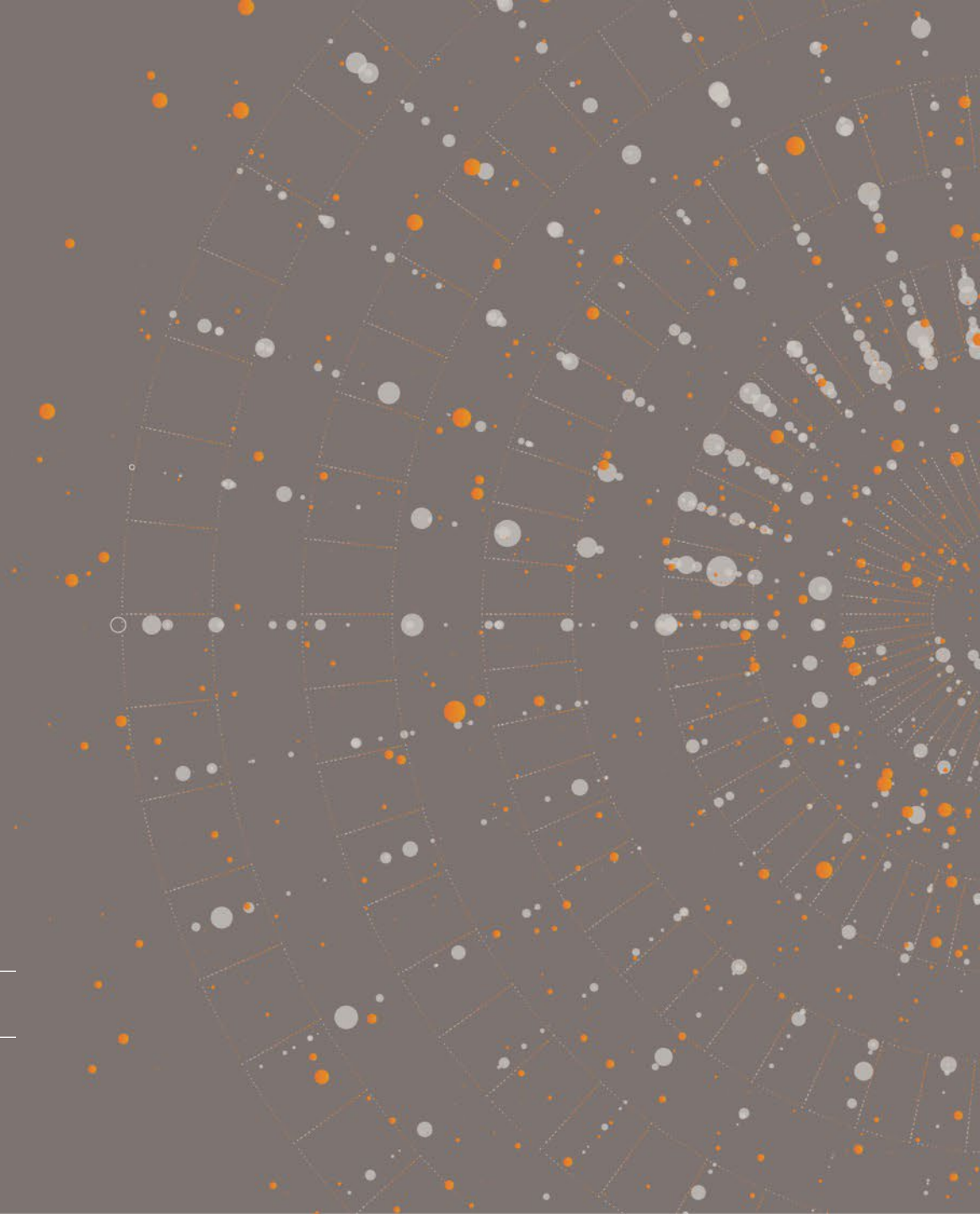
NORC at the University of Chicago

Who Are the Careless Web Respondents Identified by Machine Learning?

FCSM 2024

Ting Yan, Gizem Korkmaz, David Cantor



This work was supported by a grant awarded by the National Science Foundation [NSF-2050809] to Ting Yan (PI) and David Cantor (Co-PI).

Web surveys are a popular mode of data collection with known data quality issues

Representation

- **Coverage when used in a single web mode design**
- **Nonresponse when used in a single mode or multimode design**

Measurement

- **Measurement error when used in a single mode or multimode design**
  - Two types of web respondents of concern

Fraudulent respondents (Kennedy et al. 2021; Puleston, 2019)

- **Bot**
- **People living outside targeted area (or fake respondents)**
- **Duplicate IPs/Multi-completers**
- **Ghost respondents**

Careless respondents (Kennedy et al., 2021; Puleston, 2019; Jones et al., 2015)

- **Also called inattentive, insincere, bogus, satisficing respondents**
- **Do not read questions carefully, do not spend time and effort to carefully answer questions, multitask, not motivated**
  - Focus of this talk

During and/or after data collection

- **Attention checks, instructional manipulation checks, traps (Gummer et al., 2021)**
- **Speeding (Conrad et al., 2017)**
- **Low-incidence questions, inconsistent answers (Jones et al., 2015)**
- **Proxy indicators of data quality examined alone or together**
  - Straightlining, extreme responses, midpoint, acquiescence, missing data
  - Open-ended questions (Kennedy et al., 2021)
  - Response entropy (Tawa, 2021)

National Study of Social, Economic, and Health Experiences (NSSEHE)

- Tracks changes in opinions, lifestyle, and health of Americans
- Experiments to investigate mechanisms accounting for panel conditioning

A sample of 8000 registered voters in two states

- Invited to participate in four waves of web surveys through mailings, emails, and text messages

Fourth wave data collection between February 2023 to March 2023

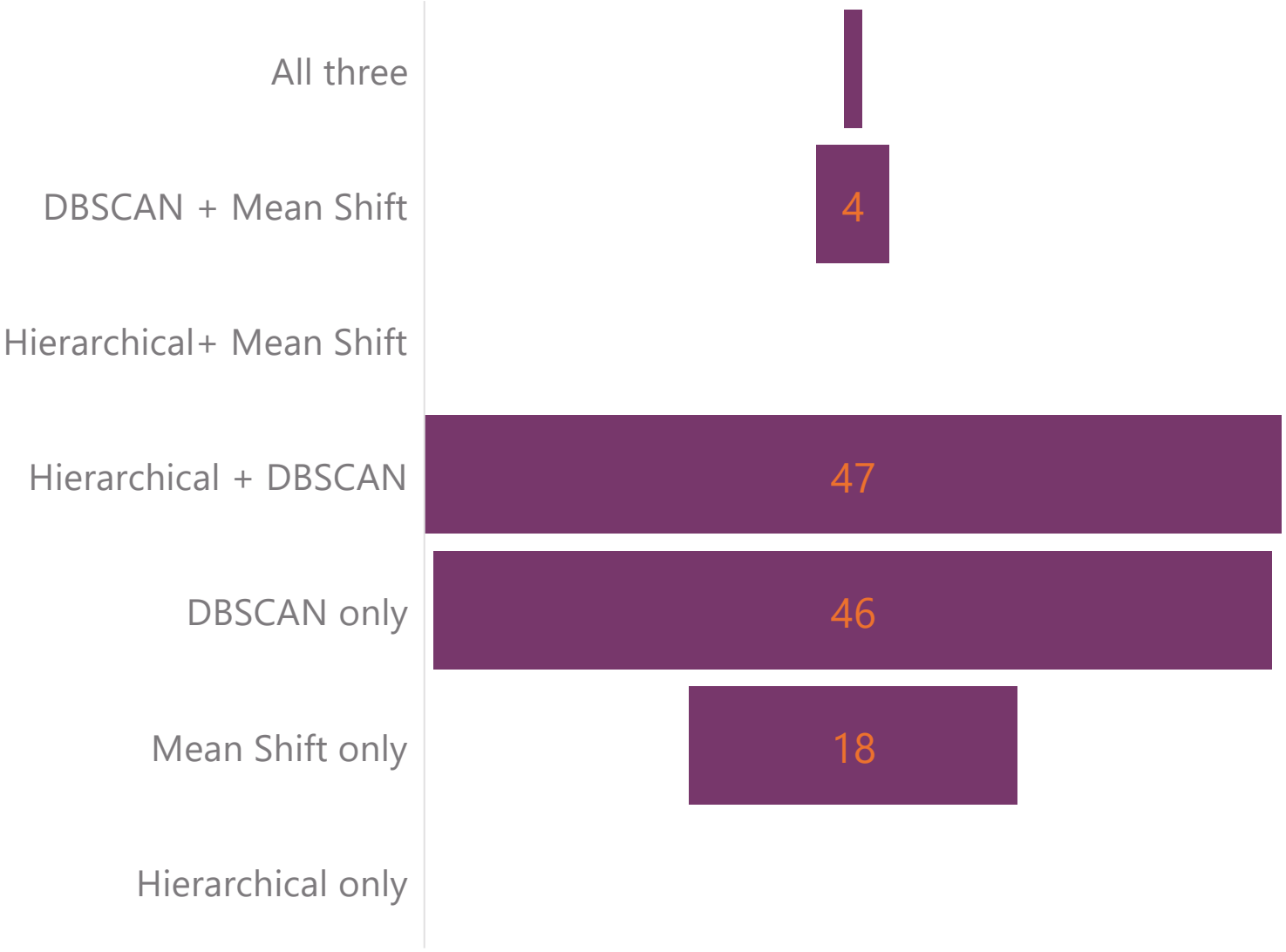
- A total of 947 completes at a response rate of 71.4% (conditional on completing the first wave)

Input Variables Used in Clustering Methods	% Rs Who Would be Flagged as Careless Respondents
Whether or not R failed trap questions	5% failed at least one trap question
Whether or not R reported multitasking	25% reported multitasking
Whether or not R answered too fast	5% fastest
Item nonresponse rate	7% with item nonresponse rate >=5%
Extreme response rate	8% with extreme response rate>=50%
Middle response rate	1% with middle response rate >=50%
Response entropy	10% with largest and smallest 5%

Clustering Methods	% Rs Identified as Careless Respondents
Hierarchical Clustering	5% (n=48) -Speeding
Density-Based Spatial Clustering of Applications with Noise (DBSCAN)	10% (n=98) -Failing both trap questions; Speeding; Item nonresponse rate
Mean Shift	2% (n=23) -Failing trap questions; Item nonresponse rate; Response Entropy
K-Means	51% (n=482) -Multitasking; Middle responses;

Among 947 respondents

- 52 flagged by 2+ methods
- 64 flagged by 1 method
- 831 not flagged

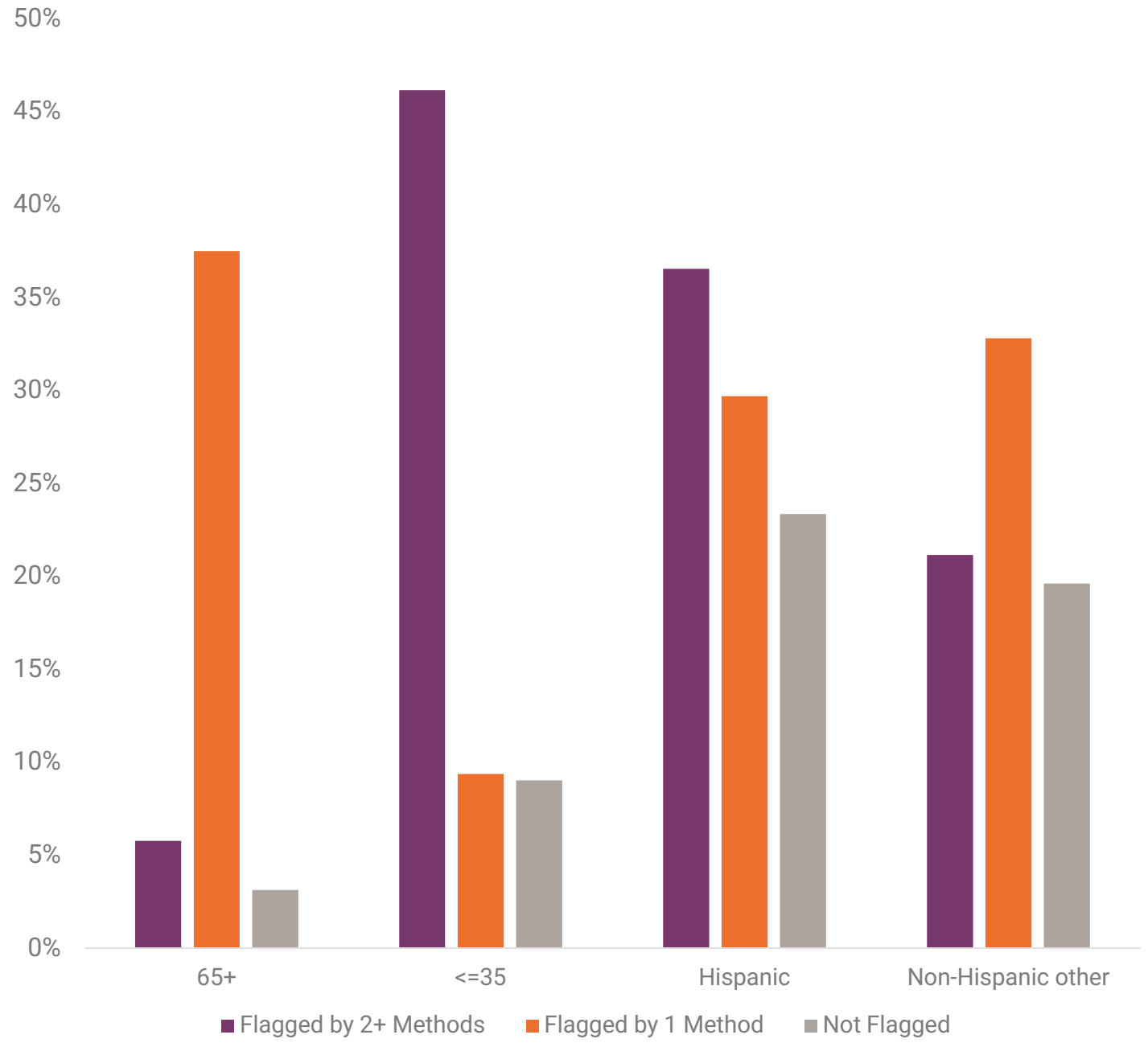


Who were flagged as careless respondents?

- Demographic characteristics related to undesirable response behaviors
- Response behaviors in prior waves
- Perception of burden of prior interviews

Careless respondents

- Older
- Younger
- Hispanic
- Non-Hispanic Other

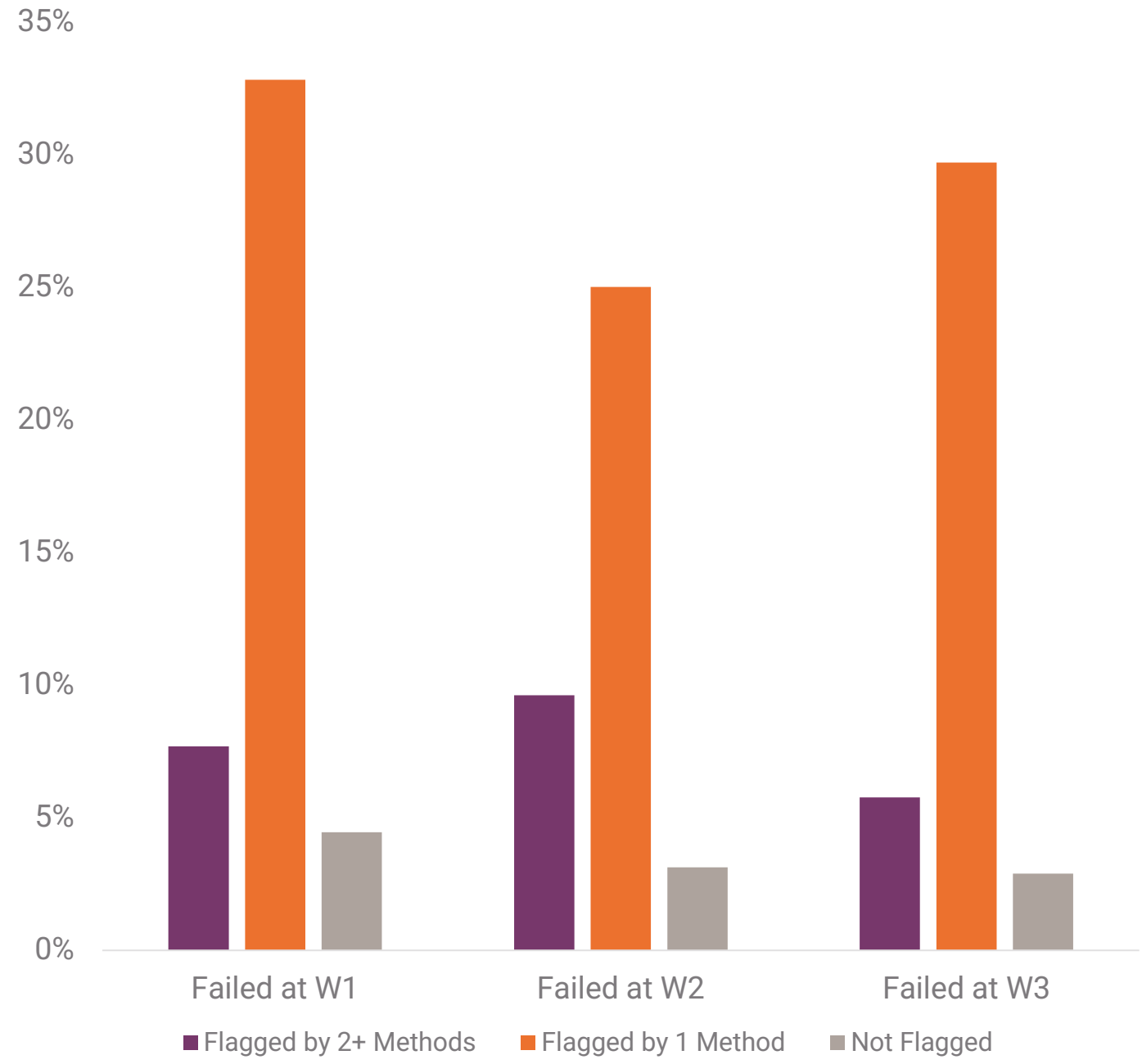


Careless respondents

- Failed attention checks at three prior waves

Paying attention and reading the instructions carefully is critical. If you are paying attention, please select "slightly worried".

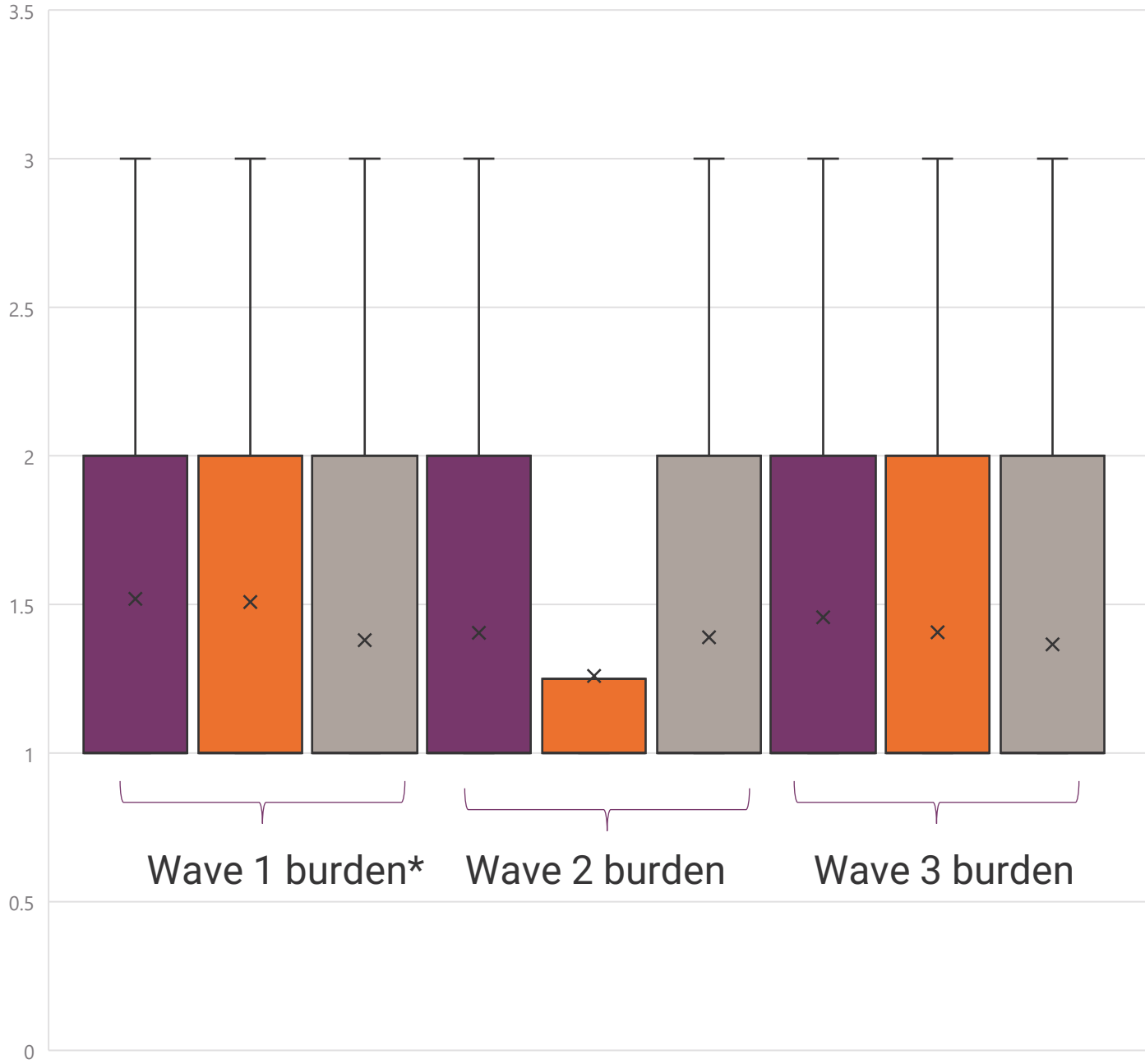
1. Extremely worried
2. Very worried
3. Somewhat worried
4. Slightly worried
5. Not at all worried



Perceived burden didn't differ

- Except for wave 1 with sig. higher burden rating for careless respondents

Overall, how burdensome was this survey to you?
1 Not at all burdensome
2 A little burdensome
3 Somewhat burdensome
4 Very burdensome



Machine learning clustering methods identified careless respondents

We found that careless respondents

- **More likely to be 35 or younger, 65 or older, Hispanic, and other racial categories**
- **More likely to fail attention check**
- **Had higher burden rating in wave 1**

Future research

- **Applying clustering methods to wave 1 data**
- **Using clusters in adaptive design**
  - Different protocol, intervention
- **What about impact on survey estimates?**

Thank you!

yan-ting@norc.org