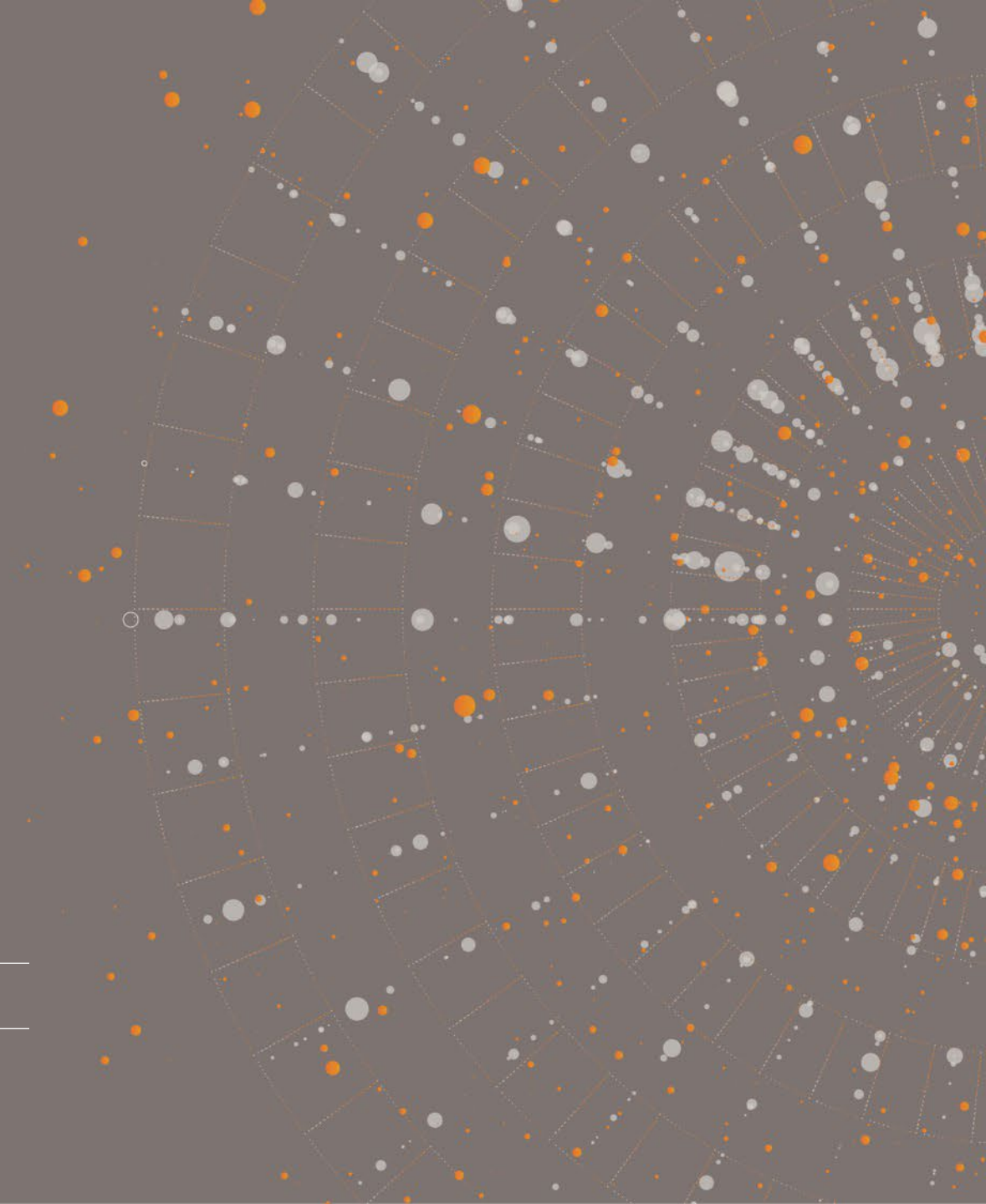# UX Principles for Integrating Textbots into Surveys

2024 FCSM Research and Policy Conference

October 24, 2024

Elizabeth Dean, Soubhik Barari, Brandon Sepulvado, Leah Christian

Why Textbots?

(and why not chatbots?)

Conversational **Generative AI** agents may improve self-administered web surveys.

1. **More informative open-ended responses**

Probe respondents to **elaborate** on their responses in real time.

2. **Increase response quality**

Follow up on insufficient, inconsistent, or missing answers.

3. **Reduce burden of complicated questions**

Support recall (residential history, medications) with a conversational structure.

A conversational approach may create higher quality survey experiences for respondents.

1. **Easier to complete**

Respondents may be more likely to stay in survey if questions can be tailored to them via AI.

2. **Better user experience**

Respondents may experience less frustration if they can ask clarifying questions, receive a personalized experience.

Chatbot

- Customer service tool
- Used when you're trying to make a complaint
- Tends to be static LLM (not generative AI)
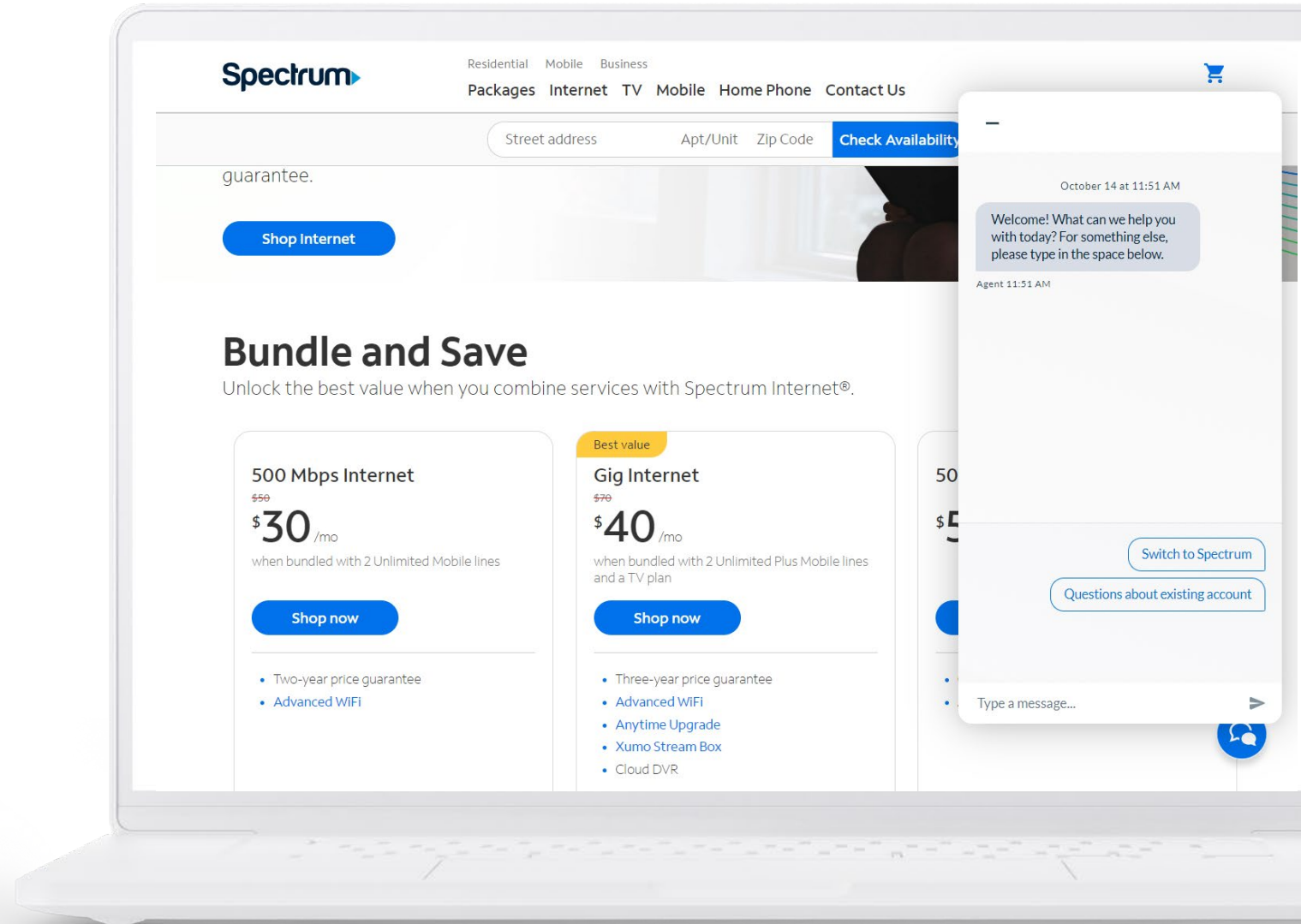
Textbot*

*working title

- Neutral term
- Avoids personification of AI
- Refers to a dynamic LLM using generative AI

What do you think is the most important problem facing this country today?

inflation

That's an interesting answer. Can you tell me more about why you think inflation is the most important problem facing the country today?

inflation affects ordinary peoples' everyday lives and may decrease quality of life

Textbots offer a new mode for conversational interviewing

1. Allowing follow-up questions (from interviewer and respondent) reduces comprehension errors.
2. Elaboration of answers facilitates more accurate coding of responses.
3. Textbots improve upon early attempts to mimic CI in web surveys:
   1. Routing and triggers for follow-up questions
   2. Hyperlinks with explanation
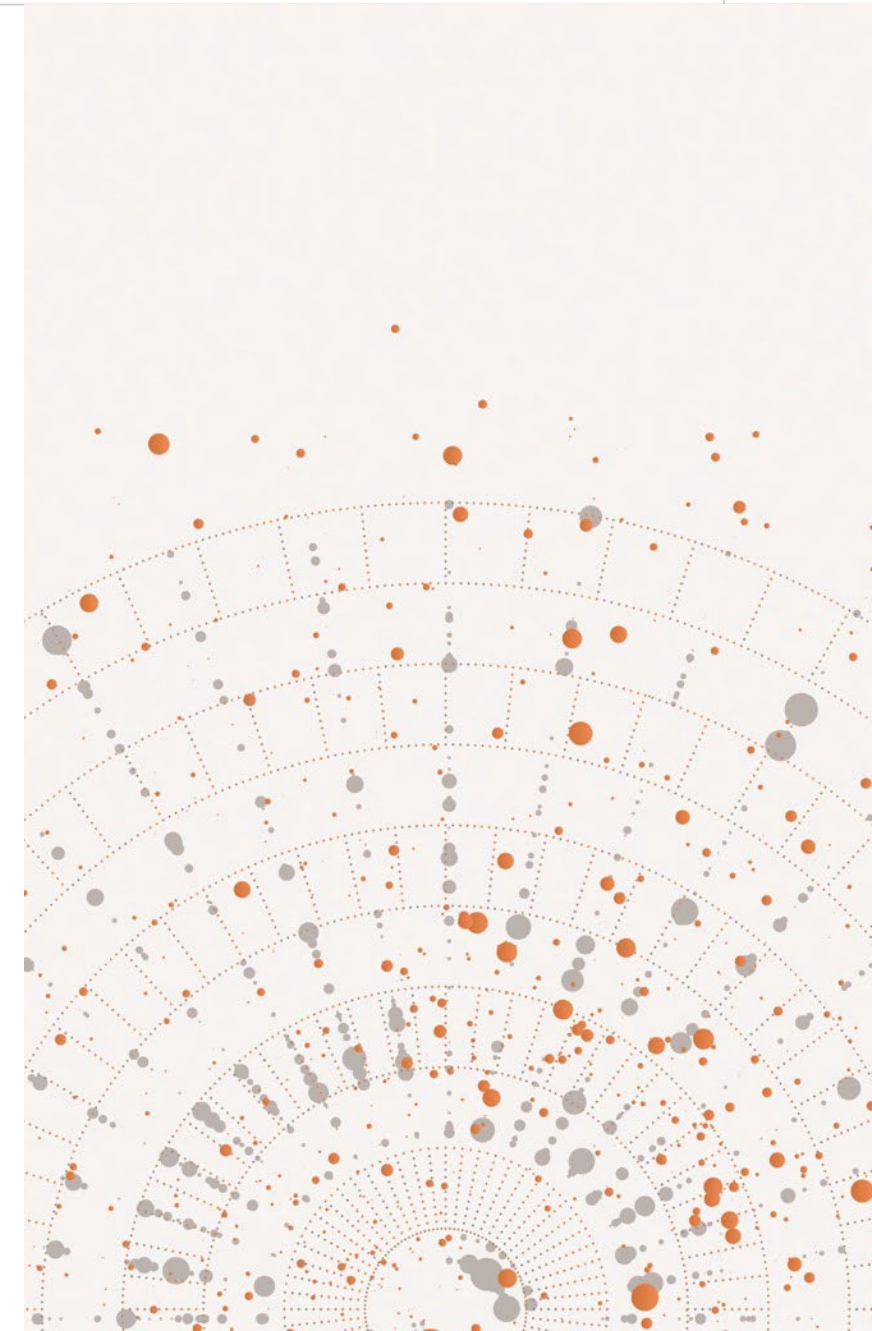   3. Integration of avatars and audio

Case Study

How do survey panel respondents react to a textbot-based survey?

Case study of findings from larger experiment

- 1800 respondents

- Three different textbot surveys designed to assess quality of responses to 3 different types of probes

- Assessed breakoffs & respondent perceptions of experience

Elaboration reprobes

Control (Status Quo)

What do you think is the most important problem facing this country today?

inflation

Treatment (Gen AI)

What do you think is the most important problem facing this country today?

inflation

That's an interesting answer. Can you tell me more about why you think inflation is the most important problem facing the country today?

inflation affects ordinary peoples' everyday lives and may decrease quality of life

Relevance reprobes

Control (Status Quo)

What do you think is the most important problem facing this country today?

asdghjasdfglasdf

Treatment (Gen AI)

What do you think is the most important problem facing this country today?

asdghjasdfglasdf

I didn't understand. Do you mind rewording what you said?

crime

Live classification with confirmation reprobes
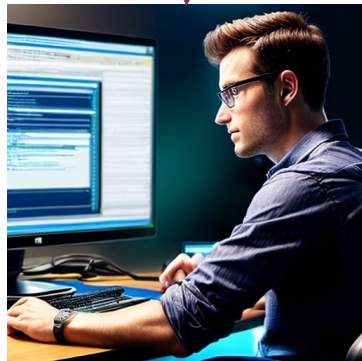
Control (Status Quo)

What do you think is the most important problem facing this country today?

too much inflateion. grocery prices r too high

cost of living

Treatment (Gen AI)

What do you think is the most important problem facing this country today?

too much inflateion. grocery prices r too high

I interpreted your answer as generally being about **cost of living**. Is that correct?

Yes   No

cost of living

- Recruitment from Prodege (a non-probability web panel provider)

- Used the Inca chatbot, which is a product of Nexxt Intelligence

- Three treatments – each programmed as own survey

|	Treatment 1 (Control)	Treatment 2	Treatment 3
Description	No reprobes of any kind	Elaboration or Relevance reprobes only	Confirmation reprobes only
Sample size	n = 600	n = 600	n = 600

- Minimal additional tuning / prompting of textbots for probing or detection

- Data collected July 15th – 17th, 2024

UX Results

Respondent experience was captured in measures of burden and satisfaction.

UX Metrics

- Breakoffs

- Self-reported experience measures:
  - Quality of survey responses
 - Ease of completion
  - Frustration
  - Overall satisfaction
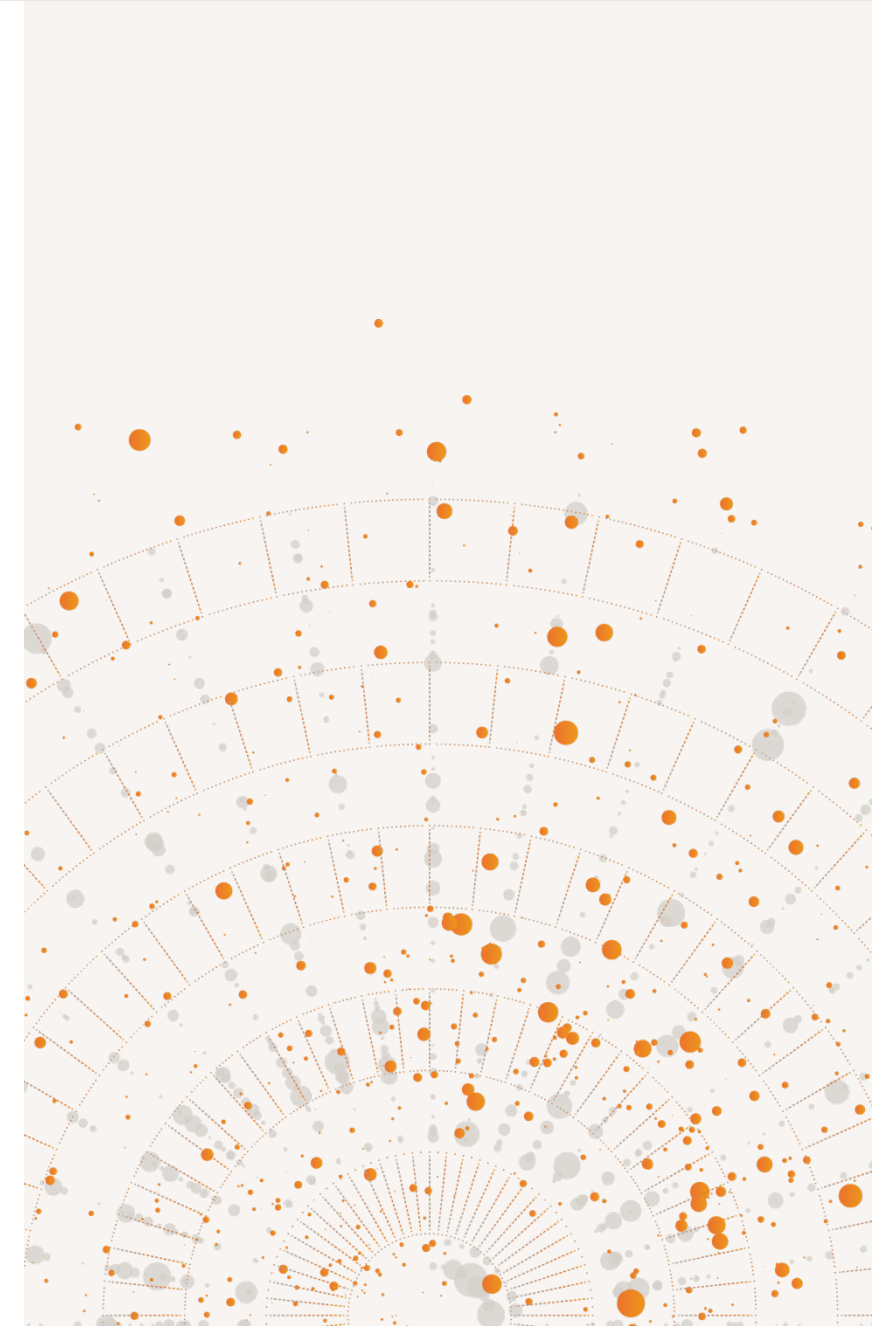
Breakoff is concentrated in early questions and among mobile respondents

- Q1 most important issue

- Q2 perception of economic conditions

- 1-4% of sample per question

- Respondents were more likely to drop out after the elaboration and relevance probes than the confirmation probes.

- Almost all of these breakoffs were on mobile
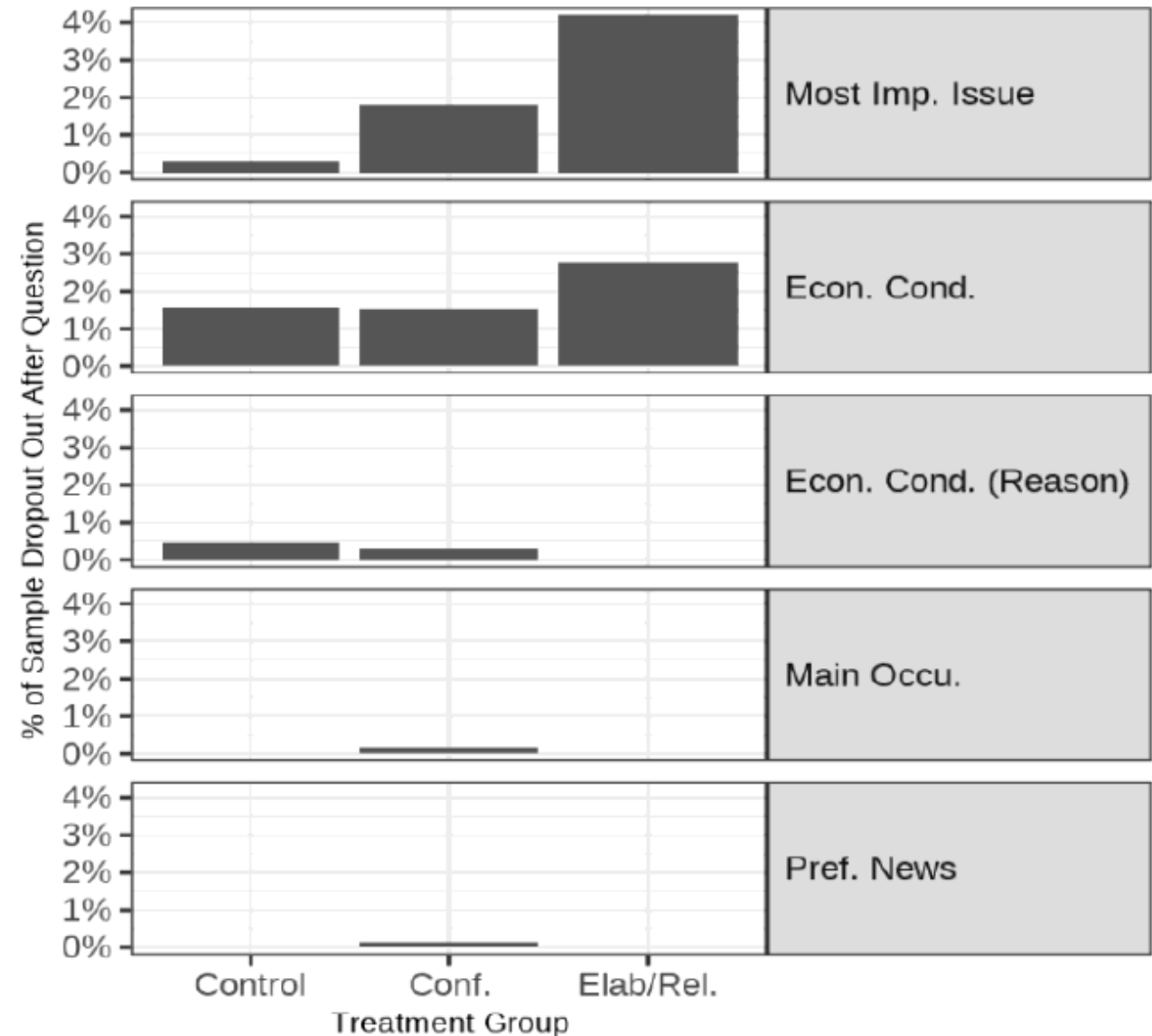
Probing slightly increases self-reported frustration, decreases satisfaction
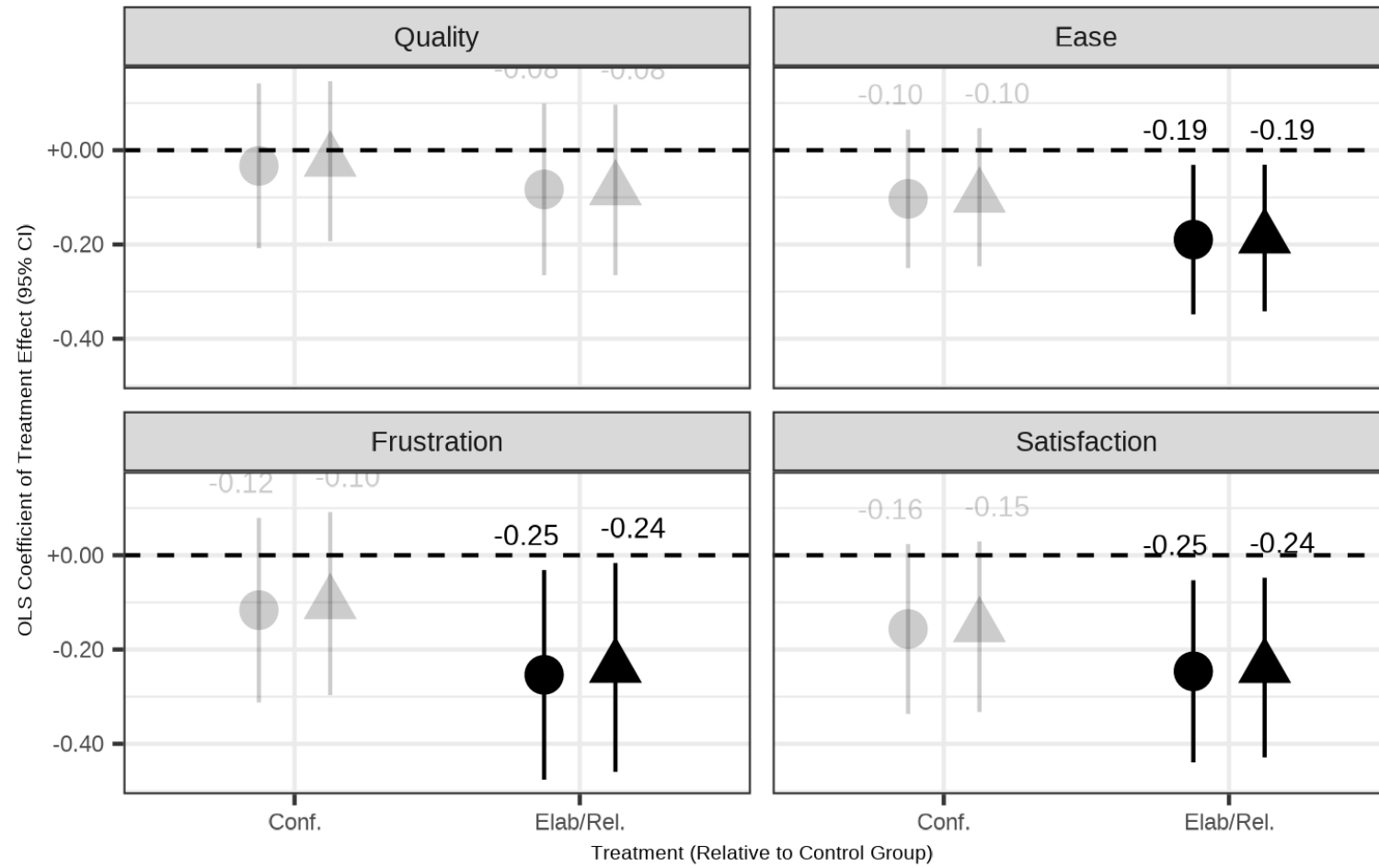
UX Principles

For Future Textbot Research

NORC
at the
University of
Chicago

Textbots can provide useful survey data with minimal respondent burden.

Findings

- Probes yielded accurate codes and more detailed, better-quality responses.

- Textbot surveys are minimally burdensome
  - They are worse for mobile respondents

- Respondents reported slightly more negative responses to ease of use, frustration, and satisfaction measures

Spoiler alert!

Basic UX Principles for Textbots (or Chatbots)

1. **Tell respondents what you're doing and why.**

    - "For some questions, this survey might use a chatbot to make answering the questions easier for you and to help us get a more accurate understanding of your experience."

2. **Limit usage of textbots to a subset of questions**

3. **Use buttons; textbot questions don't have to be open-ended**

4. **Use a chat format when a conversation makes sense**

    - True open-ended questions

    - Items where a conversational approach helps recall
        - Medical history, residential history, pregnancy and childbirth

5. **Optimize for mobile users**

    - Buttons, voice entry

Future Research

1. **Compare textbots to static web surveys**

   - Open-ended questions

   - Radio buttons vs. chatbot buttons

2. **Evaluate perceived vs actual burden**

   - Research on medical chatbots found that patients thought they were faster than a static web survey but the response time was the same.

3. **Apply a standardized UX metric like the System Usability Scale**

Thank you.

Elizabeth Dean
Principal Research Methodologist
dean-elizabeth1@norc.org

Research You Can Trust™

NORC at the University of Chicago