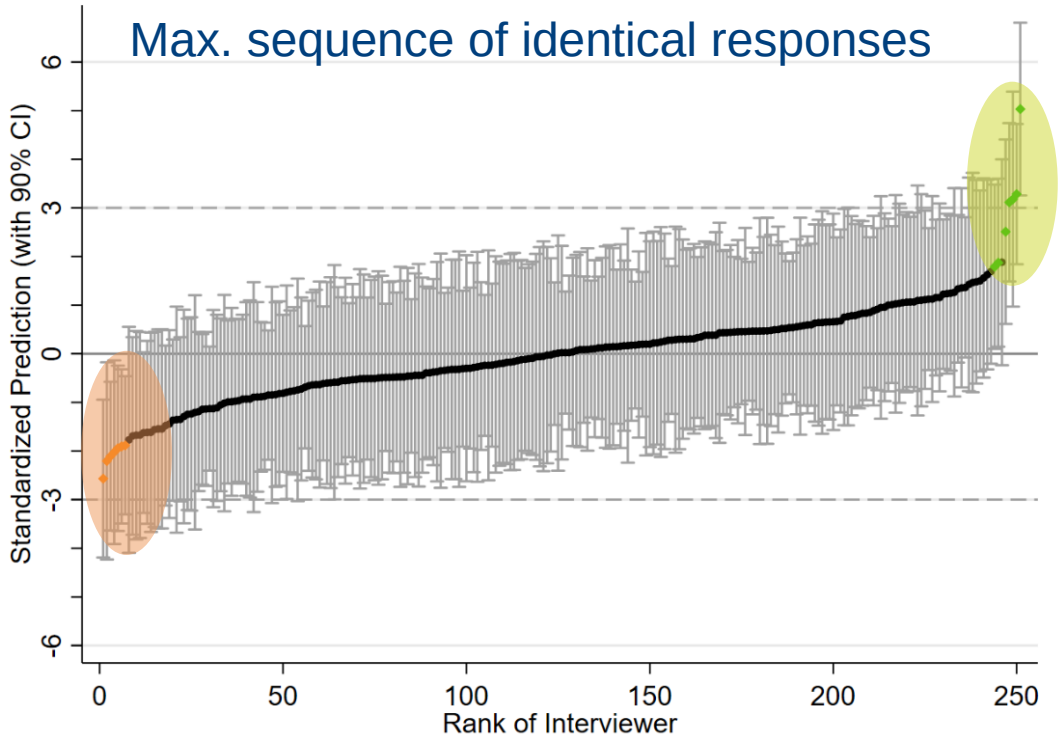


Model	ICC (Interviewer-Level)	ICC (Respondent-Level)
STRAIGHT	4.3 % (0.043)	4.7 % (0.047)
MAX STRA	1.3 % (0.013)	4.4 % (0.044)
REP	1.2 % (0.012)	4.7 % (0.047)
SD	2.3% (0.023)	10.5% (0.105)
MULL	2.1% (0.021)	6.1% (0.061)
AV DEV	1.5 % (0.015)	7.4% (0.074)
ERS	3.0% (0.030)	16.5% (0.165)
INR	9.3% (0.093)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

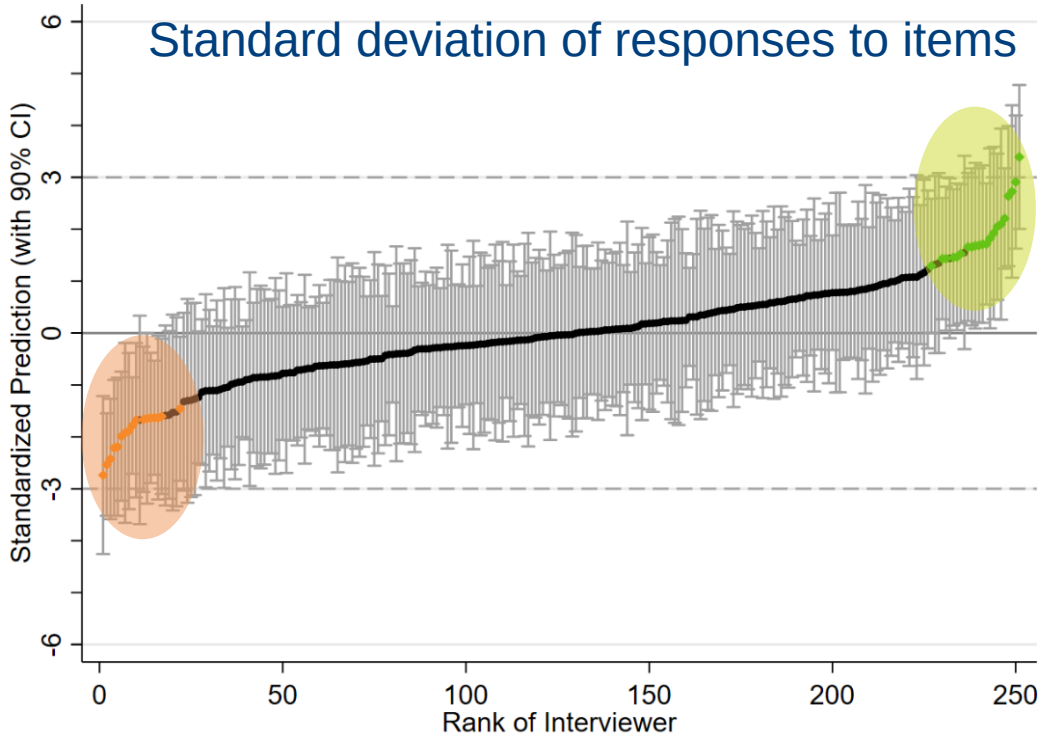
RESULTS MULTILEVEL MODELS

Estimated Random Intercept per Interviewer

Max. sequence of identical responses



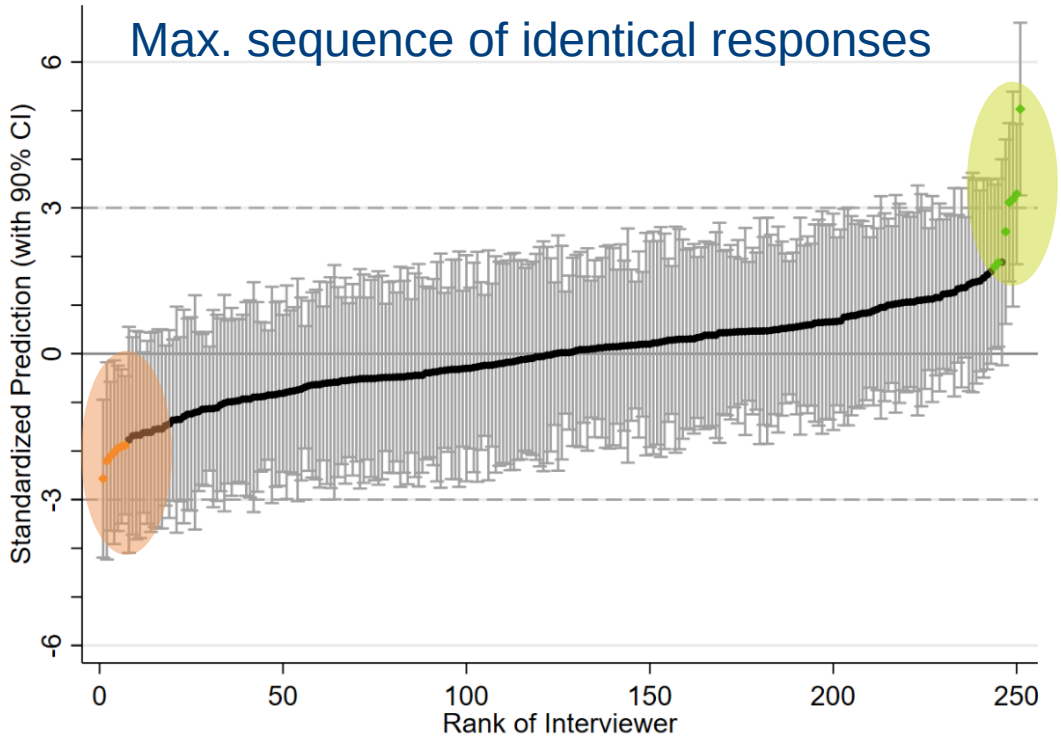
Standard deviation of responses to items



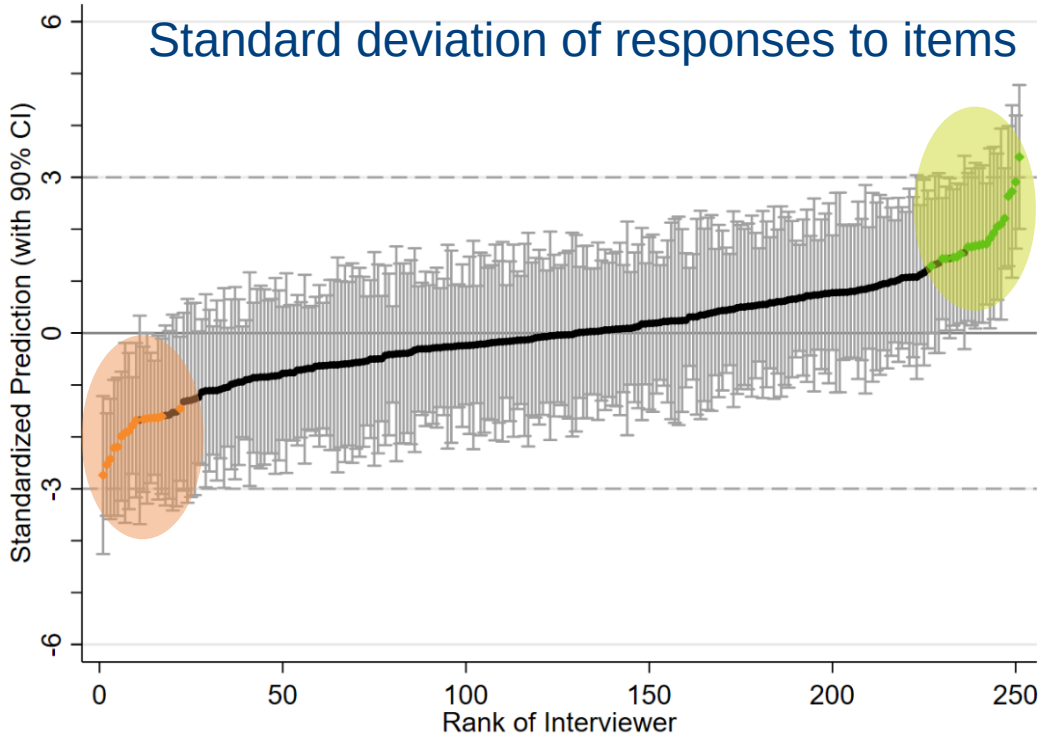
RESULTS MULTILEVEL MODELS

Estimated Random Intercept per Interviewer

Max. sequence of identical responses

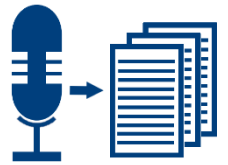


Standard deviation of responses to items



⇨ Sample for behavior coding: “Exceptional” interviewers with **low** or **high** estimates

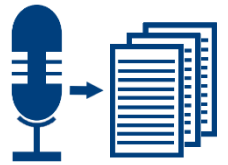
BEHAVIOR CODING



2. 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



2. 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

BEHAVIOR CODING

Sequence 1

Sequence 2

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”.

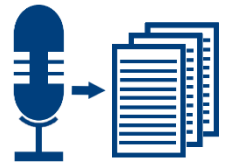
I: I tend to criticize people.

R: Could you repeat the options?

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

R: Okay, this doesn’t apply at all

BEHAVIOR CODING



Sequence 1

Sequence 2

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:** Presents response categories as scripted
- ❑ **I:** Only presents tendency: rather (doesn‘t) apply
- ❑ **I:** Only presents extreme categories
- ❑ **I:** Omits middle categorie
- ❑ **I:** Rewords response categories
- ❑ **I:** Presents response categories with irrelevant changes (e.g., flounders)

BEHAVIOR CODING

Sequence 1

Sequence 2

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.

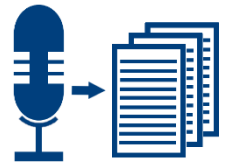
R: Could you repeat the options?

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

R: Okay, this doesn’t apply at all

- ❑ **R:** No difficulties in answering
- ❑ **R:** Request to repeat response categories
- ❑ **R:** Provides mismatch answer
- ❑ **R:** Provides answer that cannot be clearly assigned reponse categories
- ❑ **R:** Answers „dont‘ know“
- ❑ **R:** Refuses to answer
- ❑ **R:** Remark that points to uncertainty

BEHAVIOR CODING



Sequence 1

Sequence 2

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?

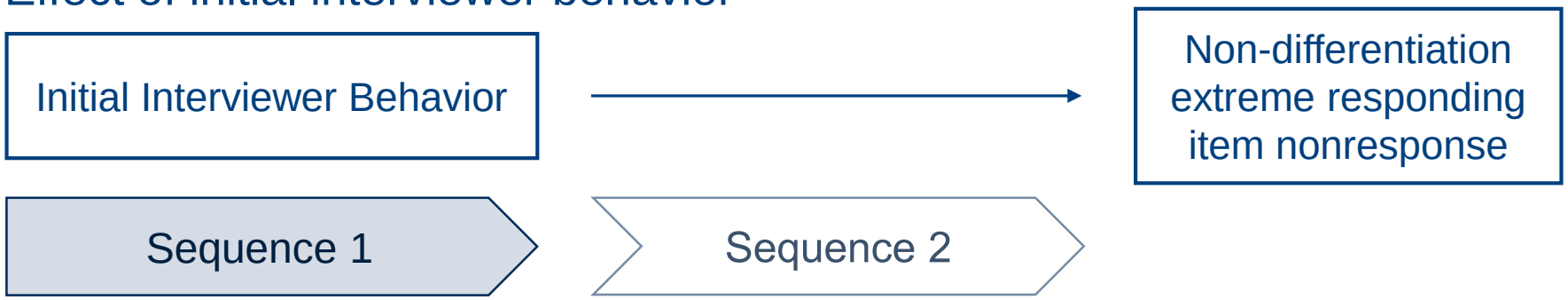
I: “does not apply at all” or “applies a lot”

R: Okay, this doesn’t apply at all

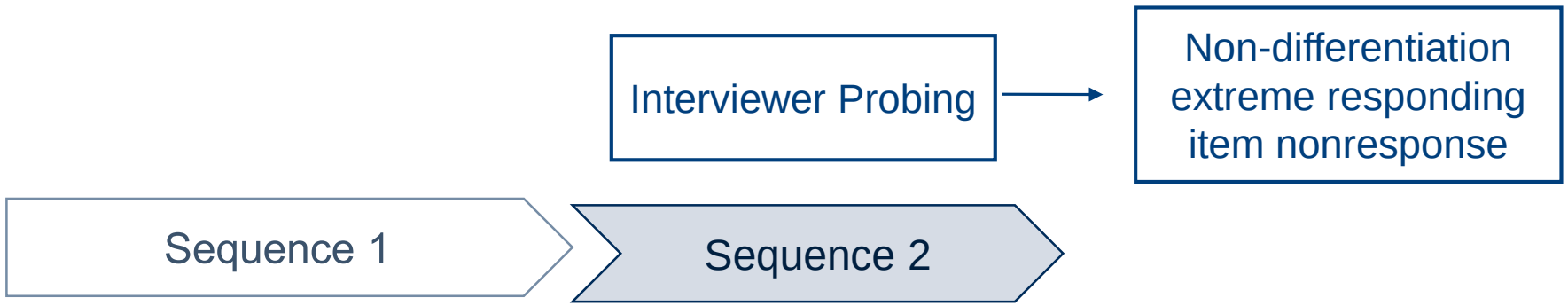
- ❑ I: Presents response categories as scripted
- ❑ I: Only presents tendency: rather (doesn‘t) apply
- ❑ I: Only presents extreme categories
- ❑ I: Omits middle categorie
- ❑ I: Rewords response categories
- ❑ I: Presents response categories with irrelevant changes (e.g., flounders)

ANALYSIS OF BEHAVIOR CODING DATA

1. Effect of initial interviewer behavior



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



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

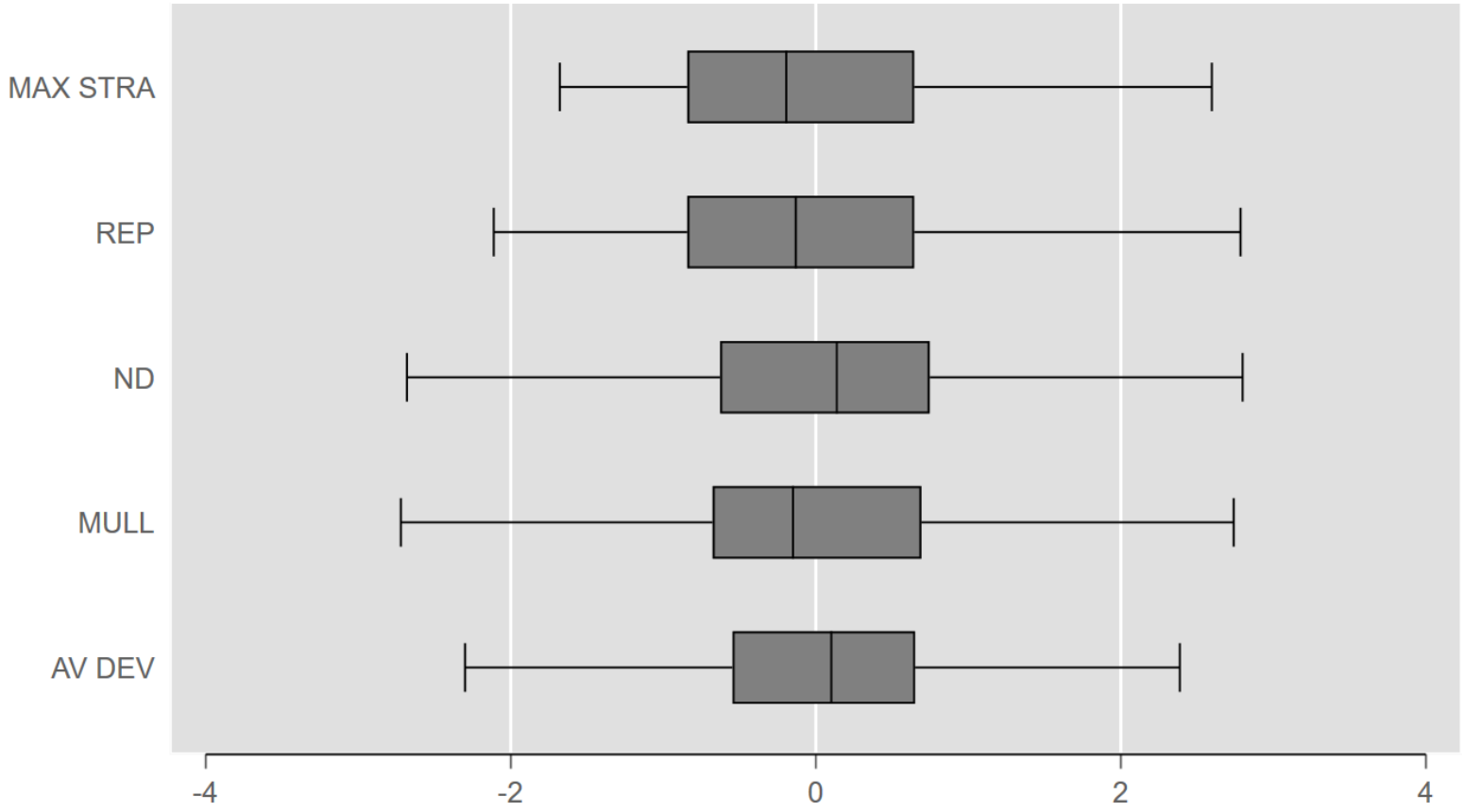
REFERENCES

- Lavin, D. & Maynard, D.W. (2001). Standardization vs. rapport: Respondent laughter and interviewer reaction during telephone surveys. *American Sociological Review*, 66(3), 453-479.
- Ongena, Y. & Dijkstra, W. (2006). Methods of behavior coding of survey interviews. *Journal of Official Statistics*, 22 (3), 419–451.
- Schaeffer, N.C. & Maynard, D.W. 2008. The contemporary standardized survey interview for social research. In F.G. Conrad & M.F. Schober (Eds.), *Envisioning the survey interview of the future* (pp.31-57), Hoboken, NJ: John Wiley & Sons, Inc.
- Schober, M.F. & Conrad, F.G. (1997). Does Conversational Interviewing Reduce Survey Measurement Error? *Public Opinion Quarterly,* 61(4), 576-602.
- Tourangeau, R. & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin,* 133(5), 859–883.

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

		Attitudes			Pol. Participation			Leisure activities			Big5		
		Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
Non-Differentiation	STRAIGHT	0.10	0.30	7,427	0.03	0.18	7,296	0.01	0.12	7,426	-	-	-
	MAX STRA	0.57	0.67	7,427	2.28	1.43	7,409	1.11	0.88	7,296	1.33	0.79	7,426
	REP	0.28	0.34	7,427	0.27	0.13	7,427	0.31	0.25	7,427	0.35	0.20	7,427
	ND	1.35	1.03	7,427	1.21	0.27	7,409	1.71	0.95	7,296	1.22	0.35	7,426
	MULL	1.05	0.55	7,427	0.94	0.15	7,409	1.16	0.44	7,296	0.98	0.24	7,426
	AV DEV	1.70	1.47	7,427	1.42	0.41	7,409	1.60	1.01	7,296	1.19	0.42	7,426
INR	INR	-	-	-	1.52	6.35	7,427	-	-	-	-	-	-
	INR DK	-	-	-	1.02	4.79	7,427	-	-	-	-	-	-
	INR NA	-	-	-	0.49	4.05	7,427	-	-	-	-	-	-
	ERS	12.87	23.75	7,427	25.88	20.31	7,427	12.88	23.67	7,427	34.98	23.79	7,427

APPENDIX | MULTILEVEL MODEL

Linear two-level random intercept model

$$INDICATOR_{ji} = \gamma_{00} + \sum_c \gamma_{c0} Control_{cji} + \mu_{0i} + \varepsilon_{ji}$$

Logistic three-level random intercept model

$$Logit(INDICATOR_{bji}) = \gamma_{000} + \sum_c \gamma_{0c0} Control_{cji} + \vartheta_{0i} + \mu_{0ji} + \varepsilon_{bij}$$