Effects of Question, Respondent, and Interviewer Characteristics on Interactional Indicators of Respondent and Interviewer Processing of Health-Related Questions

Jennifer Dykema[1], Nora Cate Schaeffer[1,2], Dana Garbarski[3], Erik V. Nordheim[4], Kristen Cyffka[5]

[1]University of Wisconsin Survey Center
[2]Department of Sociology, University of Wisconsin-Madison
[3]Center for Women's Health and Health Disparities Research, University of Wisconsin-Madison
[4]Department of Statistics, University of Wisconsin-Madison
[5]U.S. Census Bureau

Interviewer-Respondent Interaction Workshop

15-16 May 2013

UWSC
UNIVERSITY of WISCONSIN
SURVEY CENTER

Background and research objectives

- Question writers often focus on question characteristics
  - e.g., length, difficulty, response format
- Recommendations for writing questions are
  - formulated around question characteristics
  - based on research (beliefs) about impact of question characteristics on outcomes

- Know a lot about effects of some question characteristics on data quality
- Still developing a comprehensive typology in which
  - question characteristics are cataloged
  - effects on INTs' and Rs' processing are understood
 - effects on data quality are documented

Background and research objectives

- Overarching goals
  - Present conceptual model for exploring effects of question characteristics on processing, indicators of processing, and data quality
  - Examine effects of characteristics (question, respondent, and interviewer) on interviewer-respondent interactional outcomes for questions about physical and mental health status in a survey of older adults
  - Advance research on question characteristics by presenting a mixed-effects model that takes into account the complicated nested and crossed structure of the data

Conceptual Model:

Interactional Model of the Question-Answer Sequence

Interactional Model of Question-Answer Sequence
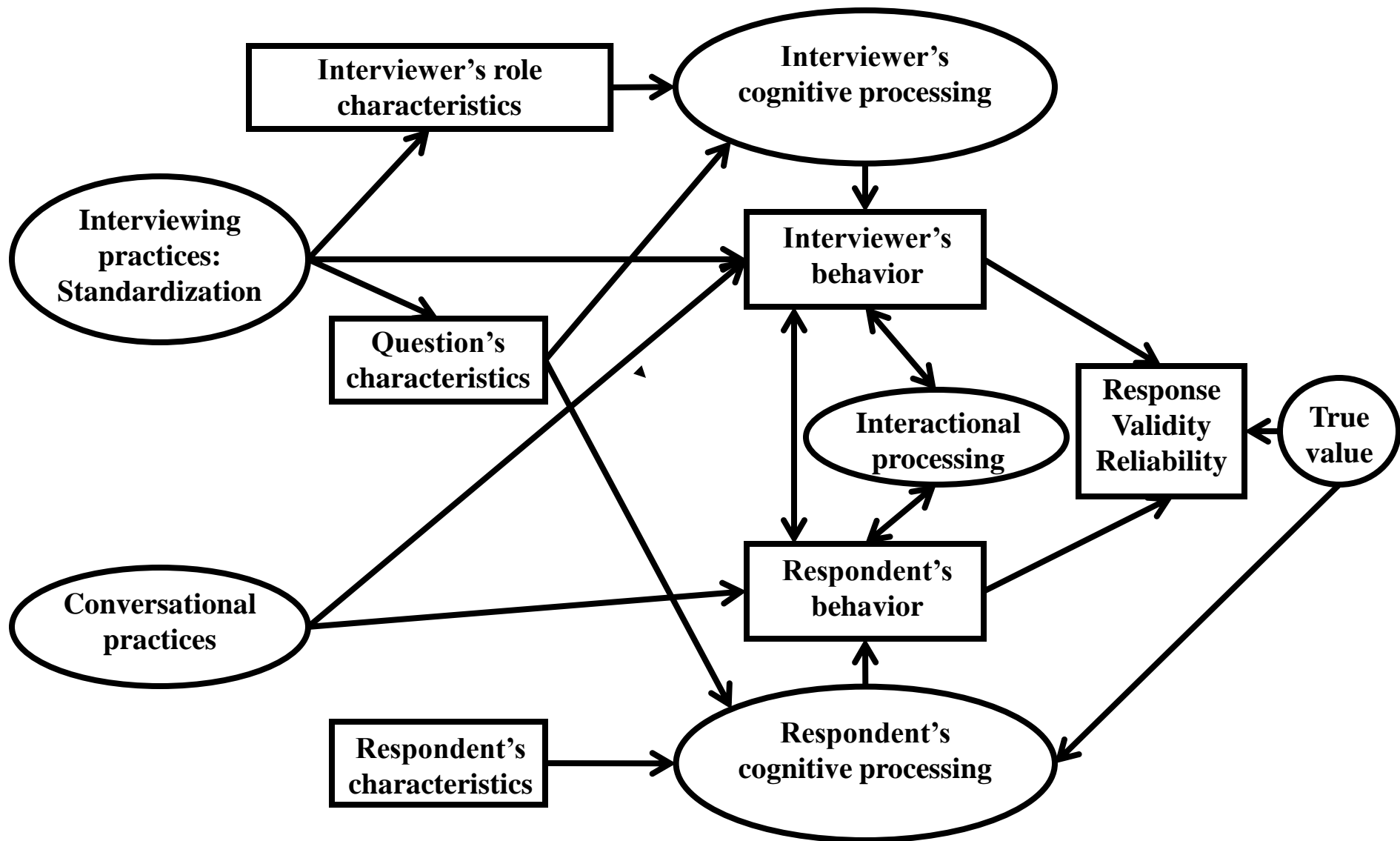
Effects of Characteristics on Interactional Outcomes:

Overview of the Data

Survey data: The Wisconsin Longitudinal Study

- 1/3 random sample of Wisconsin high school class of 1957
- Telephone interviews digitally recorded from 2003-05 wave
- 355 cases randomly sampled
  - Randomly selected INTs
  - Within INTs, stratified Rs by cognitive ability, and attempted to select 5 Rs (low, med, high ability)
- Analyze characteristics of 23 questions from the Health module

Behavioral outcomes: Interaction coding data

- Behavioral outcomes from interviewer-respondent interactions
- Interviews transcribed, coded in Sequence Viewer (Dijkstra)
- Elaborate coding scheme: Over 100 behaviors
  - Ex: pauses, tokens, uncodable answers, etc.
  - Small subset analyzed here
- Question-answer sequence
  - Unit of analysis
  - Starts with reading of the survey question by INT, ends with the last utterance spoken by INT or R before INT reads next question
  - Made up of behaviors
    - Each utterance spoken by INT or R is coded
  - 8150 question-answer sequences

Analysis: Mixed effects logistic regression

- Data have a complicated multilevel structure

Interviewers		1					2					...	78					79				
Respondents	1	2	3	4	5	6	7	8	9	10			346	347	348	349	350	351	352	353	354	355
Questions 1																						
2																						
3																						
...																						
11																						
12																						
13																						
...																						
21																						
22																						
23																						

- Use mixed effects logistic regression models
- Include random effects for INT, R within INT, Q, and INTs crossed by Qs

Effects of Characteristics on Interactional Outcomes:

Variables and Hypotheses

Behavioral outcomes: Dependent variables

- Interviewers
  - Question-reading accuracy
    - exact versus any changes
  - Tokens (any)
    - e.g. "well," "um," "oh," "er"
    - terms or phrases with a neutral connotation, may be linked to processing difficulties
  - "Okay"s (any)
- Respondents
  - Index of "problem" behaviors
    - uncodable answers, qualified answers, etc.
  - Tokens (any)
  - Questions (any)

Question characteristics: Response format

- **Yes-No**
  - Use the response categories yes or no (not read)
    - *Have you been able to see at all?*
- **List-Item**
  - Use the response categories yes or no but have a format in which categories are listed in the body of the question and may be heard as response options
    - *Have you been able to bend, lift, jump and run without difficulty and without help or equipment of any kind?*
- **Selection**
  - Use a predetermined set of response categories
    - *In general, would you say your health is excellent, very good, good, fair, or poor?*

Question characteristics: Response format

- Effects on interviewers
  - Negative outcomes more likely with list-item and selection because they may be harder to administer
- Effects on respondents
  - Negative outcomes more likely with list-item and selection because they have a more complicated structure to process

Question characteristics: Question length

- Measured as the raw number of words per question
- Effects on interviewers
  - Negative (Presser and Zhao 1992)
- Effects on respondents
  - Positive – more time to think (Blair et al. 1977)
  - Negative – comprehension difficulties (Holbrook et al. 2006)

Question characteristics: Readability

- Measured using the Flesch-Kincaid Grade Level score
- Effects on interviewers
  - Relatively unknown
  - Negative with certain question structures (Oksenberg et al. 1991)
- Effects on respondents
  - Negative with more difficult text (Knauper et al.)

Question characteristics: Problem Classification Coding Scheme (CCS) (Forsyth et al. 2004)

- 28 codes for problems
- Grouped under the 4-stage question-answer model (comprehension, retrieval, judgment, response)
- "Comprehension and communication"
  - "Question content"
    - "Vague topic/term"
- Trained coder coded each question, number of problems tallied
- Effects on interviewers
  - Higher scores associated with more behavior-coded INT problems
- Effects on respondents
  - Higher scores associated with more behavior-coded R problems

Question characteristics: Question Understanding Aid (QUAID) (Graesser et al. 2006)

- Computer tool to evaluate how difficult Qs are to comprehend
- Enter Q text and QUAID returns list of potential problems with question comprehension
  - unfamiliar technical terms, vague or imprecise relative terms, vague or ambiguous noun phrases, complex syntax, and working memory overload
- Tallied the number of problems identified
- Effects on interviewers
  - Relatively unknown
- Effects on respondents
  - Qs rated as difficult by QUAID less likely to be read
  - Expect behavior-coded R problems more likely with higher scores

Question characteristics: Inclusion of a parenthetical

- Several questions included parenthetical statements left to interviewer's discretion to read or omit
  - *(During the past four weeks) Have you been able to see at all?*
- Include indicators for whether question
  - Did not include a parenthetical
  - Included a parenthetical that was read
  - Included a parenthetical that was not read
- Effects on interviewers
  - Negative impact possibly causing more reading errors
- Effects on respondents
  - Positive impact in making the question clearer when read

Respondent and interviewer characteristics

- Respondent characteristics
  - Educational attainment
  - Cognitive ability (IQ)
  - Health status (SF-12)
- Interviewer characteristics
  - Months of experience prior to survey: GE 6 M vs less
- R-I gender match
- Other control variables

Effects of Characteristics on Interactional Outcomes:

Results

Results: Interviewer Behaviors

Fixed Effects	Exact reading	Token (any)	"OK" (any)
	OR	OR	OR
Question Characteristics			
Response format [Yes-No]			
List-item	1.16	1.02	1.34
Selection	**2.36***	1.47	**2.28+**
Question length	**0.88*****	1.00	1.03
Flesch grade	**1.07****	1.03	0.97
CCS	0.98	**1.17***	**1.22***
QUAID	1.07	1.01	**1.35***
Parenthetical [Not used]			
No parenthetical	0.84	1.09	0.96
Parenthetical used	**0.32*****	**2.06*****	**1.40***
Respondent Characteristics			
Education	**0.95+**	1.00	1.00
Cognitive ability	1.00	1.00	1.00
Health status [Highest third]			
Lowest third	**0.73+**	**1.40+**	**1.66****
Middle third	0.90	1.03	1.12
Missing	0.79	1.03	**1.70***
Interviewer Characteristics			
Mnths exp GE 6 M [vs less]	**2.14+**	0.97	**0.53***
R gender match [INT:F and R:F]			
INT:female and R:male	0.91	1.10	1.22
INT:male and R:male	1.59	1.46	0.80
INT:male and R:female	1.47	**1.82***	1.18

Results: Respondent Behaviors

Fixed Effects	Problem (any)	Token (any)	Q-Ask (Any)
	OR	OR	OR
Question Characteristics			
Response format [Yes-No]			
List-item	1.74	**2.28+**	1.25
Selection	**6.00****	**5.97****	**2.55+**
Question length	1.04	0.98	0.96
Flesch grade	0.99	0.99	1.00
CCS	1.08	**1.28+**	**1.53****
QUAID	**1.47***	1.32	**1.44***
Parenthetical [Not used]			
No parenthetical	0.60	0.88	0.68
Parenthetical used	0.94	1.07	1.36
Respondent Characteristics			
Education	1.01	1.02	**0.93+**
Cognitive ability	**0.99***	**0.99***	1.00
Health status [Highest third]			
Lowest third	**1.57****	**1.65*****	1.07
Middle third	**1.24+**	**1.29***	1.02
Missing	**1.62***	**1.46+**	0.85
Interviewer Characteristics			
Mnths exp GE 6 M [vs less]	1.06	1.08	0.99
I-R gender match [INT:F and R:F]			
INT:female and R:male	1.18	0.95	0.97
INT:male and R:male	0.88	1.00	**0.50****
INT:male and R:female	1.00	0.85	**0.65***

Summary and conclusions

- Question characteristics
  - Response format
  - Question length, readability, and coding schemes
- Respondent characteristics
  - Rs with lower cognitive ability and health status more likely to exhibit interactional problems
- Interviewer experience
  - INTs with more experience appear to maintain standardization better than those with less
  - Need better measures of interviewing quality
  - Need to examine more interviewer outcomes
- I-R gender match

Limitations

- Coding approaches are not independent (e.g., all code for question length in some way)
- Limited pool of questions examined
  - Primarily yes/no type questions about health
  - Not randomly sampled from a population of questions with many different characteristics
- Limited number of interviewer, respondent, and interactional behaviors examined
- Implicitly assume that behavioral measures are associated with poorer quality data

Future directions

- Continue work examining characteristics with a bank of questions with more varied characteristics
    - E.g., extend model to attitudinal questions
    - E.g., extend model to questions with an identification response format
- Examine more interviewer-respondent behaviors
    - E.g., indicator of how adequately interviewers administer follow ups
- Examine additional approaches for coding question characteristics
- Incorporate measures of validity and reliability as outcomes to predict

Thank You!

For copies of this presentation or more information, contact:

Jennifer Dykema

dykema@ssc.wisc.edu

Please visit us at:
www.uwsc.wisc.edu