University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

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

Sociology, Department of

2-26-2019

What Do Interviewers Learn? Changes in Interview Length and Interviewer Behaviors Over the Field Period

Kristen Olson
University of Nebraska-Lincoln, Department of Sociology, kolson5@unl.edu

Jolene Smyth
University of Nebraska-Lincoln, Department of Sociology, jsmyth2@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/sociw

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

What do interviewers learn?
An examination of interview length and interviewer behaviors

Kristen Olson and Jolene Smyth

University of Nebraska-Lincoln

Interviewers and their Effects in a Total Survey Error Framework Workshop

Lincoln, NE

February 2019

1

Acknowledgments

Interviewers set the stage for respondents

- Interviewers are important actors in telephone surveys
- By setting the pace for an interview, interviewers communicate the amount of time and cognitive effort respondents should put into their task
- But interviewers vary widely in the time they spend administering a survey
  - And it changes over the course of the data collection period as the interviewer gains within-study experience (e.g., Olson and Peytchev 2007; Olson and Bilgen 2011)
  - In particular, they speed up.
- We don't know what leads to these differences in speed of administering a questionnaire.

Three hypotheses

- Certain interviewer behaviors are omitted or shortened over the course of a field period.
  - Standardized "good" behaviors go away (Ongena and Dijkstra 2007; Olson and Bilgen 2011; Tarnai and Moore 2008; van der Zouwen, Dijkstra and Smith 1991)
- Interviewers may not change the prevalence of individual "good" behaviors, but become more efficient in them or eliminating extraneous behaviors (Olson and Peytchev 2007; Cleary, Mechanic and Weiss 1981; Houtkoop-Steenstra 1997).
  - Not directly trained, but happens over the course of interviews.
- Increased use of bad behaviors that shortcut time (van der Zouwen et al. 1991)
  - Nonstandardized bad behaviors that always happen.

Kirchner and Olson (2017, JSSAM)

- What explains interview length?
  - Interviewer Experience
    - Learning, overall experience, and interviewer cooperation rate
  - Response propensity
    - Composition: Respondent gender, age, education, race, employment status, income, HH size, parent, volunteer status
    - Contactability and cooperation: Item NR rate, ever refusal, complete at first contact, # of call attempts, time of day interview completed
    - Interaction between R and I: Word count of interview

- But there is much more to the interaction between the R and I than just the number of words that they speak

This paper

- What interviewer behaviors change over the course of the data collection period in two telephone surveys?

- Do these behaviors account for changes in survey length over the course of the data collection period?

Data – Building off Kirchner and Olson (2017)

- Work and Leisure Today 1 Survey
  - Landline RDD CATI survey
  - Conducted by AbtSRBI between July 31 and August 28, 2013
  - N=450, AAPOR RR3=6.3%
  - Questionnaire deliberately designed to have highly problematic questions
  - Data deposited at ICPSR; under review

- Work and Leisure Today 2 Survey
  - Dual Frame RDD CATI survey
  - Conducted by AbtSRBI during September 2015
  - n=902, Landline = 451, AAPOR RR3=9.4%; Cell phone = 451, AAPOR RR3=7.1%
  - Two versions – alternative experimental questionnaire designs
  - Questionnaire deliberately avoided these highly problematic questions

Question text: How much do you enjoy cooking? Not at all, A little Somewhat, A lot, or Completely?

Transcripts	Actor	Initial	Assessment	Details	# seconds
I: And how much do you enjoy cooking? Not at all, a little, somewhat, a lot, or completely?	Interviewer	iQuestion Asked	Read exact		4.7
R: Um, how, what? I didn't catch--.	Respondent	rClarification	Asks for repeat of question		4.7
I: How much do you enjoy cooking?	Interviewer	iProbes	Repeat part of Q exact		1.3
R: Cooking? I love to.	Respondent	rAnswer Provided	Uncodable answer	rElaborates no implied	1.4
I: Okay.	Interviewer	Feedback	Affirmation		0.8
R: That's, that's my favorite hobby.	Respondent	Feedback	Personal disclosure		1.3
I: Okay, so a lot or completely?	Interviewer	iProbes	Asks for explicit response	Probe directively, no mismatch	1.3
R: Uh, I'd say a lot. I'm thinking about going to culinary school.	Respondent	rAnswer Provided	Adequate answer	rAdequate w elaboration	3.4
I: Oh, good for you.	Interviewer	Feedback	Short acknowledgement		1

Behavior Codes

- 8 fields coded by trained undergraduate coders
  - 10% subsample of interviews coded by two master coders

	Actor	Initial Action	Assessment of Initial Action	Details of Action	Parentheses	Laughter	Disfluencies	Interruptions
WLT1	κ=0.998	0.90	0.55 to 0.68	0.10 to 0.77	0.92	0.96	0.87	0.94
WLT2	κ=0.998	0.93	0.36 to 0.76	0.24 to 0.83	0.95	0.97	0.83	0.93

Creating behavior measures

- Two ways of examining measures of behaviors
  - Conversational turn level – Total number of conversational turns on which a behavior occurred
    - This is a measure of how much conversation occurred due to this behavior
    - Some questions can have multiple turns with the same kind of behavior (e.g., multiple probing turns)

  - Question level – Total number of questions on which a behavior occurred
    - This is a measure of how spread out across the questionnaire each behavior was

  - Obviously highly correlated
    - Focus on questions in this presentation. Results are similar for conversational turns.

Question text: How much do you enjoy cooking? Not at all, A little Somewhat, A lot, or Completely?

Question level =
Adequate feedback = 1

Turn level =
Adequate feedback = 2

| Initial	Assessment	Details	# seconds
iQuestion Asked	Read exact		4.7
rClarification	Asks for repeat of question		4.7
iProbes	Repeat part of Q exact		1.3
rAnswer Provided	Uncodable answer	rElaborates no implied	1.4
Feedback	Affirmation		0.8
Feedback	Personal disclosure		1.3
iProbes	Asks for explicit response	Probe directively, no mismatch	1.3
rAnswer Provided	Adequate answer	rAdequate w elaboration	3.4
Feedback	Short acknowledgement		1

Dependent Variables

- Interview length in minutes, trimmed at 1st and 99th percentiles
  - WLT1: 12.65 minutes
  - WLT2: 13.36 minutes

- Interviewer behaviors
  - Standardized "good" behaviors
    - Exact question reading; Nondirective probes; Exact verification; Appropriate clarification; Appropriate feedback
  - Efficiency behaviors
    - Stuttering during question reading; Disfluencies; Pleasant talk; Task-related feedback; Laughter
  - Nonstandardized "bad" behaviors
    - Minor changes in question wording; Major changes in question wording; Directive probes; Inadequate verification (paraphrasing); Interruptions
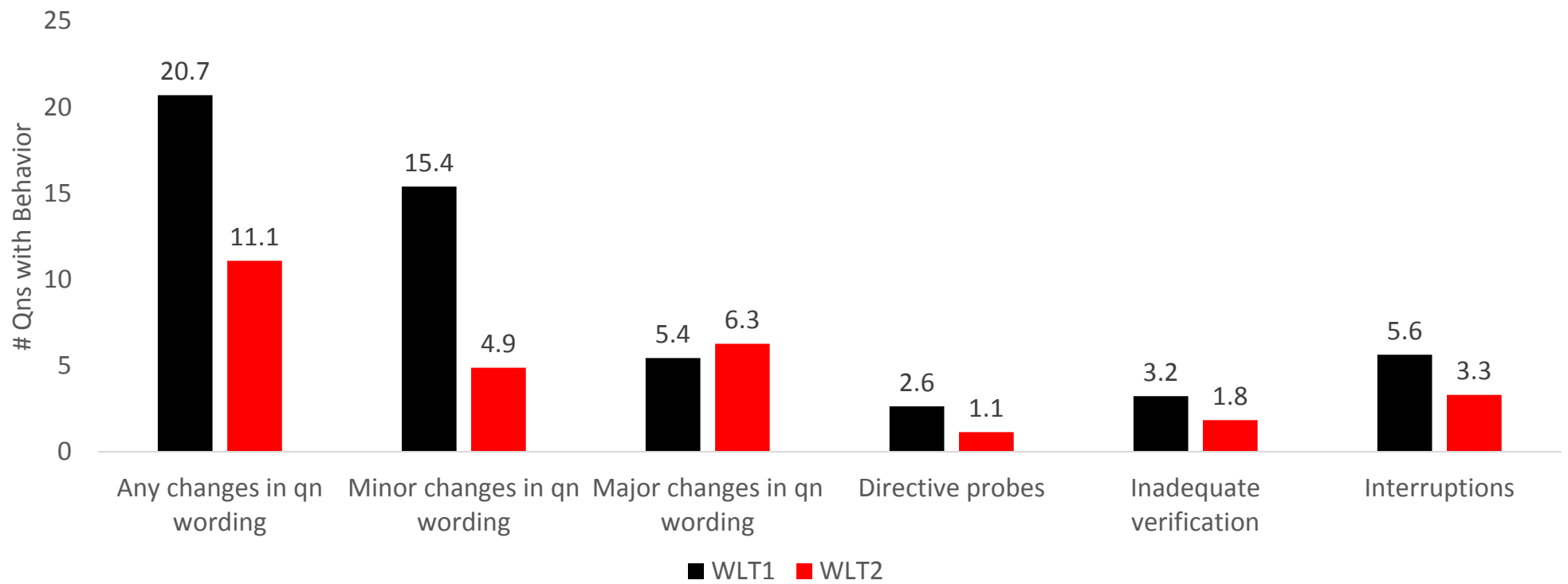
Standardized behaviors
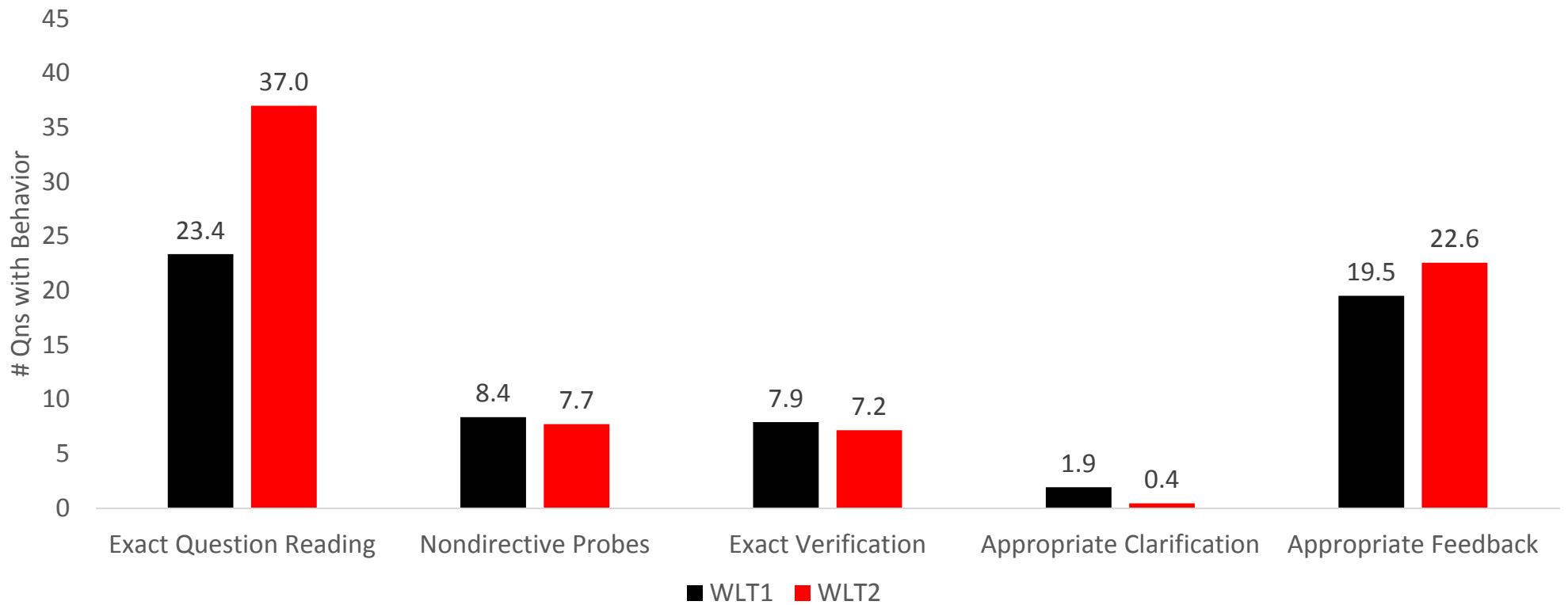
Efficiency behaviors
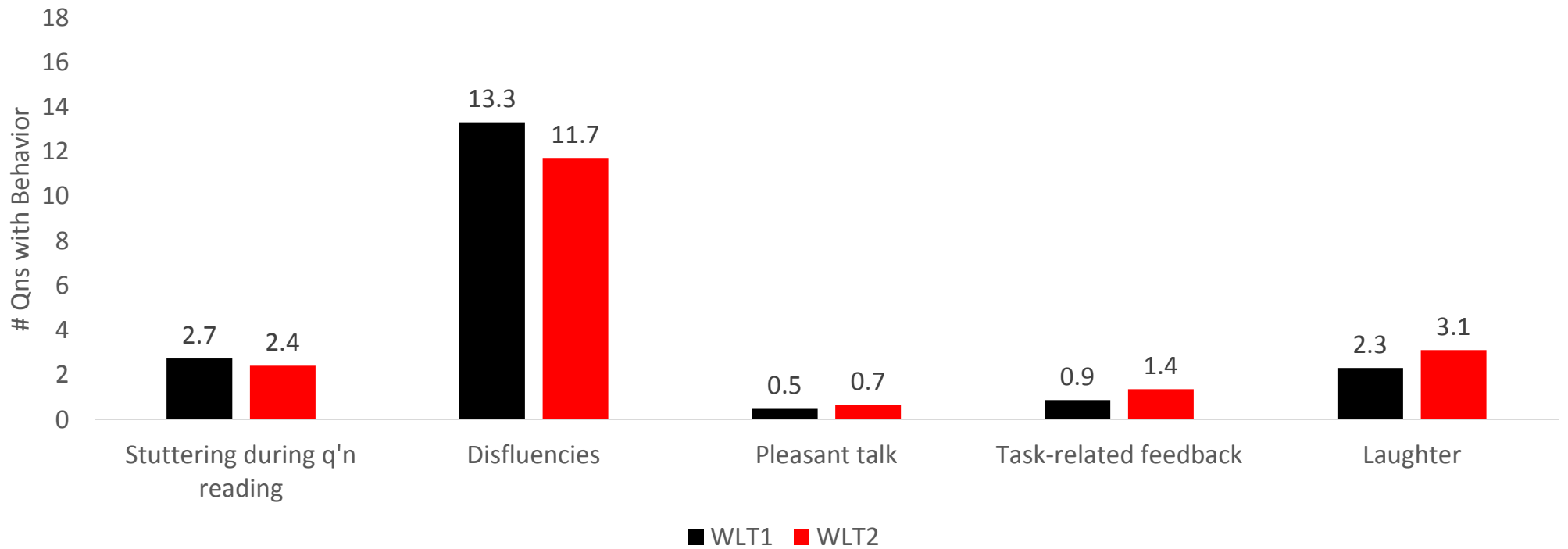
Nonstandardized behaviors

Primary Independent variable: Within-survey experience

- Log-transformed ordinal counter for within-survey experience
  - WLT1: Ranges from 1 to 27
  - WLT2: Ranges from 1 to 79

- Control variables
  - Overall interviewer experience
  - Interviewer-level cooperation rate, item NR rate, whether R ever refused, complete at first contact, # call attempts, time of day I'w completed
  - Number of questions asked
  - Number of answer changes
  - Respondent sex, age, education employment status, income HH size, parental status, volunteer status, computer usage
  - Interviewer race, gender, interviewer worked primarily weekday evening shifts
  - Version indicator and cell phone interview indicator for WLT2

Analytic strategy

- Two-level random intercept models
  - Poisson models for the interviewer behaviors
    - Number of questions as the exposure variable
  - Linear models for interview length

$$Log(IwBehaviors)_{ij} = \gamma_{00} + \beta_1 Ln(IwOrder)_{ij} + \boldsymbol{\beta_p} \textbf{Controls}_{pij} + u_{0j}$$

$$Length_{ij} = \gamma_{00} + \beta_1 Ln(IwOrder)_{ij} + \beta_2 IwBehaviors_{ij} + \boldsymbol{\beta_p} \textbf{Controls}_{pij} + u_{0j} + \varepsilon_{ij}$$

- Estimated using Stata 15.1 `mepoisson` and `mixed`

This paper

- What interviewer behaviors change over the course of the data collection period in two telephone surveys?

- Do these behaviors account for changes in survey length over the course of the data collection period?

Predicting behaviors in each study

- Focus only on interview order (within-survey experience) coefficient

$$Log(IwBehaviors)_{ij} = \gamma_{00} + \beta_1 Ln(IwOrder)_{ij} + \boldsymbol{\beta_p}\mathbf{Controls}_{pij} + u_{0j}$$

Standardized Interviewing Behavior: Associated with within-survey experience?

	WLT1	WLT2
Exact question reading	0.017	0.001
Nondirective probes	0.020	**-0.033***
Exact verification	0.020	**-0.051****
Appropriate clarification	0.091+	-0.034
Appropriate feedback	-0.010	**-0.035*****

n/s = not significant; +p<.10, * p<.05, ** p<.01, *** p<.001, **** p<.0001

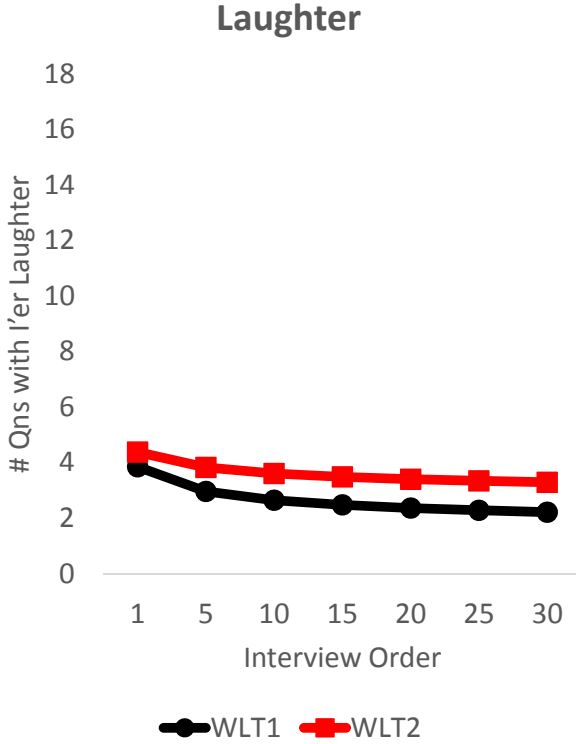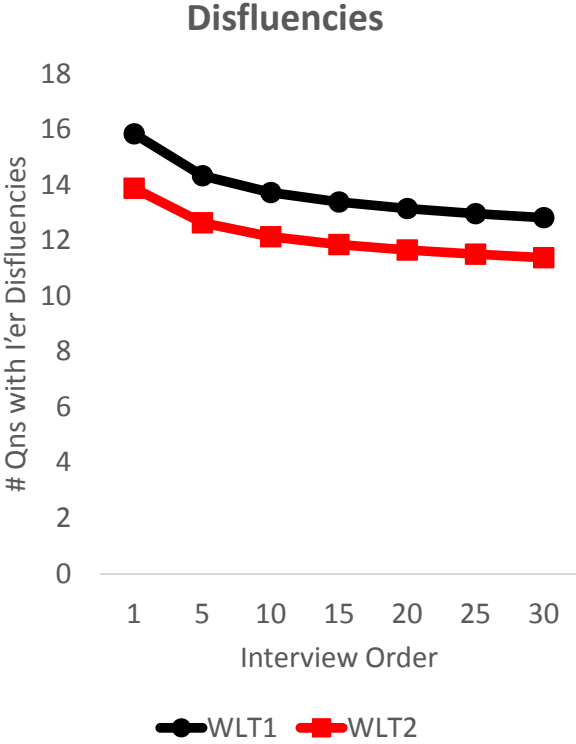# Decreases in standardized behaviors as interviewers gain within-study experience in WLT2; No change in WLT1

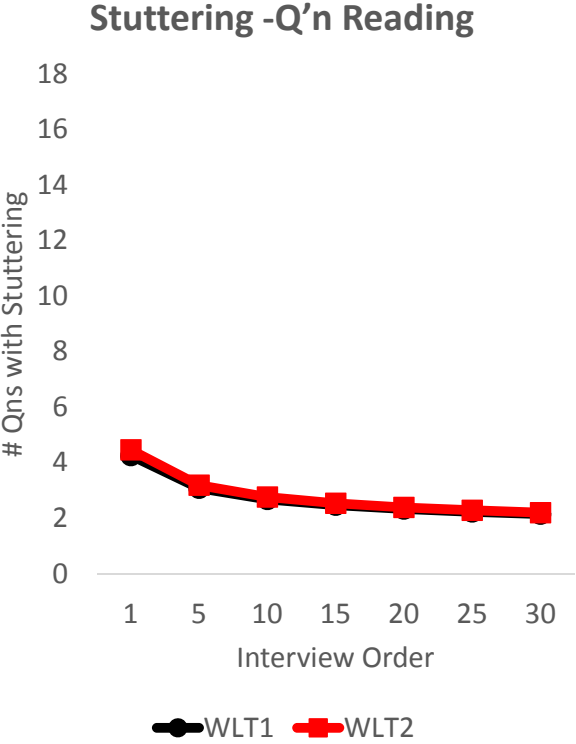**Nondirective Probes**

Exact verification

Appropriate Feedback
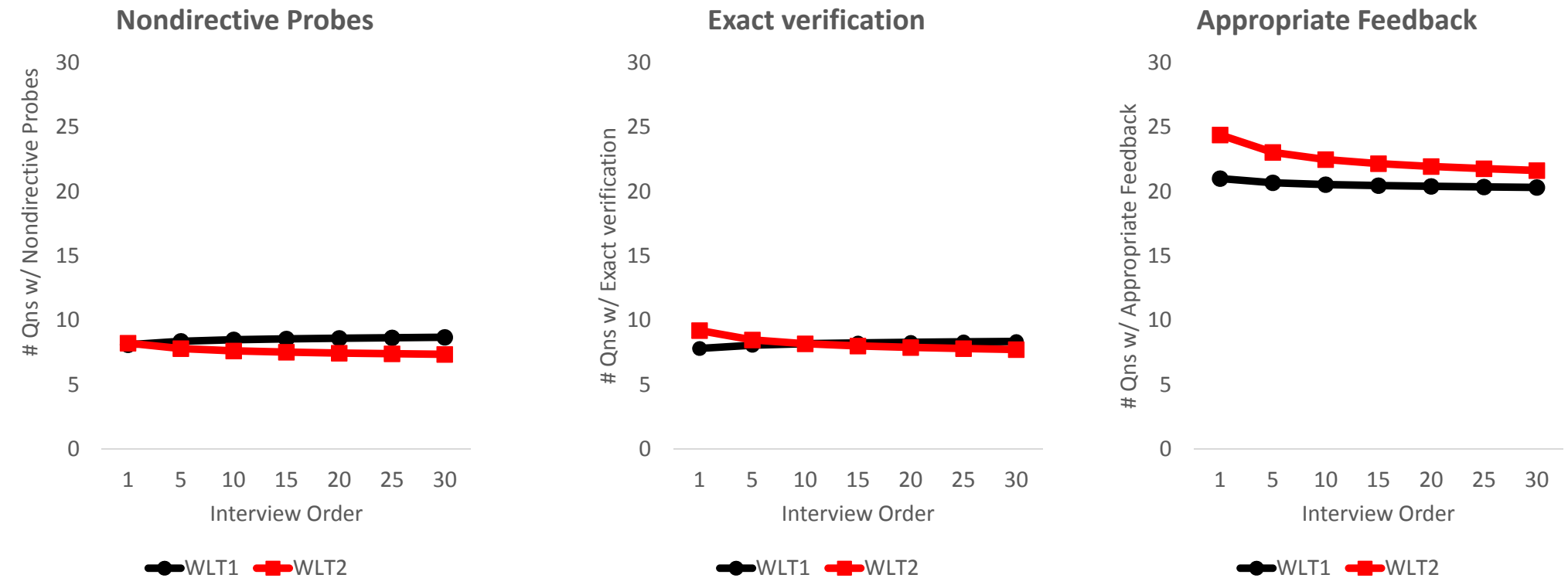
Efficiency Behaviors: Associated with within-study experience?

	WLT1	WLT2
Stuttering during q'n reading	-0.201****	-0.201****
Disfluencies	-0.062**	-0.058****
Pleasant talk	-0.086	-0.098
Task-related feedback	-0.151*	-0.052
Laughter	-0.162****	-0.084****

n/s = not significant; * p<.05, ** p<.01, *** p<.001, **** p<.0001

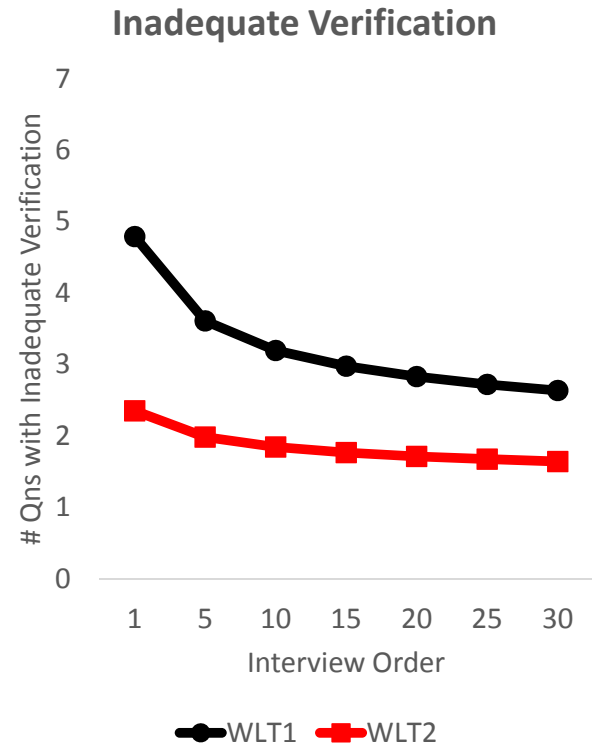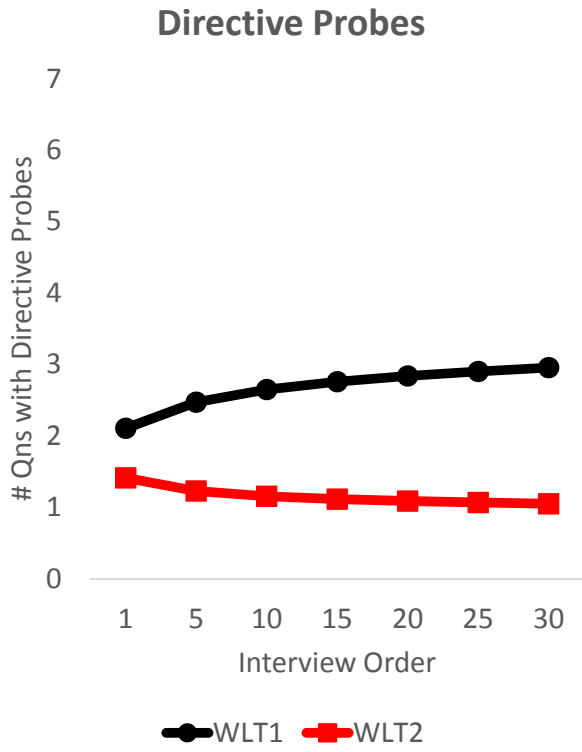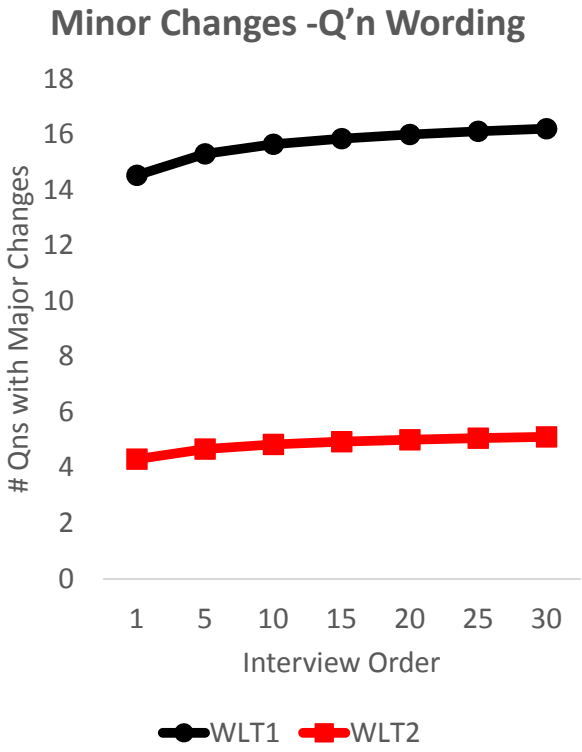# Fewer efficiency behaviors as interviewers gain within-study experience

Nonstandardized Behaviors: Associated with within-study experience?

	WLT1	WLT2
Any changes in question wording	0.009	**0.043****
Minor changes in question wording	0.032+	**0.050***
Major changes in question wording	-0.050+	**0.036***
Directive probes	**0.100***	**-0.087***
Inadequate verification	**-0.176******	**-0.106******
Interruptions	-0.046+	**-0.069****

n/s = not significant; * p<.05, ** p<.01, *** p<.001, **** p<.0001

Mixed changes in nonstandardized behaviors behaviors as interviewers gain within-study experience across the studies.

Summary: Interviewer behaviors

- Interviewers do change their behaviors as they gain experience

- Interviewers become more efficient in administering questions.
  - Have fewer questions with stutters, disfluencies, and laughter

- Interviewers experience changes in both standardized and non-standardized behaviors, although these replicate less well across studies.
  - In WLT1, few changes in standardized behaviors. In WLT2, fewer standardized behaviors.
  - Across both studies, lose inadequate verification. Other changes in nonstandardized behaviors less consistent.

This paper

- What interviewer behaviors change over the course of the data collection period in two telephone surveys?

- Do these behaviors account for changes in survey length over the course of the data collection period?

What behaviors are associated with overall interview length?

$$Length_{ij} = \gamma_{00} + \beta_1 Ln(IwOrder)_{ij} + \beta_2 IwBehaviors_{ij} + \boldsymbol{\beta_p}\mathbf{Controls}_{pij} + u_{0j} + \varepsilon_{ij}$$

- Look at the interview order coefficient as groups of behaviors are included in the model.

The interviewer behaviors partially explain interview length. Especially efficiency behaviors in WLT1.

Takeaways

- Interviewers generally don't lose their standardized behaviors over the field period.
  - This is good news. Where there are notable losses in standardized behaviors, it appears to be in feedback behaviors (ok; thank you).
  - Standardized behaviors explain between none and 20% of the change in interview length.
- Interviewers do become more efficient in administering surveys over the field period.
  - Efficiency behaviors explain between 17% and all of the change in interview length.
- Interviewers do change in their use of nonstandardized behaviors.
  - Some nonstandardized behaviors (inadequate verification) decrease. May be tradeoffs between major changes in question wording and directive probes.
  - Nonstandardized behaviors explain between 14 and 18% of the change in interview length.

Limitations

- Looked only at interviewer behaviors, but many interviewer behaviors occur in reaction to respondent behaviors.
  - Future research will examine changes in respondent behaviors as well.

- Two surveys conducted two years apart, but one organization conducting the survey.
  - Future research will add in a survey conducted by a different organization.

- Results largely replicate using turns rather than questions.
  - But some model sensitivity to the collection of behaviors included.

Summary

- Interviewer behaviors do change over the course of the data collection period.

- Interviewer behaviors are related to interview length.

- But how interviewer behaviors are related to interview length is more complicated than simply the number of questions on which the behaviors occur over the interview.
  - Are behaviors getting shortened as well as eliminated?
  - How do question characteristics themselves affect the occurrence of these behaviors?
  - More work to be done!

Thanks!

Kristen Olson

@olson_km

kolson5@unl.edu