

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

# Response Times as an Indicator of Data Quality: Associations with Interviewer, Respondent, and Question Characteristics in a Health Survey of Diverse Respondents

Dana Garbarski

Loyola University Chicago, dgarbarski@luc.edu

Jennifer Dykema

University of Wisconsin Survey Center, University of Wisconsin-Madison, dykema@ssc.wisc.edu

Nora Cate Schaeffer

University of Wisconsin Survey Center, University of Wisconsin-Madison; Department of Sociology, University of Wisconsin-Madison, schaeffe@ssc.wisc.edu

Dorothy Farrar Edwards

University of Wisconsin-Madison, dfedwards@education.wisc.edu

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



Part of the [Quantitative, Qualitative, Comparative, and Historical Methodologies Commons](#)

---

Garbarski, Dana; Dykema, Jennifer; Schaeffer, Nora Cate; and Edwards, Dorothy Farrar, "Response Times as an Indicator of Data Quality: Associations with Interviewer, Respondent, and Question Characteristics in a Health Survey of Diverse Respondents" (2019). *2019 Workshop: Interviewers and Their Effects from a Total Survey Error Perspective*. 4.  
<http://digitalcommons.unl.edu/sociw/4>

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.

# Response times as an indicator of data quality: Associations with interviewer, respondent, and question characteristics in a health survey of diverse respondents

Dana Garbarski<sup>1</sup>, Jennifer Dykema<sup>2</sup>,  
Nora Cate Schaeffer<sup>2,3</sup>, Dorothy Farrar Edwards<sup>4</sup>

<sup>1</sup>Department of Sociology, Loyola University Chicago

<sup>2</sup>University of Wisconsin Survey Center, University of Wisconsin-Madison

<sup>3</sup>Department of Sociology, University of Wisconsin-Madison

<sup>4</sup>Department of Kinesiology-Occupational Therapy,  
University of Wisconsin-Madison

Interviewer Workshop  
February 25-28 2019  
University of Nebraska-Lincoln

# Predictors of survey data quality

- Data obtained in the survey interview are a collaborative achievement accomplished through the interplay of
  - Questions
  - Respondents
  - Interviewers
- Our field is still documenting whether, when, and how each of these characteristics combine to influence data quality

## Taxonomy of Question Characteristics (Schaeffer & Dykema 2011, 2015)

| Classes of Characteristics                                    | Examples of individual features   |
|---|---|
| Question topic  | Health, politics  |
| Question type   | Event or behavior, evaluation or judgment, classification                 |
| Response dimension  | Occurrence, frequency, intensity, valence                                 |
| Conceptualization and operationalization of the target object | Labels for target object and response dimension                           |
| Question structure  | Filter and follow-up question, battery                                    |
| Response format or question form                              | Yes/no, selection, discrete value, field-coded open, record-verbatim open |
| Response categories   | Type, number, and labeling  |
| Question wording  | Length, readability   |
| Question implementation                                       | Mode, orientation of scale on screen, instructions to interviewers        |

# Question characteristics and data quality

- Experimental approaches
  - Write alternative forms of questions with particular characteristics, holding all others constant
  - Administer the questions randomly so each respondent is only exposed to questions with specific characteristics
- Observational approaches
  - Entire survey, wider range of characteristics
  - Everyone gets the questions, exposed to multiple characteristics
- Results from previous studies on the influence of specific question characteristics on data quality depend on approach, characteristics considered, and data quality measures

# Response time as a measure of data quality

- Length of time spent on an entire question-answer sequence
  - Related to yet distinct from “response latency”
- Indirect
- Nonlinear
- In general, longer response times indicate longer processing or interaction (and thus potential problems) from the respondent, the interviewer, or both

# Question characteristics and response time

- Considering just one measure of data quality (response time), how is this associated with question characteristics?
- Response times are associated with various characteristics of questions, respondents, and interviewers (where applicable) across web, telephone, and face-to-face interviews
- Differ in which characteristics they have examined, and in how those characteristics are operationalized

(Couper and Kreuter 2013; Loosveldt and Beullens 2013; Olson and Smyth 2015; Yan and Tourangeau 2008)

# Question characteristics and response time

- We as a field are still in the process of documenting both the inputs and outputs to optimal question design
  - Which question characteristics are associated with better data quality broadly defined
  - For a given measure of data quality, which question characteristics are associated with it in theoretically sensible ways



# System-based question coding schemes

- Employ a scheme to code multiple characteristics
  - Problems for interviewers and respondents
  - Problem score to predict data quality
- Produce systematic compilations of question characteristics and potential problems
  - Question Understanding Aid – QUAID
    - Graesser et al. 2006; <http://quaid.cohmetrix.com/>
  - Question Appraisal System – QAS
    - Willis 2005; <http://appliedresearch.cancer.gov/areas/cognitive/qas99.pdf>
  - Survey Quality Predictor -- SQP
    - Saris and Gallhofer 2007; <http://sqp.upf.edu/>

# System-based question coding schemes

- Unclear whether question coding schemes are associated with response times
  - Schemes could pick up problems that increase the overall response times net of effects of individual question characteristics
  - But individual question characteristics may be enough, e.g., redundancy in length and the complexity of the question

# Differential impact on response time

- Does impact of question characteristics on response time vary by
  - Interviewers' experience
  - Respondents' race/ethnicity

# Interviewers' experience, questions, and response times

- Interviewers' experience leads to familiarity with questions in ways that develop routines, improve fluency, anticipate problems, etc.
  - Within a given study
- Interviewers' task complexity
  - Instructions given to the interviewer
  - Emphasis in the question wording
  - Parenthetical statements included in questions

(Kirchner and Olson 2017; Olson and Peytchev 2007; Olson and Smyth 2015)

# Respondents' race/ethnicity, questions, and response times

- Differences across racial/ethnic groups in how respondents process survey concepts, answer survey questions, and interact with interviewers
- Given what we know about various question characteristics and their relationship to various measures of data quality, we know even less about how these effects might vary by race/ethnicity

(Holbrook, Cho, and Johnson 2006; Johnson, Shavitt, and Holbrook 2011; Warnecke et al. 1997)

# Research aims

- Examine how response times are associated with characteristics of questions, interviewers, and respondents
- Broad set of question characteristics
- System-based question coding schemes
- Differential impact on response times
  - Interviewers' experience
  - Respondents' race/ethnicity
- Study of racially/ethnically diverse respondents answering questions about trust in medical researchers, participation in medical research, and health

# Data

- “Voices Heard” CATI survey
- N=410 completed interviews using a quota sampling design
  - American Indian, Black, Latina/o, White respondents
- Conducted in Wisconsin from Oct 2013 – Mar 2014
- 96 question survey
- Average interview 25.21 minutes



## Individual question characteristics (N=96)

|                            |   |
|----------------------------|---|
| Word count                 |   |
| Flesch-Kincaid grade level |   |
| Question type              | Event/behavior, subjective, demographic   |
| Question form              | Yes/no, nominal, open, selection  |
| List item ("and," "or")    |   |
| Battery                    | First in battery, later in battery,<br>first in series, later in series, standalone |
| Definition                 |   |
| Interviewer instructions   |   |
| Parenthetical              |   |
| Emphasis                   |   |
| Sensitive                  |   |
| Race-focused               |   |



# System-based coding of questions

- Question Understanding Aid – QUAID
  - Graesser et al. 2006; <http://quaid.cohmetrix.com/>
- Question Appraisal System – QAS
  - Willis 2005; <http://appliedresearch.cancer.gov/areas/cognitive/qas99.pdf>
- Survey Quality Predictor – SQP
  - Saris and Gallhofer 2007; <http://sqp.upf.edu/>

# Respondents' and interviewers' characteristics

## **Respondents (N=410)**

- Race/ethnicity
- Gender
- Age
- Education
- Household income

## **Interviewers (N=24)**

- Race/ethnicity
- Gender
- Age
- Prior experience
- Number of interviews completed

# Methods: Unit of analysis and outcomes

- 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
    - Total n = 39,053 sequences (410 Rs x 95 or 96 Qs)
- Analysis
  - Cross-classified random-effects linear regression models to predict the log-transformed response times using Stata 15.1 and the *mixed* command with restricted maximum likelihood (*reml*)

|   | Mean  | Std. Dev. | Minimum | Maximum |
|---|-------|-----------|---------|---------|
| <b>Response time (trimmed)</b>            | 13.04 | 8.42      | 1.00    | 92.00   |
| <b>Response time (trimmed and logged)</b> | 2.37  | 0.66      | 0.00    | 4.52    |

# Results

## Regress response time on question characteristics

- 1) QUAID
- 2) QAS
- 3) SQP
- 4) Set of individual question characteristics
- Full model: QUAID, QAS, SQP, set of question characteristics

|          | 1, 2, 3 | Full model |
|----------|---------|------------|
| 1) QUAID | +       |            |
| 2) QAS   |         |            |
| 3) SQP   |         |            |

|          | 1, 2, 3 | Full model |
|----------|---------|------------|
| 1) QUAID | +       |            |
| 2) QAS   |         |            |
| 3) SQP   |         |            |

|          | 1, 2, 3 | Full model |
|----------|---------|------------|
| 1) QUAID | +       |            |
| 2) QAS   |         |            |
| 3) SQP   |         |            |



|                                  | 4         | Full model |
|----------------------------------|-----------|------------|
| 4)                               |           |            |
| Word count                       | +         | +          |
| Flesch-Kincaid grade level score | +         | +          |
| Question type                    |           |            |
| Event/behavior                   | Reference |            |
| Subjective                       |           |            |
| Demographic                      |           |            |
| Question form                    |           |            |
| Yes/no                           | Reference |            |
| Nominal                          |           |            |
| Open                             |           |            |
| Bipolar selection                | +         | +          |
| Unipolar selection               | +         | +          |

|                                  | 4         | Full model |
|----------------------------------|-----------|------------|
| 4)                               |           |            |
| Word count                       | +         | +          |
| Flesch-Kincaid grade level score | +         | +          |
| Question type                    |           |            |
| Event/behavior                   | Reference |            |
| Subjective                       |           |            |
| Demographic                      |           |            |
| Question form                    |           |            |
| Yes/no                           | Reference |            |
| Nominal                          |           |            |
| Open                             |           |            |
| Bipolar selection                | +         | +          |
| Unipolar selection               | +         | +          |

|                                  | 4         | Full model |
|----------------------------------|-----------|------------|
| 4)                               |           |            |
| Word count                       | +         | +          |
| Flesch-Kincaid grade level score | +         | +          |
| Question type                    |           |            |
| Event/behavior                   | Reference |            |
| Subjective                       |           |            |
| Demographic                      |           |            |
| Question form                    |           |            |
| Yes/no                           | Reference |            |
| Nominal                          |           |            |
| Open                             |           |            |
| Bipolar selection                | +         | +          |
| Unipolar selection               | +         | +          |

|                                  | 4         | Full model |
|----------------------------------|-----------|------------|
| 4)                               |           |            |
| Word count                       | +         | +          |
| Flesch-Kincaid grade level score | +         | +          |
| Question type                    |           |            |
| Event/behavior                   | Reference |            |
| Subjective                       |           |            |
| Demographic                      |           |            |
| Question form                    |           |            |
| Yes/no                           | Reference |            |
| Nominal                          |           |            |
| Open                             |           |            |
| Bipolar selection                | +         | +          |
| Unipolar selection               | +         | +          |

|                          | 4         | Full model |
|--------------------------|-----------|------------|
| Battery                  |           |            |
| First in battery         |           |            |
| Later in battery         | Reference |            |
| First in series          | +         | +          |
| Later in series          | +         |            |
| Standalone               | +         |            |
| Definition               |           |            |
| Interviewer instructions | +         |            |
| Parenthetical            | -         | -          |
| Emphasis                 | -         | -          |
| Sensitive                |           |            |
| About race               |           |            |

|                          | 4         | Full model |
|--------------------------|-----------|------------|
| Battery                  |           |            |
| First in battery         |           |            |
| Later in battery         | Reference |            |
| First in series          | +         | +          |
| Later in series          | +         |            |
| Standalone               | +         |            |
| Definition               |           |            |
| Interviewer instructions | +         |            |
| Parenthetical            | -         | -          |
| Emphasis                 | -         | -          |
| Sensitive                |           |            |
| About race               |           |            |

|                          | 4         | Full model |
|--------------------------|-----------|------------|
| Battery                  |           |            |
| First in battery         |           |            |
| Later in battery         | Reference |            |
| First in series          | +         | +          |
| Later in series          | +         |            |
| Standalone               | +         |            |
| Definition               |           |            |
| Interviewer instructions | +         |            |
| Parenthetical            | -         | -          |
| Emphasis                 | -         | -          |
| Sensitive                |           |            |
| About race               |           |            |

|                          | 4         | Full model |
|--------------------------|-----------|------------|
| Battery                  |           |            |
| First in battery         |           |            |
| Later in battery         | Reference |            |
| First in series          | +         | +          |
| Later in series          | +         |            |
| Standalone               | +         |            |
| Definition               |           |            |
| Interviewer instructions | +         |            |
| Parenthetical            | -         | -          |
| Emphasis                 | -         | -          |
| Sensitive                |           |            |
| About race               |           |            |



|                          | 4         | Full model |
|--------------------------|-----------|------------|
| Battery                  |           |            |
| First in battery         |           |            |
| Later in battery         | Reference |            |
| First in series          | +         | +          |
| Later in series          | +         |            |
| Standalone               | +         |            |
| Definition               |           |            |
| Interviewer instructions | +         |            |
| Parenthetical            | -         | -          |
| Emphasis                 | -         | -          |
| Sensitive                |           |            |
| About race               |           |            |

# Results

- Respondents' characteristics
  - Latino (vs. White) +
  - Age +
- Interviewers' characteristics
  - Women (vs. men) +

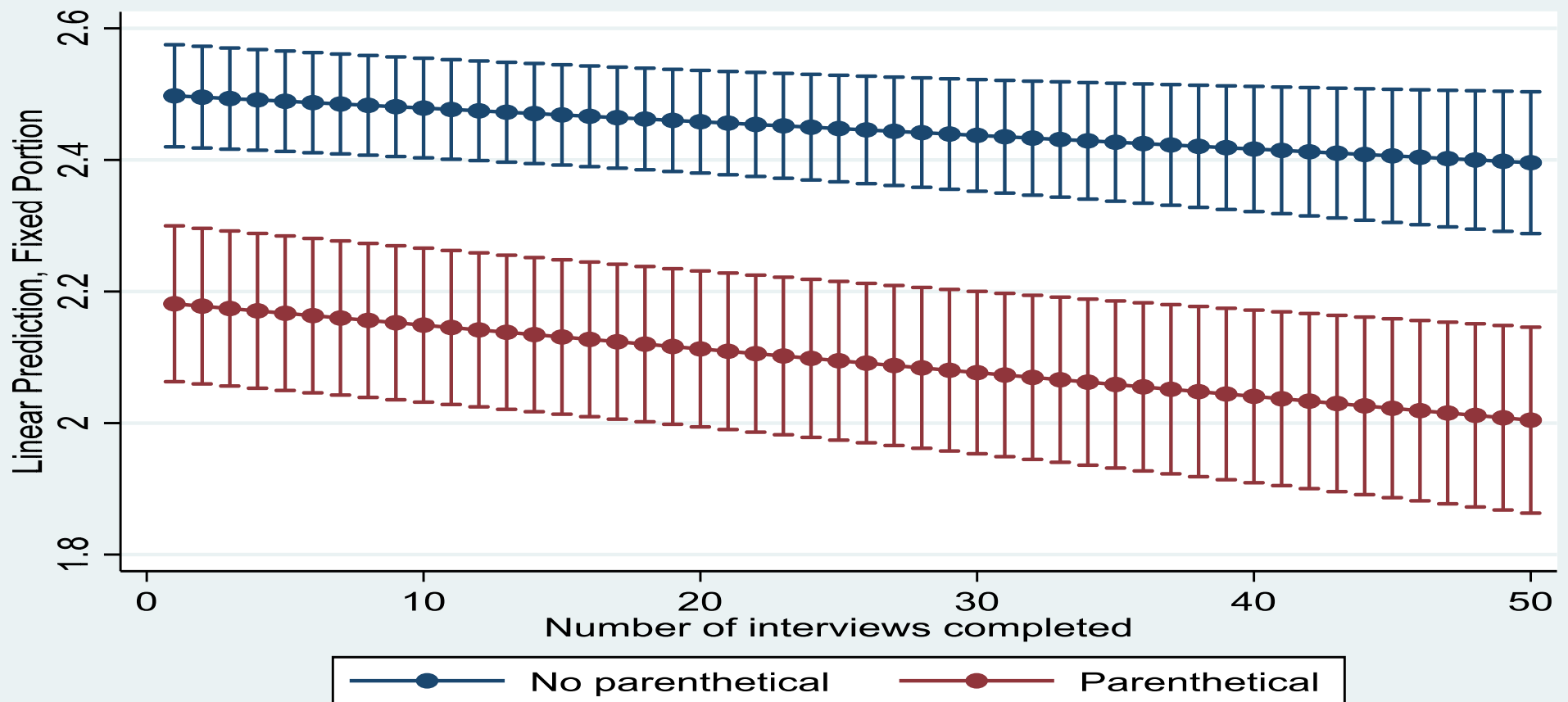
# Interviewers' experience

- Interactions of number of interviews completed with
  - Parenthetical
  - Interviewer instructions
  - Definitions

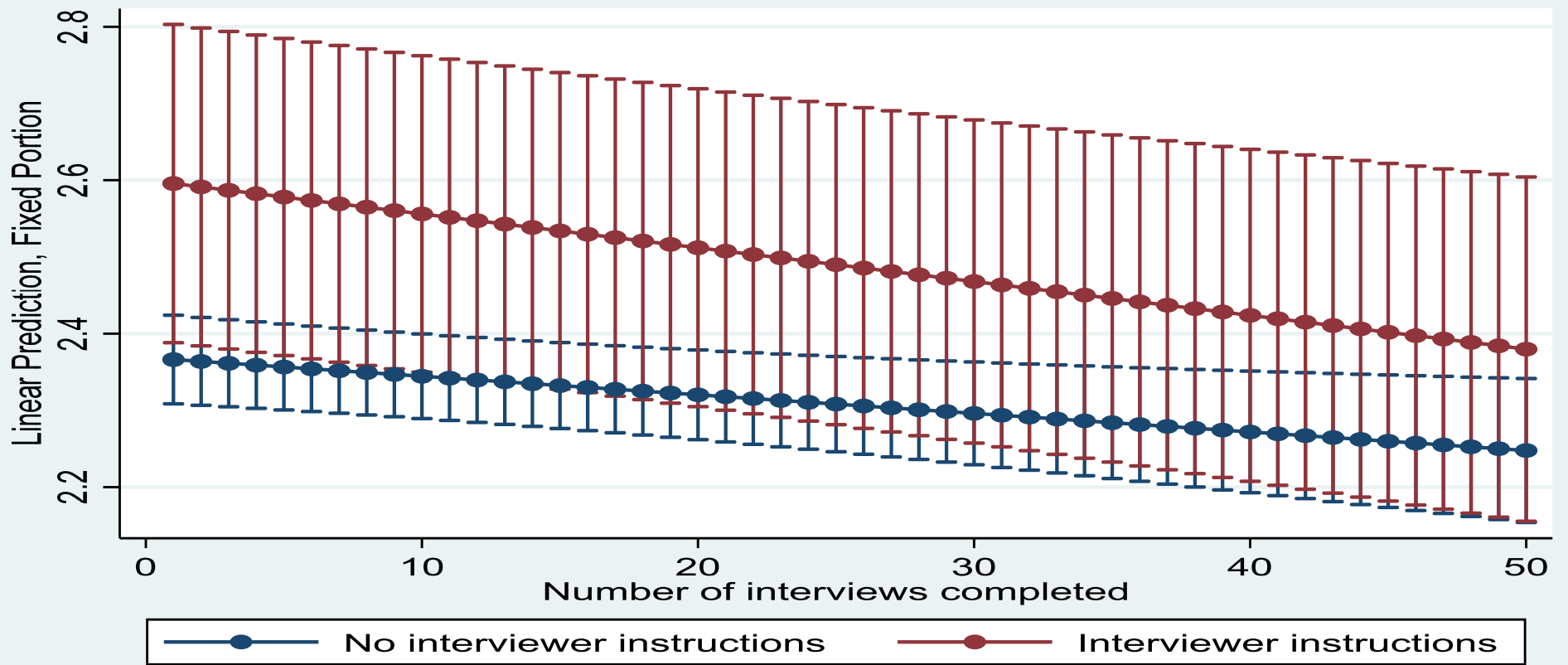
# Interviewers' experience

- Interactions of number of interviews completed with
  - **Parenthetical**
  - **Interviewer instructions**
  - Definitions

Estimated Marginal Means of (Log-transformed) Response Times by Number of Interviews Completed and Parenthetical Statements



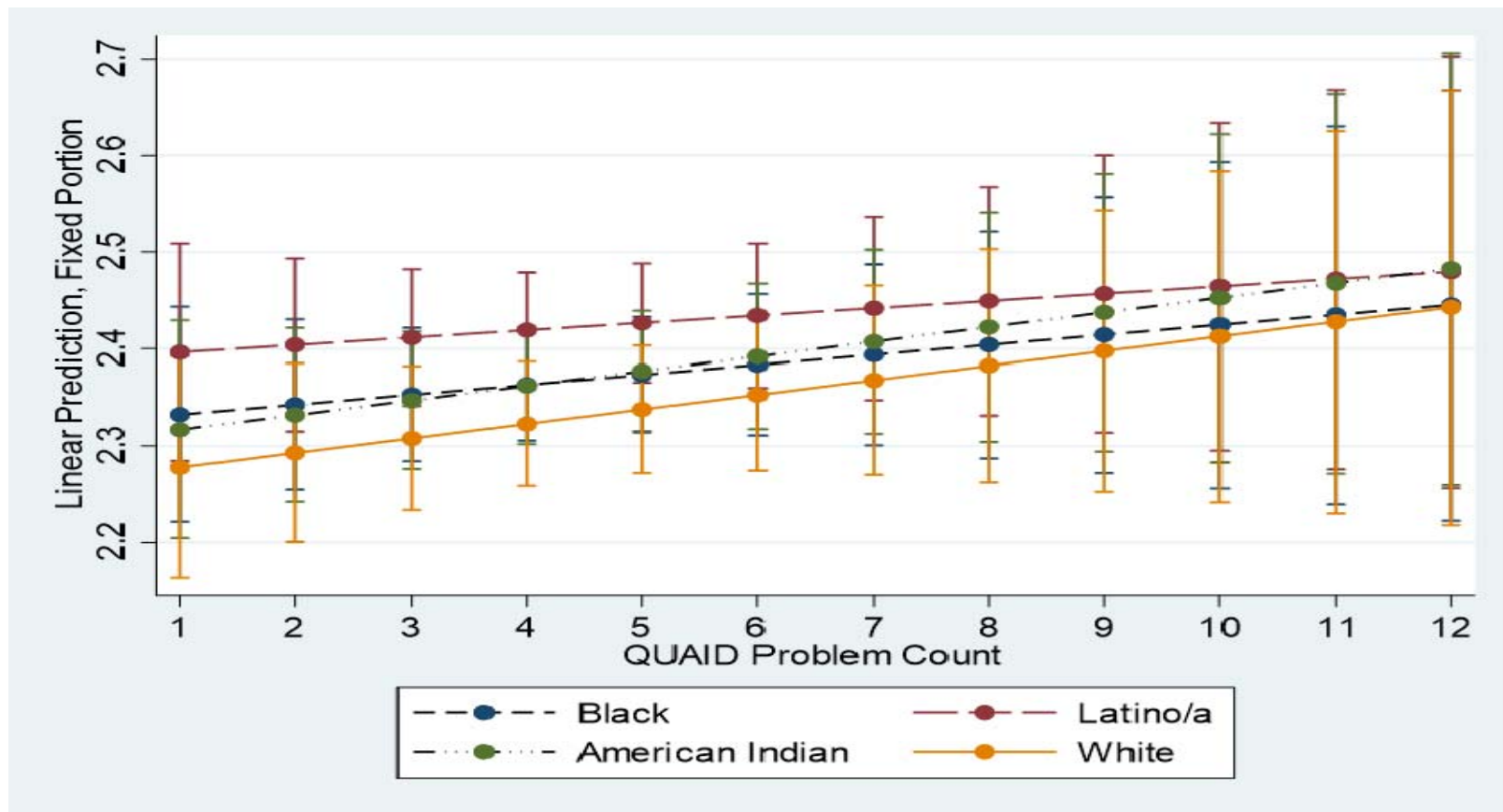
Estimated Marginal Means of (Log-transformed) Response Times by Number of Interviews Completed and Interviewer Instructions



# Respondents' Race/Ethnicity

- Interactions with all question characteristics
  - Significant for QUAID, QAS, SQP, word count, grade level, question type, question form, and whether the question took a list form, battery form, contained interviewer instructions, and about race
- All confidence intervals around estimated marginal means overlapped for a given level of the question characteristic of interest

# Estimated Marginal Means of Response Times by QUAID Problem Count and Race/Ethnicity





# Summary

- Study adds to the body of knowledge we are accumulating about both the inputs and outputs to question writing
  - Which question characteristics are associated with better data quality
  - Within a given measure of data quality (in this case response times), which question characteristics are associated with response times in theoretically sensible ways
- Approaches for testing questions with systems has expanded in recent decades
  - These did not really tell us anything about response time
- Structural dependency of question characteristics

Thank you!