

# Does typographic cueing improve the processing of information from survey questions on a mobile device?

Brian Falcone  
U.S. Census Bureau

AAPOR 72<sup>st</sup> Annual Conference  
May 20, 2017  
New Orleans, Louisiana

This presentation is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

# What is Typographic Cueing?

- Reveals text content structure through changes in weight, size, case, typeface, etc. (Keyes, 1993)
- Differentiates information categories

# Background Literature

- Use dark print for questions and light print for answer choices
- Separate optional or occasionally needed instructions from the question statement by font or symbol variations

Source: Tailored Design Method (Dilman, 2007)

# Examples in Practice - PC

## National survey of college graduates

NATIONAL SURVEY OF COLLEGE GRADUATES

**Save & Exit**

**Part A - Employment Situation**

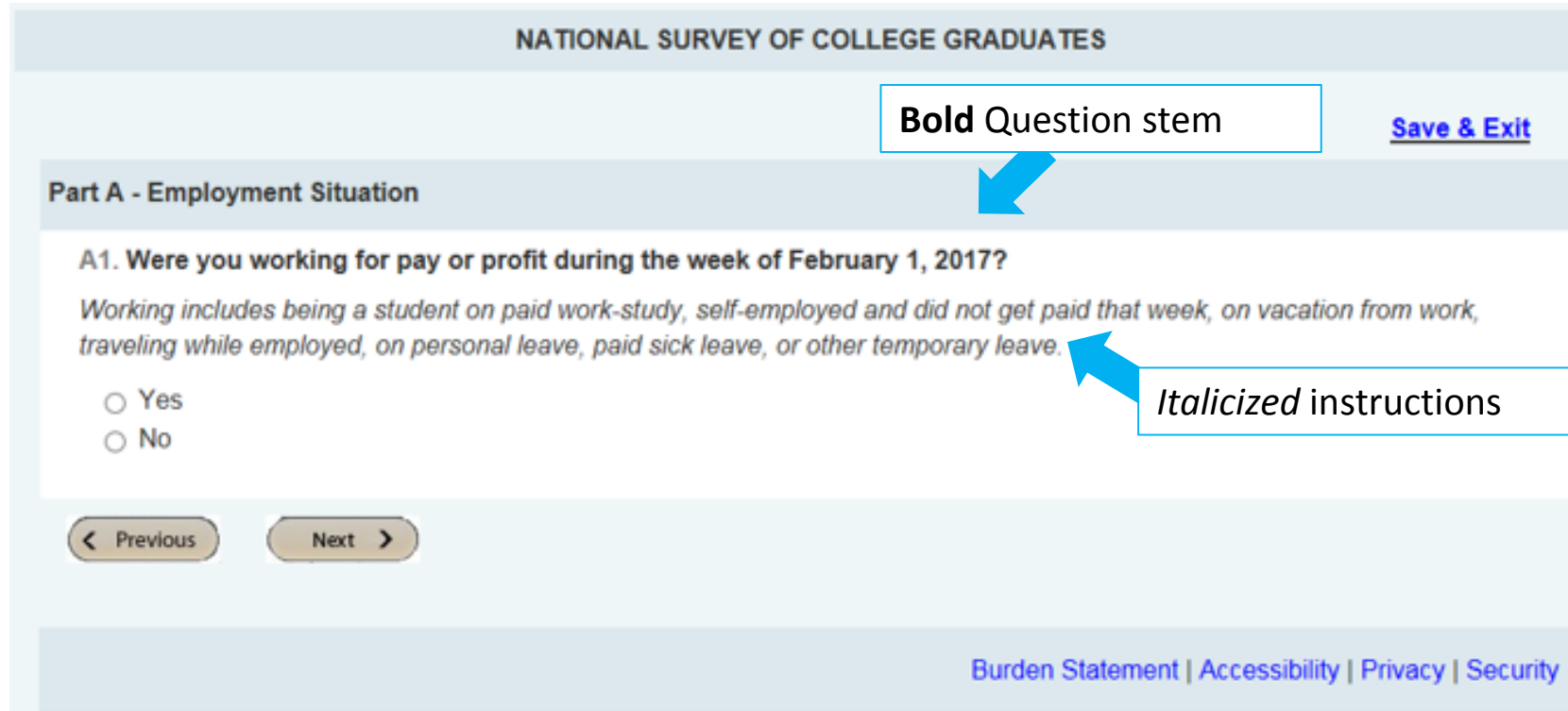
**A1. Were you working for pay or profit during the week of February 1, 2017?**

*Working includes being a student on paid work-study, self-employed and did not get paid that week, on vacation from work, traveling while employed, on personal leave, paid sick leave, or other temporary leave.*

Yes  
 No

[Previous](#) [Next](#)

[Burden Statement](#) | [Accessibility](#) | [Privacy](#) | [Security](#)



# Examples in Practice - Mobile

## American Community Survey

The screenshot displays a mobile browser interface for the American Community Survey. The address bar shows 'acs.respo'. The page title is 'American Community Survey'. The main content area features question 46: 'What were Jane Doe's most important activities or duties? (For example: patient care, directing hiring policies, supervising order clerks, typing and filing, reconciling financial records) (Help)'. A text input field is provided below the question. Two blue arrows point to specific text elements: one points to the bolded question stem 'What were Jane Doe's most important activities or duties?' and the other points to the italicized instructions '(For example: patient care, directing hiring policies, supervising order clerks, typing and filing, reconciling financial records)'. Below the input field are two green navigation buttons with left and right arrows. The bottom of the screen has a green bar with 'Contact Us' and 'Accessibility Privacy' links.

**Bold Question stem**

*Italicized instructions*

# Motivation

- Lack of empirical evidence in support of theory and recommendations.
- Has not been tested for mobile.

# Research Questions

- Does **bolding** question-stems make them more distinguishable and easier to find than other text on a mobile survey?
- Does *italicizing* instructions make them more distinguishable and easier to find than other text on a mobile survey?

# General Method

- Quantitative Controlled Study
- Setting: In-person tests at community/senior centers in Washington DC area from Dec. 2016 to Jan. 2017
- Convenience sample
- Participants ages 60-75
  - Had to have at least one year of mobile phone experience
- Procedure
  - Participant completed background paper questionnaire (demographic questions)
  - Test administrator loaded survey app on iPhone 5s, handed phone to participant, instructed him/her to complete survey
  - 3-5 experiments run during 1 hour session



# Method for this Experiment

- Self administered survey app
  - 5 questions to assess time
  - 2 questions to assess accuracy of responses
  - Satisfaction and preference data
- Between-subjects design: 4 conditions
- 30 participants
  - 7-8 participants per condition/group
  - 14-16 participants per factor level

# Design

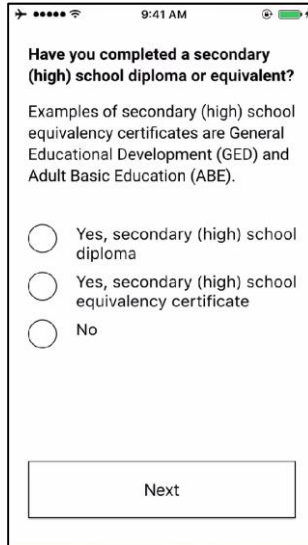
## 2 x 2 Factorial Between-Subjects

		Question stem	
		Bolded	Not Bolded
Instructions	Italicized	<b>Bolded</b> + <i>Italicized</i> ( <i>n</i> = 7)	Non-bolded + <i>Italicized</i> ( <i>n</i> = 8)
	Not Italicized	<b>Bolded</b> + Non-italicized ( <i>n</i> = 8)	Non-bolded + Non-italicized ( <i>n</i> = 7)

# Design (cont.)

## Four Conditions:

A: **Bold** + No Italics



9:41 AM

Have you completed a **secondary (high) school diploma or equivalent?**

Examples of secondary (high) school equivalency certificates are General Educational Development (GED) and Adult Basic Education (ABE).

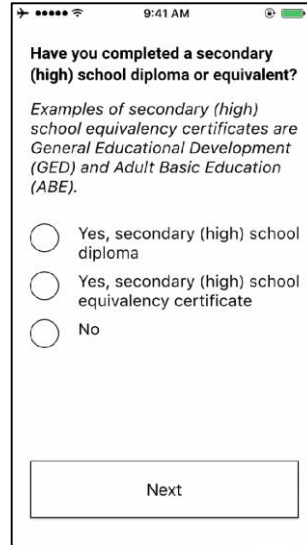
Yes, secondary (high) school diploma

Yes, secondary (high) school equivalency certificate

No

Next

B: **Bold** + *Italics*



9:41 AM

Have you completed a **secondary (high) school diploma or equivalent?**

*Examples of secondary (high) school equivalency certificates are General Educational Development (GED) and Adult Basic Education (ABE).*

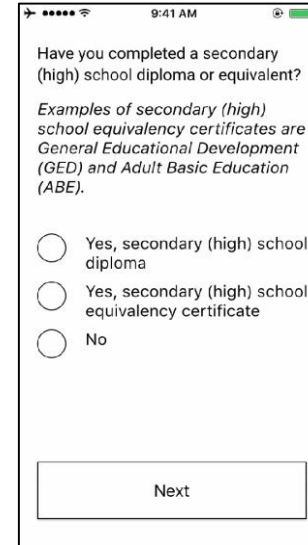
Yes, secondary (high) school diploma

Yes, secondary (high) school equivalency certificate

No

Next

C: No Bold + *Italics*



9:41 AM

Have you completed a secondary (high) school diploma or equivalent?

*Examples of secondary (high) school equivalency certificates are General Educational Development (GED) and Adult Basic Education (ABE).*

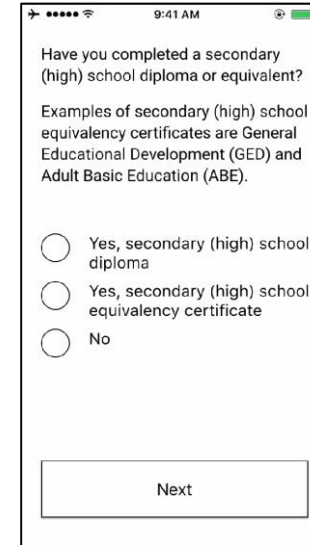
Yes, secondary (high) school diploma

Yes, secondary (high) school equivalency certificate

No

Next

D: No Bold + No Italics



9:41 AM

Have you completed a secondary (high) school diploma or equivalent?

Examples of secondary (high) school equivalency certificates are General Educational Development (GED) and Adult Basic Education (ABE).

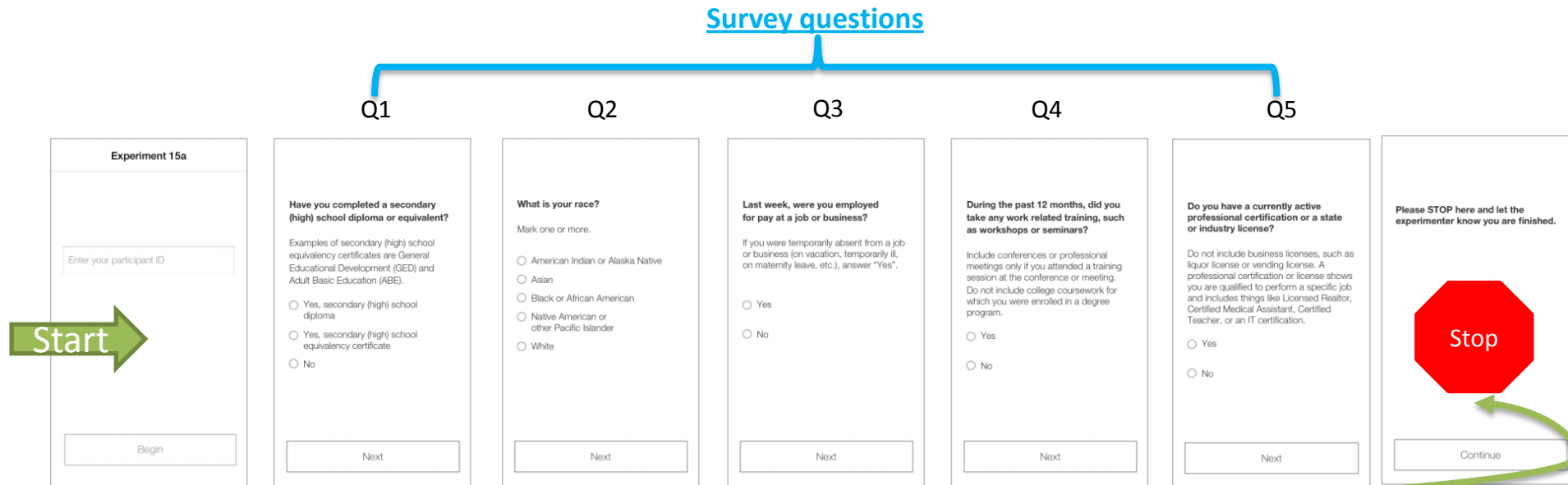
Yes, secondary (high) school diploma

Yes, secondary (high) school equivalency certificate

No

Next

# Experimental Procedure



# Word Recognition

- Indirect measure of how carefully question stem and instructional text was read, if at all.
  - Less words recognized = Less text read
- Word List
  - Participants instructed to circle keywords they remembered seeing
- Rationale: Allows us to address why survey completion times differ and would imply visual filtering of survey content.

## Question 1: Education

Secondary  
Adult Basic Education (ABE)  
Adult Commensurate Education (ACE)  
Primary

## Question 2: Origins

Race  
Mark One or More  
Mark One Only  
Ethnicity

## Question 3: When Employed for Pay

Last Week  
Sabbatical  
Vacation  
Next Week

## Question 4: Past 12 Months Training

Workshop  
Seminars  
Conferences  
Internship

## Question 5: Certifications or Licenses

Certified Medical Assistant  
Certified Practitioner  
Industry License  
Technical License

# Vignettes

- Vignettes provide information needed to answer survey question
- Chose a survey question with complex instructions
  - One correct answer: based on info in vignette & instruction
- Rationale: Allows us to investigate if faster survey completion time comes at a cost to accuracy.

**Vignette 2:** “You live in a home that contains two bedrooms, a bathroom, a hallway, a kitchen, an unfinished basement, and a living room and dining room separated by an archway that extends 5 inches from the wall.”

**How many separate rooms are in your house, apartment, or mobile home?**

Rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling. Include bedrooms, kitchens, etc. Exclude bathrooms, porches, balconies, foyers, halls or unfinished basements.

Number of Rooms:

# Metrics

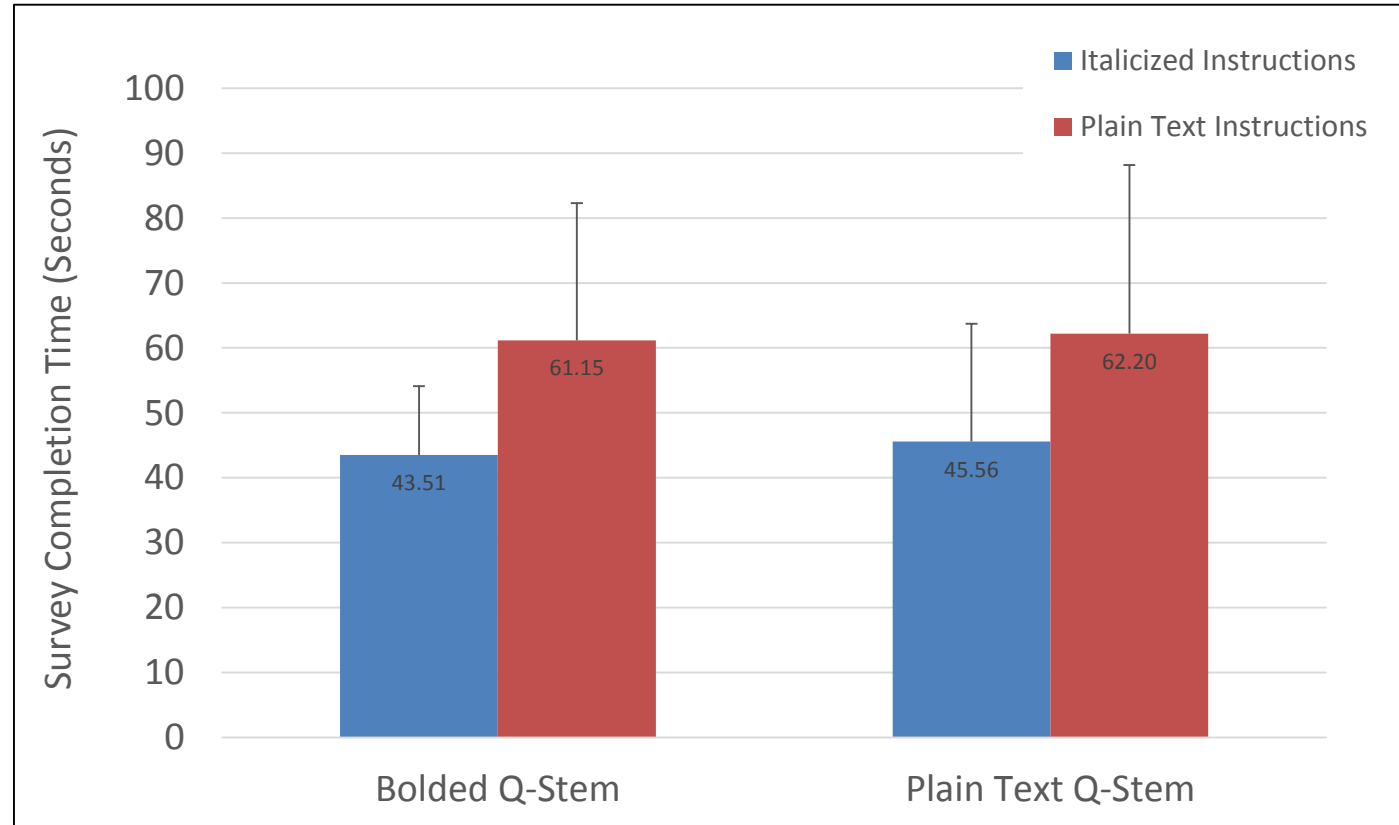
1. Efficiency
  - Survey completion time
2. Effectiveness
  - Response Accuracy
3. Satisfaction
  - Task-Difficulty Rating
  - Overall Preference
4. Other metrics
  - Word Recognition – List of words
    - “Old”: words that appeared in survey
    - “New”: words that did not appear in survey

# Analysis and Results



# Efficiency: Survey Completion Time

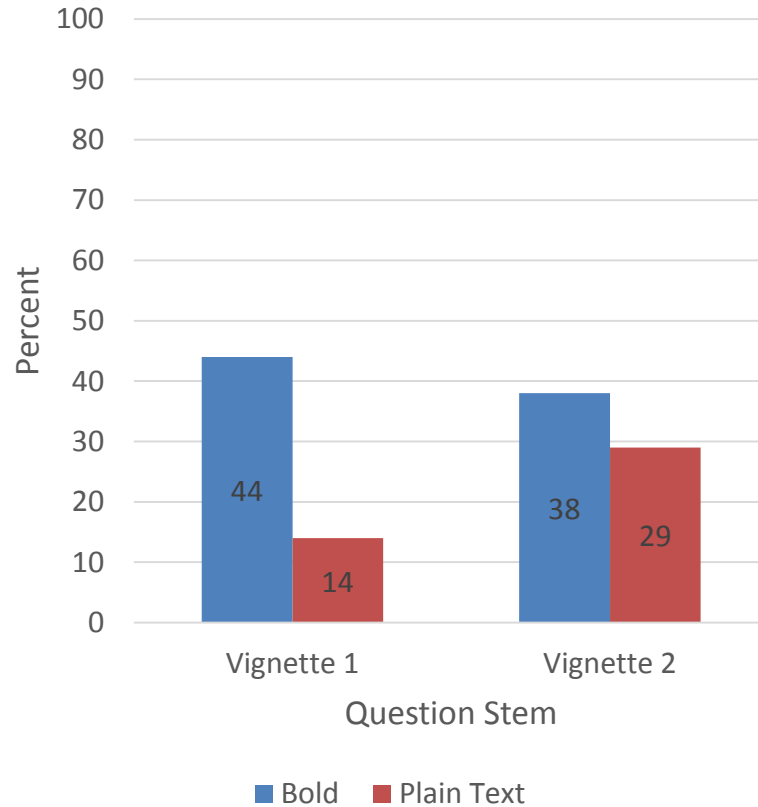
Main effect of instructions: *Italicized* instructions result in faster survey completion times compared to plain text.



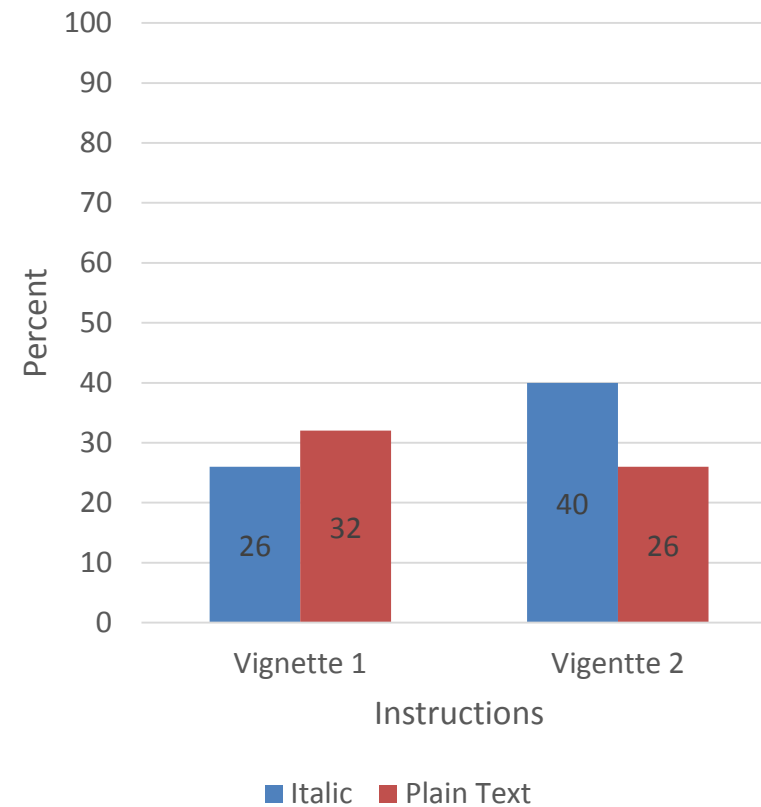
- Bold –  $F(3,26)=.04$ ,  $p = ns$
- Ital –  $F(3,26)=4.94$ ,  $p < .05$
- Bold x Ital –  $F(3,26)= .00$ ,  $p = ns$

# Effectiveness: Accuracy

Response accuracy when the question stem was...



Response accuracy when the instructions were...



**Bold:**

Vignette 1:  $\chi^2(1) = 3.08, p = n.s.$

Vignette 2:  $\chi^2(1) = 0.26, p = n.s.$

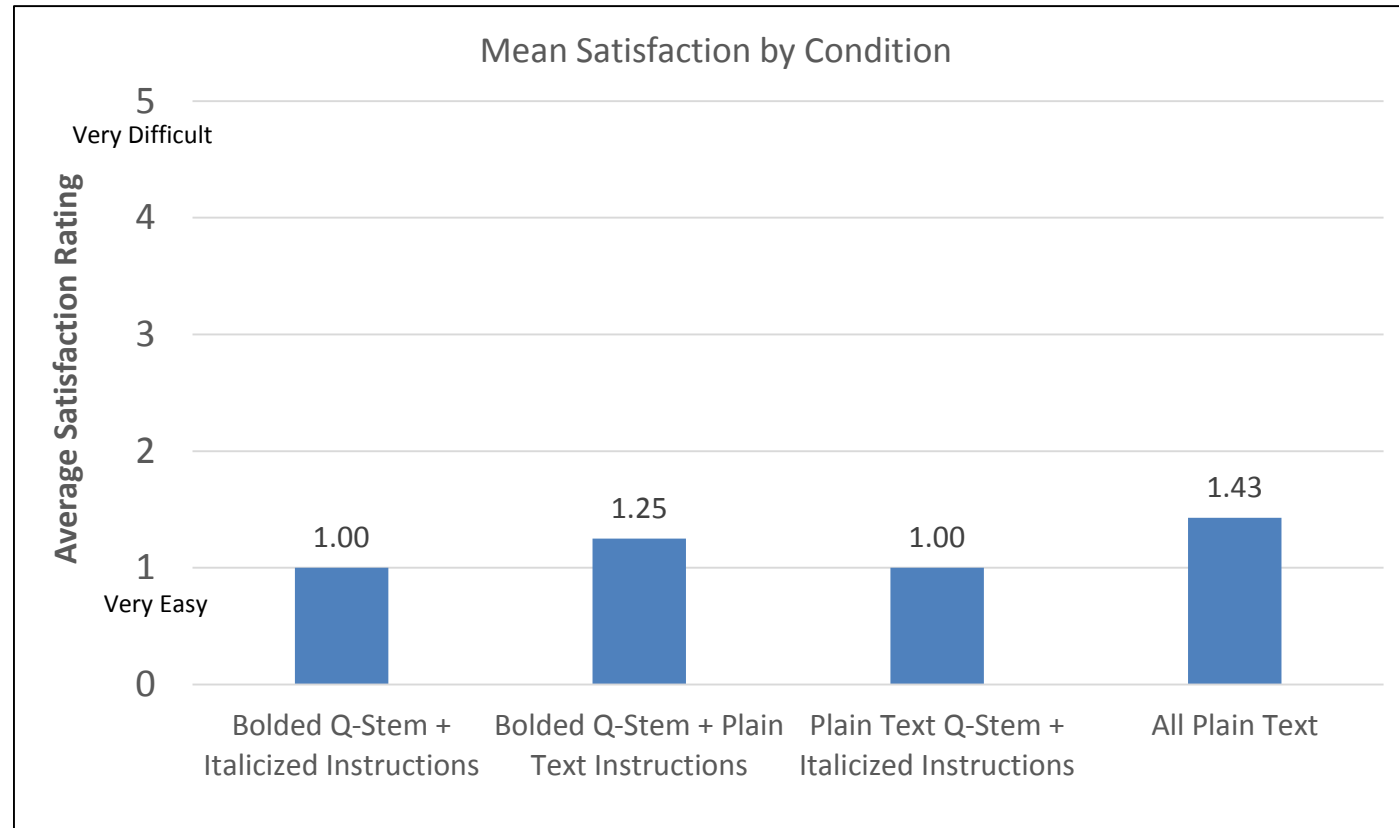
*Italic:*

Vignette 1:  $\chi^2(1) = 0.15, p = n.s.$

Vignette 2:  $\chi^2(1) = 0.60, p = n.s.$

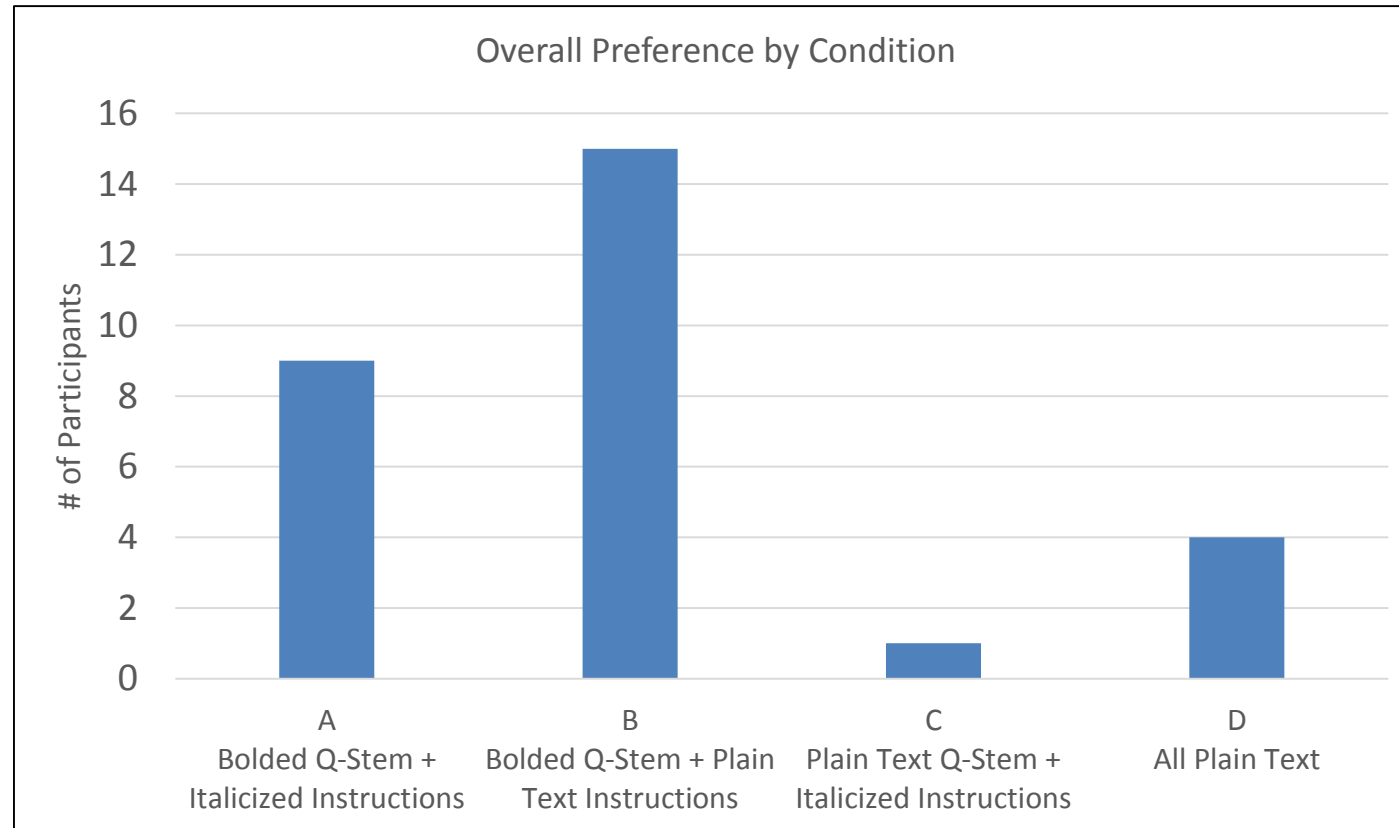
# Satisfaction: Task Difficulty

There were no significant differences in task difficulty ratings between conditions.



# Satisfaction: Overall Preference

- Participants overwhelmingly preferred conditions with a **bolded** question stem.
- Popular comment: “Easy to read”



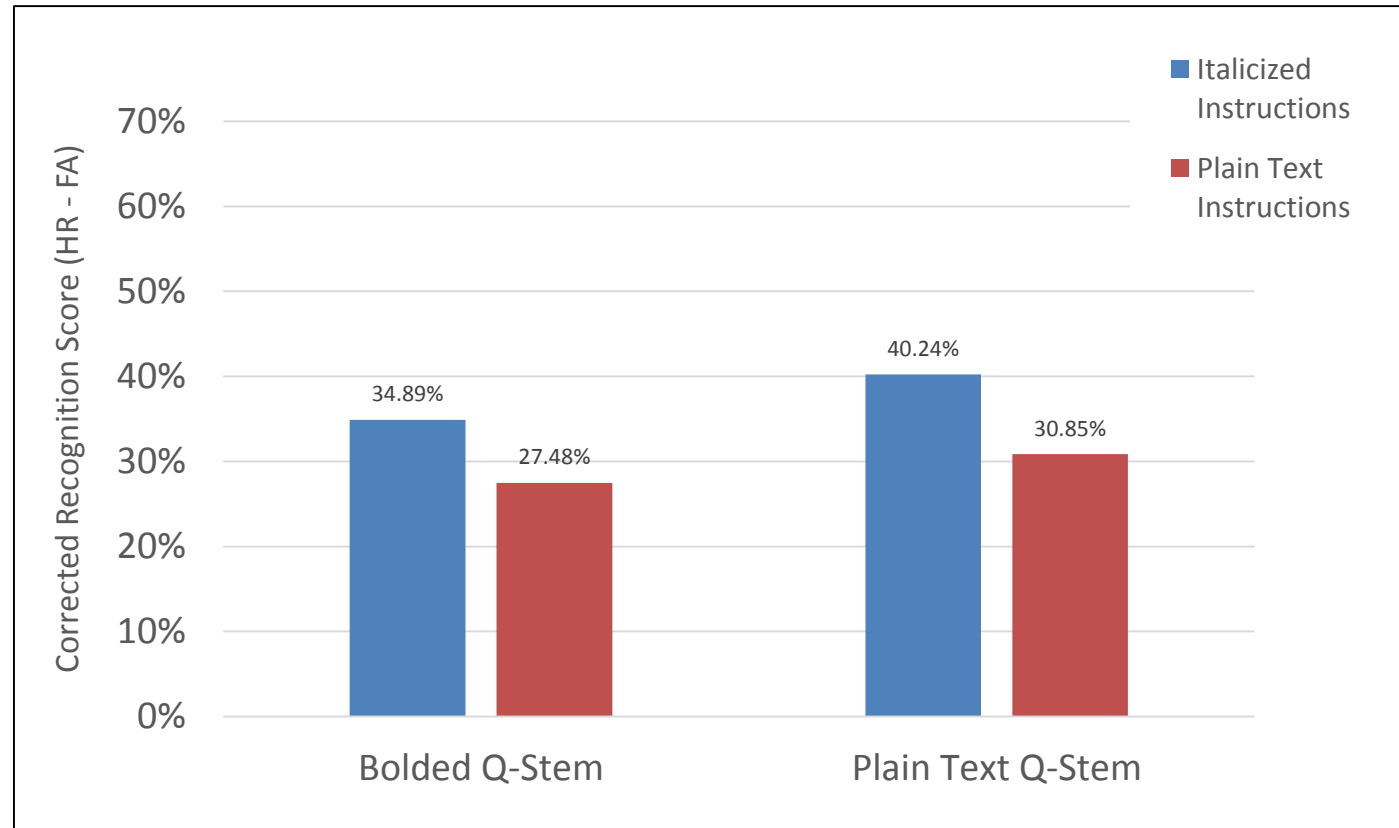
$\chi^2(3) = 14.03, p < .01$

Pairwise comparisons:

- A vs C and D,  $p < .05$
- B vs C,  $p < .05$

# Other Metric: Word Recognition

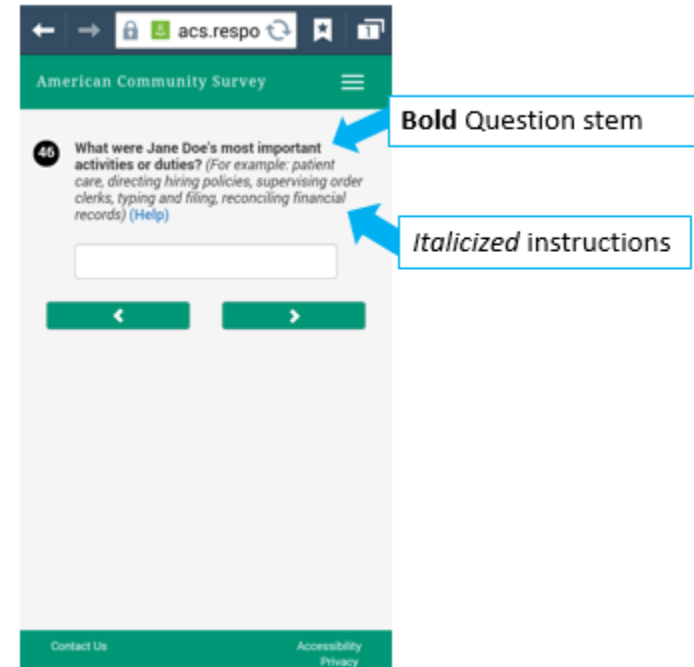
Neither **bold** q-stem nor *italicized* instructions significantly affected the amount of text read.



- Bold -  $F(3,26) = .26, p = n.s.$
- Italics -  $F(3,26) = .97, p = n.s.$
- Bold x Ital -  $F(3,26) = .01, p = n.s.$

# Recommendation

- *Italicize* instructions – This resulted in faster overall survey completion time
  - Speed benefit did not come at a cost to response accuracy
  - No evidence found for skipped instructions via filtering out of irrelevant text
- **Bold** question stems – Strongly preferred by study participants
  - No benefit or cost to speed or accuracy
- Some empirical support for the common practice of *italicizing* instructions and **bolding** q-stem



# Limitations

- Our word recognition measure may not be sensitive enough to uncover the mechanism underlying faster survey completion times for *italicized* instructions due to high task difficulty.
- Accuracy results were based on an artificial task (vignettes)

# Future Directions

- Eye-Tracking to investigate filtering of instructional text
- Test scrolling design to test possible benefits of **bolding** q-stem on mobile device survey

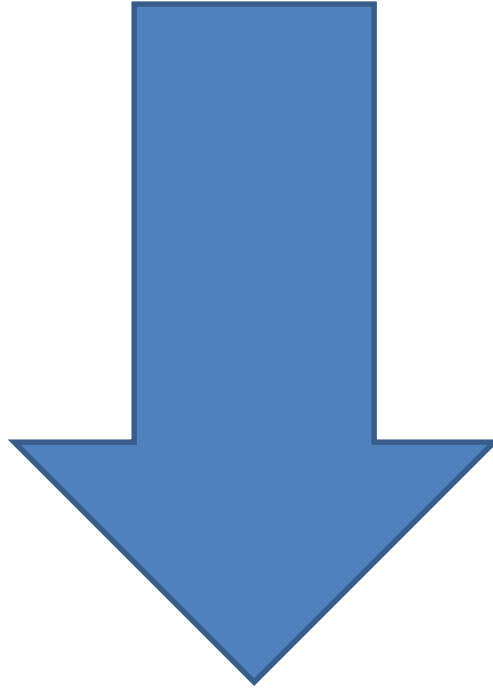


# Questions?

Brian Falcone

Brian.Falcone@census.gov

# Extra Slides

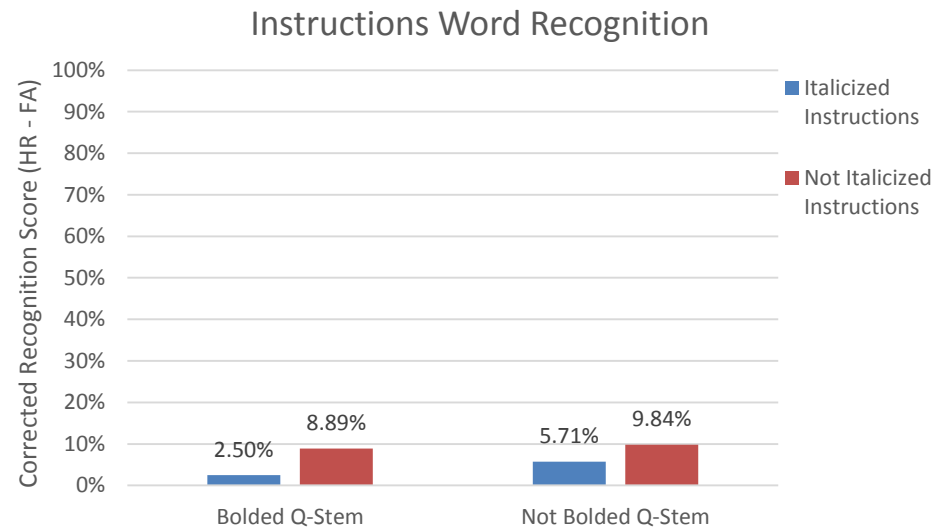
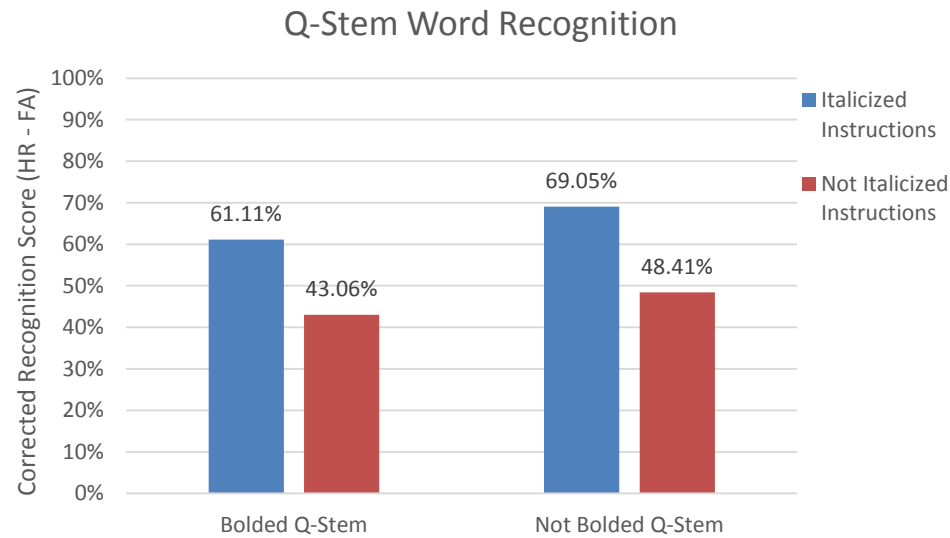


# Word Recognition: The Details

- 11 “old” words, 9 “new” words
- Hit Rates (Recognition)
  - % of “old” words circled
- False Alarm Rates (Guessing)
  - % of “new” words circled
- Measure Used: Corrected Recognition Score
  - Hit Rate minus False Alarm Rate

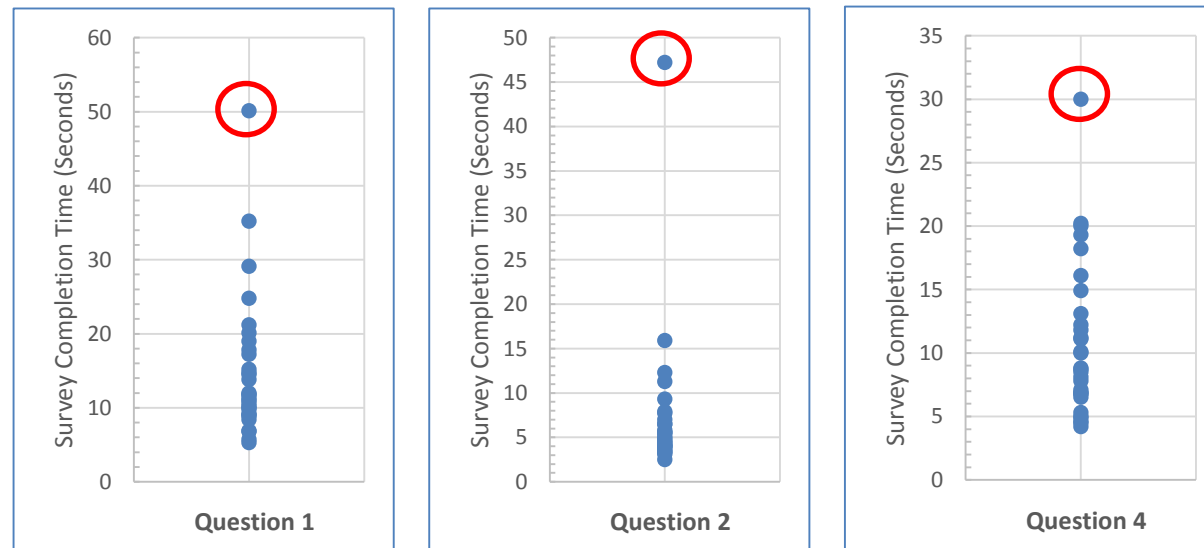
# Word Recognition - Results

- No significant effects of Bold or Italics on recognition of words from the question stem or instructions



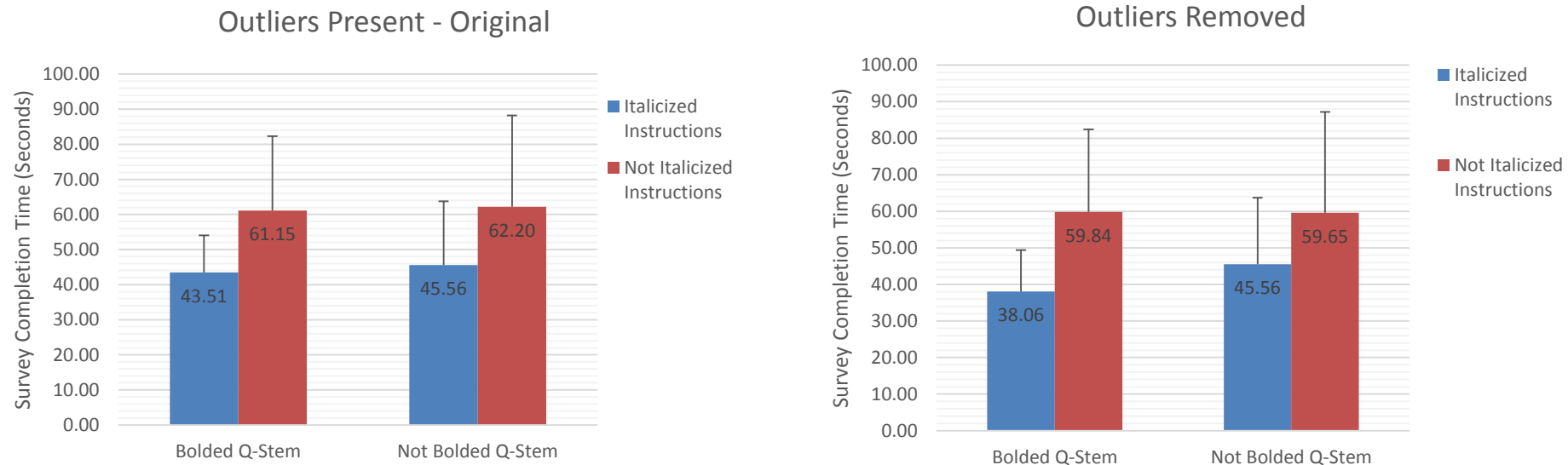
# Outliers

- 3 participants talked to TA during task
- Common outlier detection methods using conservative criteria were used to determine the need for sensitivity analyses:
  - 3 of 3 were  $>3$  SDs above mean
  - 2 of 3 were  $>3$  IQRs (Inter-Quartile Ranges) outside range



# Sensitivity Analysis

- Main effect of italicized instructions remains significant:
- Subjects Removed – 2 x 2 ANOVA,  $p=.03$



- Additional Confirmatory Analyses:
  - Mean imputation – 2 x 2 ANOVA,  $p=.03$
  - Nearest neighbor – 2 x 2 ANOVA,  $p=.01$
  - Only question-level data removed – Random Effects Analysis,  $p=.02$

# Satisfaction by factor

