

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

2019 Workshop: Interviewers and Their Effects
from a Total Survey Error Perspective

Sociology, Department of

2020

Chapter 12: Differences in interaction quantity and conversational flow in CAPI and CATI interviews. Appendix 12

Yfke Ongena

Marieke Haan

Follow this and additional works at: <https://digitalcommons.unl.edu/sociw>



Part of the Quantitative, Qualitative, Comparative, and Historical Methodologies Commons

This Article is brought to you for free and open access by the Sociology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in 2019 Workshop: Interviewers and Their Effects from a Total Survey Error Perspective by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Chapter 12: Differences in interaction quantity and conversational flow in CAPI and CATI
interviews

Appendix 12

Yfke Ongena and Marieke Haan

Full book reference:

Olson, K., J. D. Smyth, J. Dykema, A. L. Holbrook, F. Kreuter, B. T. West. 2020. *Interviewer Effects from a Total Survey Error Perspective*. Boca Raton: CRC Press.

Appendix 12A Differences in Length of Interviewer-Respondent Interactions in CAPI and CATI Interviews

Model Description

Following notation by Beretvas (2010), the base model predicts the log transformed of number of words uttered ($Y_{i(j1, j2)k}$) as a function of an overall mean (γ_{0000}), a random effect due to the respondent (u_{0j10k}), a random effect due to the question (u_{00j2k}), a random effect due to the interviewing mode (v_{000k}), and a residual term ($e_{i(j1, j2)k}$), where u_{0j10k} , u_{00j2k} and v_{000k} are normally distributed with mean zero and variance τ_{uj10} , τ_{uj2} , and τ_{uk} , respectively, and $e_{i(j1, j2)k}$ is normally distributed with mean zero and variance σ_e^2 (Beretvas 2010, p.330):

$$Y_{i(j1, j2)k} = \gamma_{0000} + v_{000k} + u_{0j10k} + u_{00j2k} + e_{i(j1, j2)k}$$

The base model was used for obtaining intraclass correlation coefficients, for the variance due to respondents calculated as:

$$\rho_{resp} = \frac{\hat{\tau}_{uj10}}{\hat{\tau}_{uj10} + \hat{\tau}_{uj2} + \hat{\tau}_{uk} + \hat{\sigma}_e^2}$$

Intraclass correlation coefficients for the variance due to questions and mode modified with the appropriate variance due to questions and mode in the numerator (Raudenbush and Bryk 2002, cited in Olson and Smyth 2015).

We expanded the base model to include question and respondent characteristics:

$$Y_{i(j1, j2)k} = \gamma_{0000} + \sum_{m=1}^p \beta_m \text{Respondent_char}_{j10} + \sum_{s=1}^q \beta_s \text{Question_char}_{j10} + v_{000k} + u_{0j10k} + u_{00j2k} + e_{i(j1, j2)k}$$

For occurrence of filled pauses and respondent uncertainty, we fitted a mixed effects logistic regression model, again taking the variance at three levels into account. Adjusting a mixed effects regression model to a logistic regression entails embedding of the stochastic linear predictor:

$$\eta = Y_{i(j_1, j_2)k} = \gamma_{0000} + \sum_{m=1}^p \beta_m \text{Respondent_char}_{j_{10}} + \sum_{s=1}^q \beta_s \text{Question_char}_{j_{10}} + v_{000k} + u_{0j_{10}k} + u_{00j_{2}k} + e_{i(j_1, j_2)k}$$

in the binomial error function:

$$\mu = \frac{e^{Y_{i(j_1, j_2)k} = \gamma_{0000} + \sum_{m=1}^p \beta_m \text{Respondent_char}_{j_{10}} + \sum_{s=1}^q \beta_s \text{Question_char}_{j_{10}} + v_{000k} + u_{0j_{10}k} + u_{00j_{2}k} + e_{i(j_1, j_2)k}}}{1 + e^{Y_{i(j_1, j_2)k} = \gamma_{0000} + \sum_{m=1}^p \beta_m \text{Respondent_char}_{j_{10}} + \sum_{s=1}^q \beta_s \text{Question_char}_{j_{10}} + v_{000k} + u_{0j_{10}k} + u_{00j_{2}k} + e_{i(j_1, j_2)k}}}}$$

All models are estimated using restricted maximum likelihood in the R-package lme4 (Bates, Mächler, Bolker and Walker 2015).

References

- Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of statistical software* 67(1):1-48.
- Beretvas, S. N. 2010. Cross-classified and multiple-membership models. In *Handbook of advanced multilevel analysis*, ed J. J. Hox and J. K. Roberts, 313-334. New York: Routledge.
- Olson, K. and J. D. Smyth. 2015. The effect of CATI questions, respondents, and interviewers on response time. *Journal of survey statistics and methodology* 3:1-36.
- Raudenbush, S. W. and A. S. Bryk. 2002. *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Table A12A.1 Response Rates (AAPOR RR1)

Contact Mode	Face-to-Face	Telephone	Telephone
Mode choice	No random assignment: CAPI or Web	No random assignment: CATI or Web	Random assignment: CAPI, CATI or Web
AAPOR RR1	54.9	40.6	28.8

Table A12A.2 Descriptive Statistics for Questions and Respondent Characteristics

	n	%/ mean overall	SD	%/mean CAPI (n=53)	%/mean CATI (n=55)
Trimmed number of turns per QA sequence	6629	3.23	(1.99)	3.42	3.06
Trimmed number of events per QA sequence	6629	3.80	(3.01)	4.31	3.36
Number of words per QA sequence	6629	36.71	(31.82)	35.57	35.81
Log Number of words per QA sequence	6629	3.32	(0.75)	3.28	3.36
Interviewer filled pauses	6629	16.93%		27.67%	7.92%
Respondent filled pauses	6629	23.81%		24.49%	23.24%
Respondent uncertainty markers	6629	9.15%		9.07%	9.22%
Necessary question features					
Question length (#words in question)	81	21.19	(14.46)	19.61	22.59
# Response options	81	5.93	(3.82)	-	-
Type of Question					
Behavior (ref)	81	29.63%		-	-
Attitude/opinion	81	44.44%		-	-
Demographic	81	25.93%		-	-
Format of response options					
Open-ended text (ref)	81	9.88%		-	-
Open-ended numeric	81	6.17%		-	-
Closed numeric	81	23.46%		-	-
Closed-nominal	81	4.94%		-	-
Closed-ordinal	81	30.86%		-	-
Yes/No	81	24.69%		-	-
Respondent task complexity					
Question reading level	81	8.79	(3.51)	8.76	8.81
Mismatch Q and response options	81	1.39%		-	-
Sensitive question	81	48.61%		-	-
Unknown terms	81	8.33%		-	-
Showcard question	81	56.94%		-	-
Respondent characteristics					

Education HS degree or less	108	69.47%		77.36%	61.82%
Age	108	39.27	5.31	38.01	40.71
Interviewers	27				

Table A12A.3 Means and Correlations Trimmed number of Turns and Events in CAPI and CATI by Question Characteristics

	Turns CAPI	Turns CATI	W	Events CAPI	Events CATI	W
Mode	3.42	3.06	4474207 ***	4.31	3.36	4581698 ***
Necessary question features						
Question length (correlation)	0.22	0.10		0.25	0.06	
# Response options (correlation)	0.09	0.07		0.09	0.05	
Type of Question						
Behavior	3.38	2.94	387525 ***	4.24	3.33	377742 ***
Attitude	3.50	3.24	1315410 ***	4.52	3.46	1244041 ***
Demographic	3.25	2.71	120808 ***	3.73	3.05	125252 ***
Format of response options						
Open-ended text (ref)	4.54	3.50	1720 **	5.61	4.05	1648.5 ***
Open-ended numeric	6.81	3.52	5.5	8.14	4.01	5
Closed numeric	3.49	3.19	384668 ***	4.58	3.59	373032 ***
Closed-nominal	4.64	3.69	4525.5 ***	5.75	4.26	3935 ***
Closed-ordinal	3.66	3.32	534418 ***	4.65	3.28	484761 ***
Yes/No	2.64	2.60	442019	3.10	3.13	468633
Respondent task complexity						
Question reading level (correlation)	-0.03	-0.02		-0.04	0.04	
Mismatch Q and response options	3.62	3.65	1054.5	3.83	4.40	1276.5
Sensitive question	3.71	3.23	1174388 ***	4.79	3.38	1070122 ***
Unknown terms	3.63	3.16	50254 ***	4.77	4.20	55512 *
Showcard question	3.63	3.31	1651708 ***	4.66	3.34	1500480 ***

Note n = 108 respondents, 81 questions, 2 modes, total n = 6629

* p<0.05, ** p <0.01, ***p<0.001

Table A12A.4 Means and Correlations Trimmed number of Turns and Events by Question Characteristics

	Turns	W	Events	W
Necessary question features				
Question length (correlation)	0.105	-	0.163	-
# Response options (correlation)	0.108	-	0.093	-
Type of Question				
Behavior	3.13	4434939 *	3.74	4354266
Attitude	3.36	4738846 ***	3.96	4535382 ***
Demographic (ref)	2.97	3090242 ***	3.38	3376417 ***
Format of response options				
Open-ended text	3.94	348522 ***	4.71	387382 **
Open-ended numeric	6.44	21672 ***	7.63	21624 ***
Closed numeric	3.33	4099525 **	4.05	3827736 ***
Closed-nominal	4.38	492326 ***	5.34	523619 ***
Closed-ordinal	3.49	4129617 ***	3.94	4236090 ***
Yes/No	2.62	5371604 ***	3.12	5443858 ***
Respondent task complexity				
Question reading level (correlation)	-0.03	-	-0.00	-
Mismatch Q and response options	3.63	230784 ***	4.16	266580 *
Sensitive question	3.46	4414946 ***	4.05	4261374 ***
Unknown terms	3.37	1862680 ***	4.46	1695909 ***
Showcard question	3.46	4237064 ***	3.98	4289800 ***

* p<0.05, ** p <0.01, ***p<0.001