Please Answer This Important Question: Can Web Prompts Reduce Item Nonresponse?

Robin Kaplan and Erica Yu
Office of Survey Methods Research
U.S. Bureau of Labor Statistics
AAPOR
May 12, 2023

Imagine you were asked…

“What has been your household’s usual monthly expense for alcohol, including beer and wine, to be served at home?”

Imagine you were asked…

"What are all the things you have done to look for work during the last 4 weeks?"

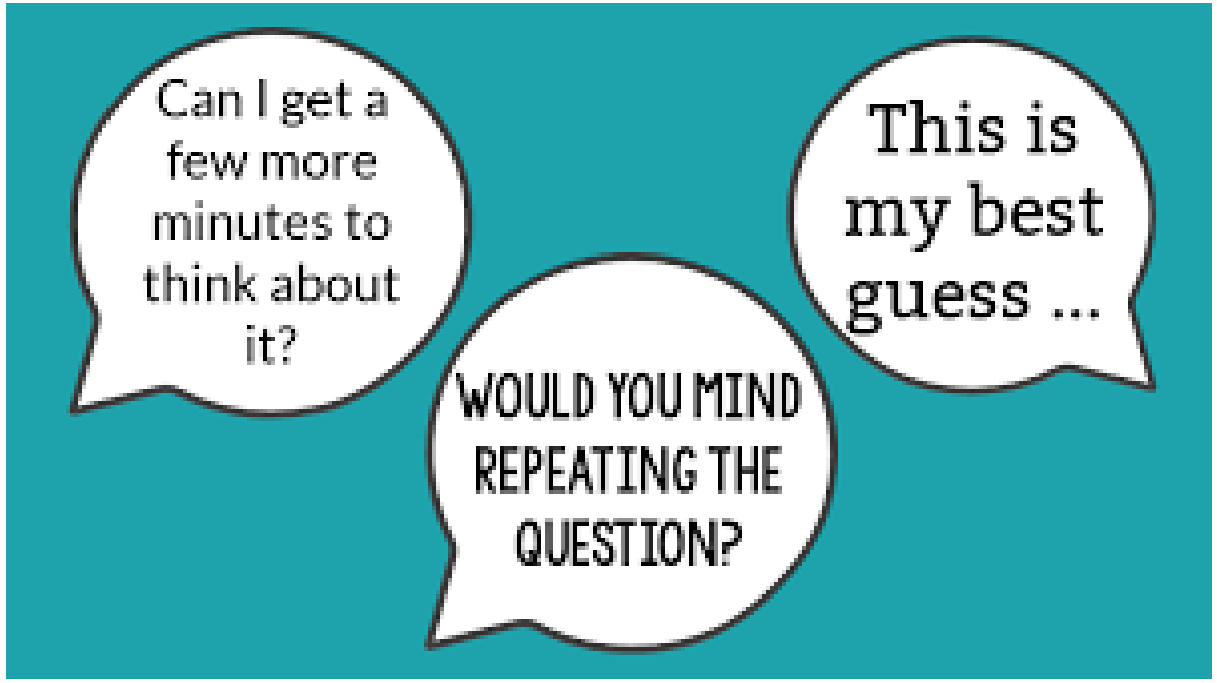
"What are all the things your housemate has done to look for work during the last 4 weeks?"

Imagine you were asked…

"Last week, how many hours did your housemate actually work at his/her main job?"

"Last week, how many hours did you actually work at your main job?"

Your Answer?


Can I get a few more minutes to think about it?
WOULD YOU MIND REPEATING THE QUESTION?
This is my best guess ...


I don't know

What do interviewers *really* do?

- “Your best guess is fine”
- “We are not looking for a precise value, just a category”
- “Your response is helpful because…”

Kaplan & Yu (2020)

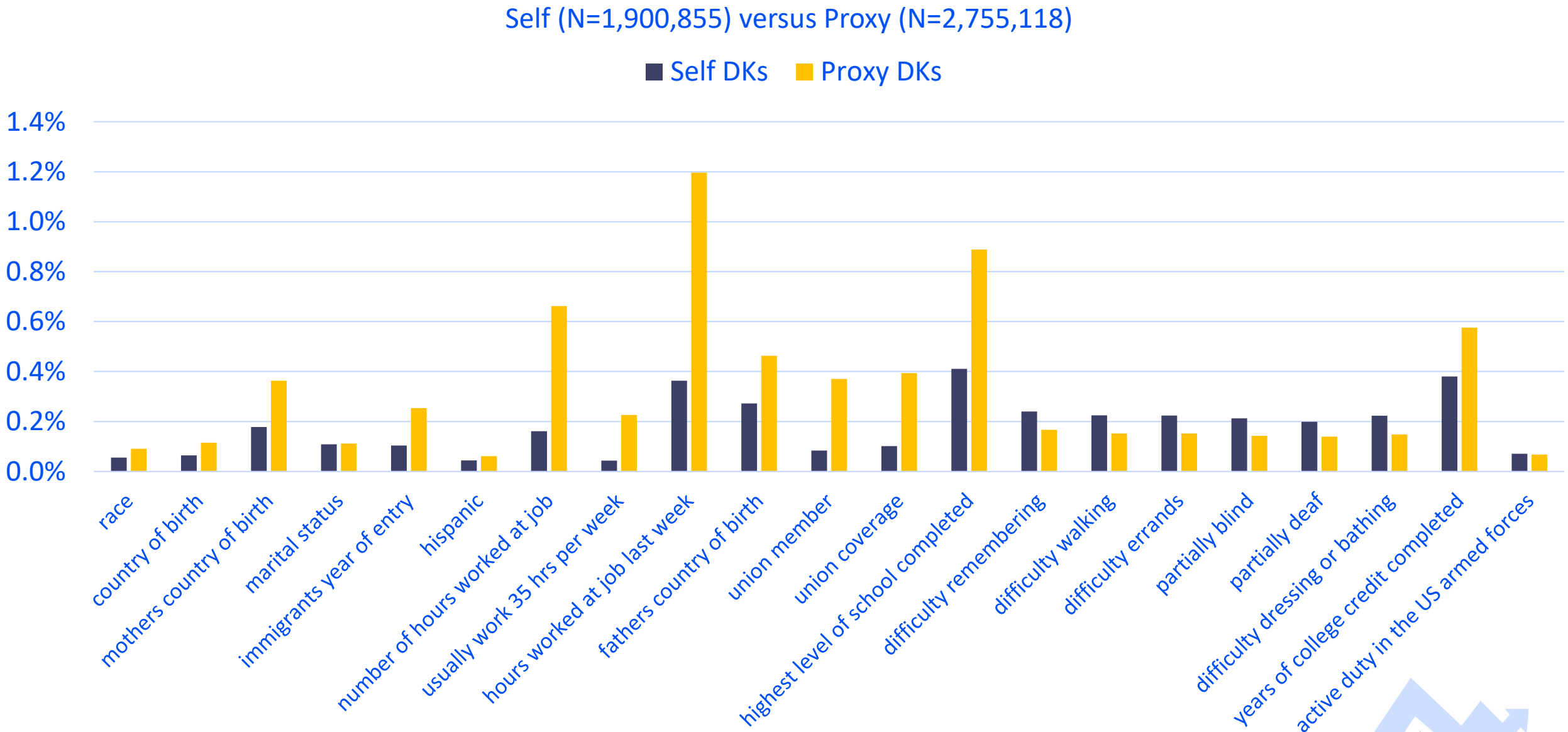
Interviewer-administered mode

“Last week, how many hours did [you/your housemate] actually work at [your/his/her] main job?”

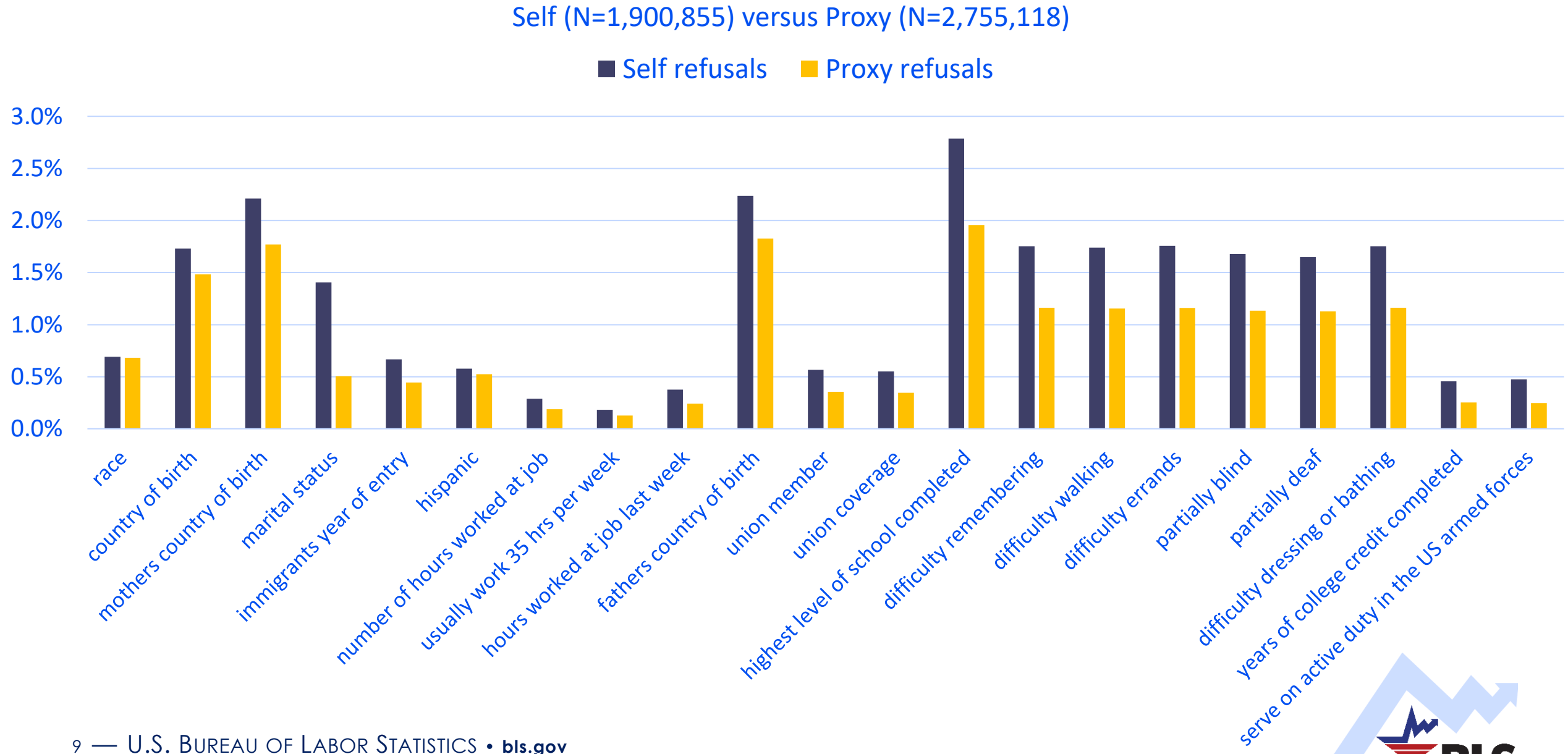
Enter number of hours (0-99)

- *DK*
- *RF*

DKs by self vs. proxy response in the Current Population Survey (2015-2017 CPS basic)



Refusals by self vs. proxy response in the CPS (2015-2017 CPS basic)



Considerations for web mode

- Offer Don't Know (DK) or Refuse (RF) response options?
  - Recorded in the CAPI/CATI mode
  - DK can be a valid response, especially for proxies
  - On the web, can't tell if skip was due to RF or speeding
- How to handle DK/RF/skips?
  - Important for instrument skip patterns, logic, classification, or data analysis

Background

- Web respondents select DK more often when offered on initial question (explicit) than when hidden (implicit) (McGee et al., 2019)
  - Hidden = separated spatially or only offered when an item is skipped
- Hidden DKs may not be advisable when DK is a valid response
  - Omitting DKs can lead to higher break-offs (Lemcke et al., 2019)
- Type of nonsubstantive response can be informative
  - DKs indicate knowledge issues; RFs indicate sensitivity (Cobb, 2018)

Web Prompts

- Some evidence that web "prompts" encouraging a substantive response can lead to lower item nonresponse (Baghal & Lynn, 2015; Derouvray & Couper, 2002; de Leeuw et al., 2015), e.g.:
  - *Please answer this important question. (American Community Survey)*
  - *If possible, please provide an answer to this question, as this is one of the key questions in this study. Please be assured that the information you give us will be treated confidentially.*
  - *These questions are important to researchers and we would be grateful if you could try your best to answer them.*
  - *We would very much like to have your answer to this question. If you would like to choose one of the proposed answers, please select "Back."*
  - *Thank you, we have recorded your answer. Perhaps you can express a preference for either yes or no? That would help us very much.*

Research Questions

- Are web prompts effective at increasing substantive response when respondents provide a nonsubstantive response (i.e., answer DK, RF, or Skip a question?)
- Does nonsubstantive response vary for self vs. proxy?

Procedure

- Survey design
  - Web survey
  - Mix of BLS survey questions
  - Self and proxy response for up to one other randomly selected HH member
- Survey questions
  - Household roster
  - Demographics (e.g., age, marital status, education)
  - Labor force questions (employment status, looking for work)
  - Household-level questions (income, disability, expenditures)

Study 1 Design (N=4960)

- 2x2 Design
  - Response option conditions:
    - Explicit DK
    - Explicit DK and RF

Last week, were you on layoff from a job?

- ○ Yes
- ○ No
- ○ Don't know
- ○ Prefer not to say

  - Prompt conditions:
    - Prompt when DK, RF, or Skip → *"If possible, please provide an answer to this question, as this is one of the key questions in this survey."*
    - No Prompt

Participant Demographics (N=4680)

- Mean age = 38.2 years (SD = 12.3)
- Median HH size = 2
- Median education level
  - Bachelor's Degree
- Gender = 57% male
- Ethnicity:
  - 20% Hispanic/Latino
- Race:
  - 70 % White
  - 7 % Asian
  - 19% Black or African American
  - 1% American Indian or Alaskan Native

Results

- Over half (n=2373; 50.7%) of participants provided substantive responses to **all** questions

- Of the 2373 participants who provided substantive responses to all questions:
  - 52.7% were in the DK only response option condition
  - 47.3% were in the DK and RF response option condition

$p < .001$

Results:
Prompt Condition (n=2359)

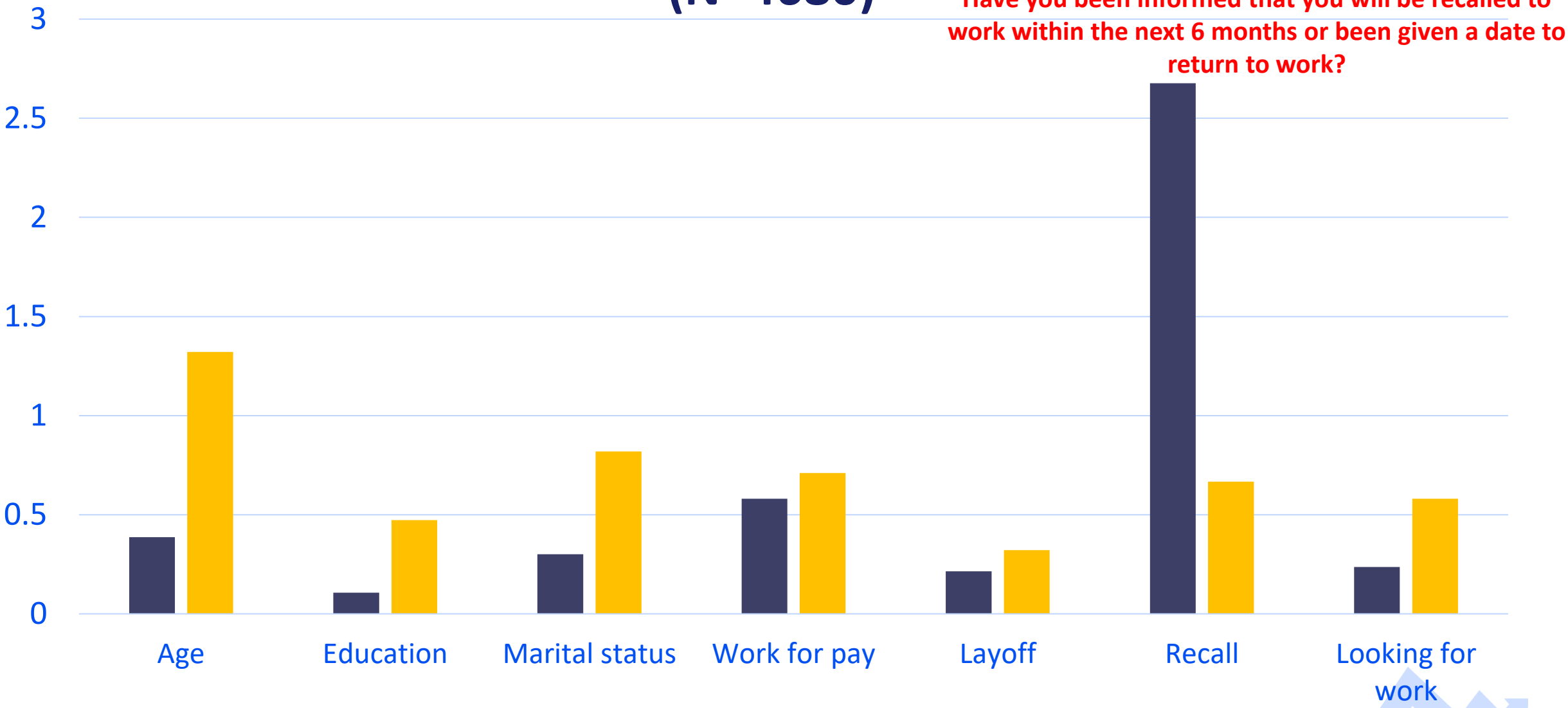
- 1,132 nonsubstantive responses were initially given
- *"If possible, please provide an answer to this question, as this is one of the key questions in this survey."*
- Post-prompt:
  - 883 (78%) were changed to a substantive response
  - 249 (22%) remained a nonsubstantive response

Results

	% Participants with DKs			% Participants with RF			% Participants with Skips		
	0	1	2 +	0	1	2 +	0	1	2 +
Prompt (n=2359)	56.6%	19.0%	24.4%	90.2%	6.7%	3.1%	95.2%	4.0%	0.8%
No Prompt (n=2321)	57.8%	15.7%	26.5%	92.4%	4.5%	3.1%	92.7%	5.9%	1.4%

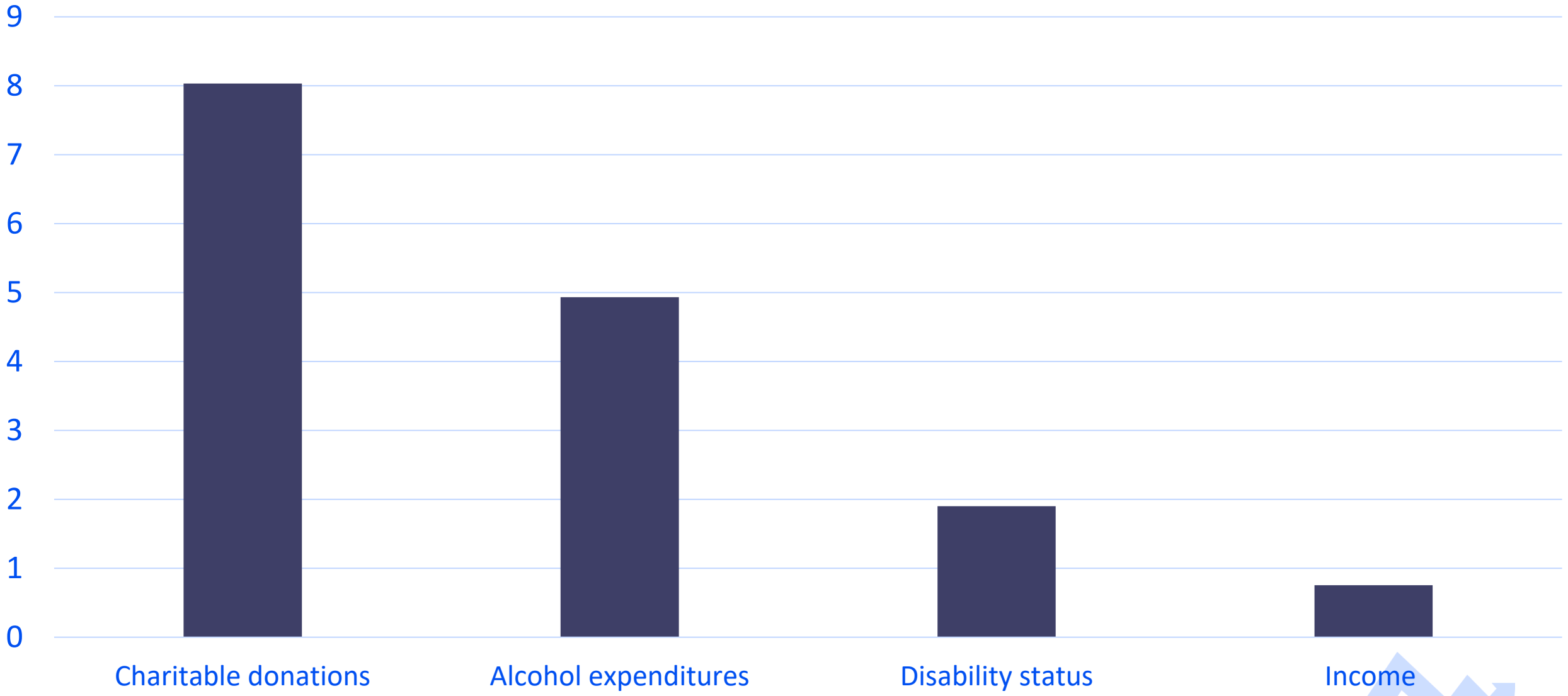
- Prompt had a marginal effect on reducing DK responses in the survey (p =.09), and reducing skips (p=.06)
- No effect on RF (p > .05)

% nonsubstantive response by select items for self vs. proxy (N=4680)



Note that not every participant got every question here due to skip patterns.

% nonsubstantive response for select household-level items (N=4680)



Study 2 Design

- Design similar to Study 1
  - Condition 1: Explicit DK offered on initial question
  - Condition 2: Hidden DK (offered only if item was skipped)
- Both conditions received the following prompt when an item was **skipped**:
  - *Please answer this important question. If you do not have the exact answer, give your best estimate or leave it blank.*
- Used 3 nonprobability panels
  - Can vary widely in demographics, representativeness, attitudes, behaviors (Kennedy et al. 2016)
  - Assess demographics and data quality

Panel demographics

	Panel A (N=1699) Condition 1 = 834 Condition 2 = 865	Panel B (N=1732) Condition 1 = 875 Condition 2 = 857	Panel C (N=1820) Condition 1 = 894 Condition 2 = 926
Age	Mean=40.1	Mean=48.7	Mean=50.2
Education	Median = Bachelor's	Median=Associate's	Median=Masters
Marital status	50.9% married	54.7% married	50.7% married
Sex	55.7% male	51.2% male	44.5% male
Ethnicity	9.1% Hispanic	10.7% Hispanic	6.6% Hispanic
Race	81.8% White 9.9% African American 1.1% American Indian/AL Native 9.4% Asian 0.4% Native Hawaiian/Other Pacific Islander	83.7% White 8.3% African American 3.0% American Indian/AL Native 6.8% Asian 1.3% Native Hawaiian/Other Pacific Islander	87.9% White 7.4% African American 2.0% American Indian/AL Native 4.8% Asian 0.8% Native Hawaiian/Other Pacific Islander
Work for pay last week	27.7% ONLY online surveys 70.0% yes 2.3% no	11.4% ONLY online surveys 43.8% yes 44.4% no	12.9% ONLY online surveys 40.5% yes 46.4% no

Results:
DK response across panels (Explicit DK condition)

	Panel A N=835	Panel B N=875	Panel C N=894
Total % of DK responses (n=1,507)	16.7%	45.3%	37.9%
% of *participants* with one or more DK response by panel	21.1%	39.2%	37.4%

Condition 1: Explicit DK offered

Panel A had significantly fewer DK responses compared to Panel B and Panel C, $p < .001$

Results:
Skipped items across panels (Conditions 1 and 2)

	Panel A N=1699	Panel B N=1732	Panel C N=1820
Total % of skips (n=1,261)	17.0%	43.4%	39.0%
% of *all participants* who skipped one or more questions by panel	10.2%	20.5%	19.2%

Panel A had a significantly fewer skips, $p < .001$

Results: Skips changed to substantive response (Conditions 1 and 2)

Please answer this important question. If you do not have the exact answer, give your best estimate or leave it blank.

	Panel A N=1699	Panel B N=1732	Panel C N=1820
% of skipped *items* changed to substantive responses after prompt	14.6%	11.2%	10.0%

Substantive response = Anything but DK or Skip
Differences were not significant by panel, p > .05

Summary

- Prompt was effective at getting substantive responses
  - Most effective for reducing DK/skips when DK/RF **explicitly offered** (Study 1)
    - Might help catch "speeders" or accidental skips
    - Prompts may not be helpful for RF
  - Small effect at getting substantive responses when **implicit DK** offered following skips (Study 2)
- Some demographic and data quality differences observed across different nonprobability panels

Conclusions/Limitations

- DKs common; nonsubstantive response higher for proxy vs. self
  - Knowledge issues
  - Hidden DKs may be advisable for some items in web surveys
- Researchers should consider the nonprobability panel used
- Limitations:
  - Nonprobability samples
  - Don’t know the “true value” – cannot assess data quality
  - Questions are factual/behavioral; might be different for opinion/attitude questions

Contact Information

Robin Kaplan and Erica Yu
Research Psychologists
Office of Survey Methods Research
U.S. Bureau of Labor Statistics
Kaplan.Robin@bls.gov