

Interactive interventions in Web surveys can increase response accuracy

Frederick G. Conrad*†
Roger Tourangeau*†
Mick P. Couper*†
Chan Zhang*

*Program in Survey Methodology, University of Michigan
†Joint Program in Survey Methodology, University of Maryland

2011 AAPOR Conference

1

Acknowledgments

- NIH grant # R01 HD041386-04A1
- Reg Baker

Two Approaches to Improving Data Quality

- The *blanket* approach
 - Provide instructions or motivation to all *R*s
 - E.g., incentive, instruction to be conscientious
- The *interactive* approach:
 - Based on particular action by *R* that indicates risk of reduced data quality, provide information and motivation to address the potential problem
 - E.g., when *R* is *inactive* for prolonged period, offer to clarify question meaning (Conrad, Schober & Coiner, 2007)
 - Take action before *R* advances!
 - Web (or other computerized administration) ideal for this approach

3

Prompt for Speeding

At AAPOR (2009), we presented two studies that slowed speeders by presenting the following prompt when their response time < 350 msec per word (typical reading time)

You seem to have responded very quickly. Please be sure you have given the question sufficient thought to provide an accurate answer. Do you want to go back and reconsider your answer?

This is *interactive* in the sense that the prompt is triggered by a particular R behavior

4

Previous Findings

• Prompt triggered by fast response times (RTs) for each of 7 questions asking for frequencies or other quantities, Embedded within larger questionnaire :

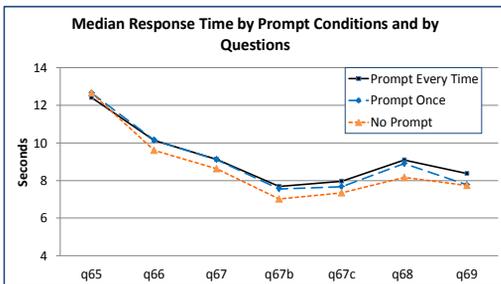
1. Overnight trips past 2 years
2. Calories per day in past year
3. \$\$ on alcohol per month
4. # drinks last 7 days
5. Binge drinks last 30 days
6. Days in bed sick during past year
7. Traffic tickets last 10 years

• Major Findings:

- Prompts slow responses (increase RT)
- Prompts reduce straightlining in later grid items
- Prompts do not affect break-offs

5

Previous Findings: Prompts Slow Down Responses



6

Current Research

- Although prompting slowed speeders in prior work, we could not directly measure response accuracy because answers were not verifiable
- Possible that *Rs* are just answering more slowly but not more thoughtfully, i.e., more accurately
- In current study, we explore relationship between slow-down due to prompting and response accuracy
- *Rs* answered 7 *numeracy* (quantitative literacy) questions for which we could determine response accuracy
 - decimal arithmetic and simple probability
 - Similar in structure to frequency/quantity *Qs* used previously to help generalize current to previous results
- Prompt triggered by speeding as in earlier studies

7

Example Numeracy Items

- *Different sports vary in the risk of injury. Which of the following numbers represents the biggest risk of getting injured while playing a sport? (1 in 200, 1 in 10, 1 in 50, 1 in 100)*
- *Chronic diseases, such as heart disease, stroke, cancer, diabetes, and arthritis, are among the most common, costly, and preventable of all health problems in the U.S. If the estimated prevalence of a particular type of chronic disease is 1 out of 1,000, what percent of people would get the disease? (1%, 0.1%, 0.01%, 0.001%)*
- Prompting when *Rs* answer < 350 ms intended to increase RT and improve accuracy

8

Commitment

- Because prompting has punitive character, we also tried motivating *Rs* to be conscientious
 - Appeal to their better nature
 - This is the blanket approach
- Inspired by Cannell, Oksenberg & Converse (1977), we presented a statement about importance of careful responding and asked *Rs* to commit to this:

9

Commitment statement

"It is very important that you read each question carefully and think about your answer before you give it. We rely on our respondents being thoughtful and taking this task seriously.

"I commit to reading each question carefully and thinking about my answer before I give it."

Yes

No, but I will participate anyway

- 99% selected "Yes"
- Commitment is *not interactive* in sense that does not depend on Rs' behavior; we think of it as a "blanket" approach to improving response quality

10

- No Commitment statement (Neutral)

"Thank you very much for participating in our study. We are grateful that you are willing to contribute to our project and hope you find it a stimulating and worthwhile experience.

11

Experimental Design

Two main factors

1. Prompting:
 - Prompt (every time Rs answer too quickly)
 - No prompt
2. Commitment:
 - Commitment statement
 - No commitment (Neutral statement)

12

Recruitment and Assignment of Rs

- 2565 Rs from 2 Opt-In Panels
 - SSI and Authentic Response
- Rs entered into lottery upon completion
- Experiment embedded in omnibus questionnaire on “health and lifestyles”
- Rs randomly assigned to each of 4 conditions

	Prompt	No Prompt
Commitment	n=640	n=598
No Commitment	n=692	n=623

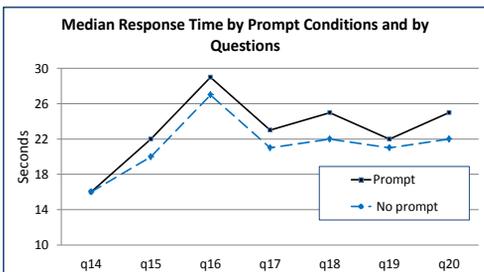
13

Research Questions

- Does prompting both slow Rs and increase their response accuracy?
- Does commitment have the same effect?
- If so, do they affect all Rs in the same way?
- Are their effects independent or does the presence of one treatment affect the other?

14

Effects of Prompting on Response Times



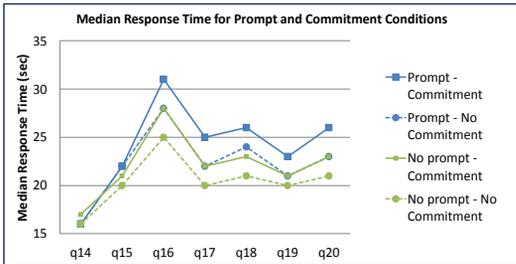
- Replicates result in earlier studies using numeracy instead of frequency/quantity Qs

15

- Prompting reduced straightlining
- and did not increase breakoffs
- Replication of prior results supports applying accuracy results to prompts with earlier Qs

16

Effects of Prompting & Commitment on Response Times



17

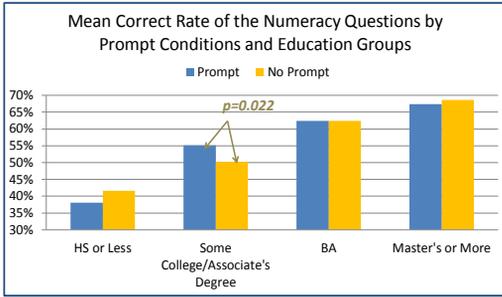
Response Accuracy: Prompting

% Correct Answers	Speeders		
	Prompt Every Time	No Prompt	P-value
q14	51.6%	54.9%	0.1569
q15	78.6%	73.0%	0.0053
q16	59.2%	58.9%	0.8906
q17	37.1%	37.2%	0.9821
q18	53.1%	49.0%	0.0814
q19	55.4%	55.3%	0.9518
q20	45.6%	50.3%	0.0463

- Prompting improves accuracy for two questions but reduces accuracy for one question

18

Response Accuracy: Prompting



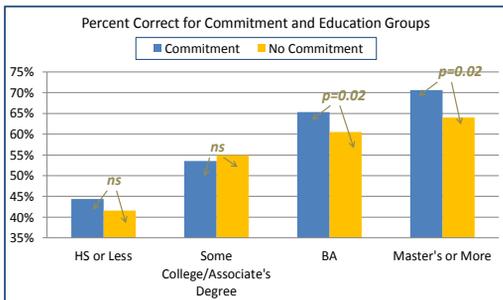
19

Accuracy and Education: Prompting

- Prompting does not improve accuracy for the most educated *Rs*
 - They may find task easy enough that don't feel need to rethink answers after being prompted
- but prompting does improve accuracy for *Rs* with Some College/Associate's degree
 - These *Rs* might find the numeracy questions hard enough that more time helps
- High School educated *Rs* may find numeracy questions so hard that, even with more time, they cannot solve them

20

Response Accuracy: Commitment



21

Accuracy and Education: Commitment

- Commitment may complement prompting by increasing accuracy for the most educated *Rs*
 - possible these *Rs* have sufficient executive control to be more accurate when they want to be
- Less educated *Rs* do not improve with commitment
 - Possible that trying harder is just not sufficient to improve accuracy

22

Conclusions

- Both interactive and blanket approaches to improving data quality can be effective, although more so for some groups than others
 - May complement each other by each helping groups that the other approach does not help
- Analogous to flu prevention with a vaccine given to everyone (blanket approach) and also treating people who develop symptoms anyway (interactive approach)

23

Next Steps

- Given success of prompting in reducing speeding and increasing accuracy, should experiment with other undesirable *R* behaviors
 - e.g., primacy effects, straightlining, conditioning
 - Lack of impact on breakoffs suggest *Rs* will tolerate intervention

24

Thank You!

25

BACK UP SLIDES START FROM HERE

26

Numeracy Items

Now for some questions on your perception of health-related risk...

- Q1. Different sports vary in the risk of injury. Which of the following numbers represents the biggest risk of getting injured while playing a sport? (1 in 200, 1 in 10, 1 in 50, 1 in 100)
- Q2. Heart disease is the leading cause of death in the United States. If Person A's chance of getting a particular type of heart disease is 1% in ten years, and person B's risk is half that of A's, what is B's risk? (2%, 5%, 0.2%, 0.5%)
- Q3. Viruses cause some of the most familiar infectious diseases, such as the common cold and the flu. If the chance of getting a particular type of viral infection is .0005, about how many out of 10,000 are expected to get infected? (0.5, 5, 50, 500)
- Q4. Chronic diseases, such as heart disease, stroke, cancer, diabetes, and arthritis, are among the most common, costly, and preventable of all health problems in the U.S. If the estimated prevalence of a particular type of chronic disease is 1 out of 1,000, what percent of people would get the disease? (1%, 0.1%, 0.01%, 0.001%)

27

Numeracy Items (cont'd)

- Q5. If Vaccine A's chance of causing serious side effects is 0.1%, and Vaccine B's chance is double that of A's, what percent of people are NOT expected to experience serious side effects after taking Vaccine B? (98%, 99%, 99.9%, 99.8%)
- Q6. Vision loss is a public health problem in the U.S. If Persons A's chance of vision loss is 1 in 100 in twenty years, and person B's risk is double that of A's, what is B's risk? (2 out of 200, 1 out of 200, 1 out of 50, none of the above)
- Q7. Iron deficiency is a condition resulting from too little iron in the body. If the risk of iron deficiency in a certain demographic group is 0.2%, how many people out of 1,000 in that group would be expected to have iron deficiency? (0.2, 2, 20, 200)

28

Regression Analysis - RT

- Dependent VAR: log transferred average RT across the numeracy items (first item excluded)
- Independent VARs:
 - Education levels
 - Prompt conditions
 - Interactions

29

Regression Analysis – RT (cont'd)

	Coef.	SE	
Intercept	2.985***	0.038	1. N=1441
HS or Less	-0.162**	0.062	2. Conditional on Rs who deserved/received at least one prompt;
Some College & Associate	-0.080	0.050	3. ***:p<.0001; **p<.01; *p<.05
BA	-0.050	0.057	4. No sig main effect of prompts
Prompt Every Time × HS or Less	0.151*	0.068	5. Coef. Of the interaction terms do not sig differ
Prompt Every Time × Some College & Associate	0.223***	0.048	
Prompt Every Time × BA	0.187**	0.059	

Scale parameter is estimated.

30

Regression Analysis - Accuracy

- Cumulative logit Model

$$\log\left(\frac{\Pr(Y \geq j)}{\Pr(Y < j)}\right) = \text{intercept}_j + \tilde{\beta}\text{cov}$$

Y: number of correct answers across numeracy items, first item not included, j=1, 2, ..., 6

Covariates: education levels, prompt conditions, and interactions

31

Regression Analysis – Accuracy (cont'd)

	Coef.	SE
intercept 1	-1.213***	0.125
intercept 2	-0.128	0.118
intercept 3	0.707***	0.119
intercept 4	1.585***	0.124
intercept 5	2.594***	0.133
intercept 6	3.878***	0.159
HS or Less	-1.732***	0.188
Some College & Associate	-1.177***	0.151
BA	-0.425*	0.168
Prompt Every Time × HS or Less	-0.213	0.195
Prompt Every Time × Some College & Associate	0.313*	0.137
Prompt Every Time × BA	0.027	0.169

Scale parameter is held fixed.

32

- N=1659
- Conditional on Rs who deserved/received at least one prompt;
- ***=p<.0001; **p<.01; *p<.05
- No sig main effect of prompts
- Differences among the coef. for the education are sig.

Additional Evidence that Impacts of Prompts Differ across Educational Groups

- For HS or less, by the time of the last numeracy questions (i.e., Q20), having been previously prompted seems to be related to lower accuracy

% Accuracy in Q20 *	Not Prompted prior to Q20	Prompted prior to Q20	p-value of CHISQ
HS or Less	42.3% (n=71)	31.7% (n=161)	0.1194
Some College & Associate	44.5%	44.6%	0.9863
BA	58.8%	52.5%	0.3191
Master or More	55.9%	59.2%	0.7242

* Prompt Every Time condition only

34

Commitment

- Few break-offs in both conditions
- Most Rs, when asked if would like to make the commitment, say "YES" (99%).
- We examine if making Rs committed affect their response behaviors.
 - Not Committed: All Rs in "Thank-You Msg" condition
 - Committed: 99% Rs in "Warn Rs to Go Slow" condition who make the commitment

35

Results: Overall Effects of Prompting: Break-offs

- We also replicate the findings that prompting has no impact on break-offs

	Prompt Every Time	No Prompt
% Break-offs during the experiment	3.76 (n=47)	4.41 (n=58)

36

Results: Overall Effects of Prompting – Straightlining

- We also find that prompting reduces straightlining in the later grid questions.

% ANY Straightlining in the Two Grids *	Prompt Every Time	No Prompt	p-value of CHISQ
All Rs			
Current Study	7.9% (n=1255)	10.4% (n=1176)	0.0331
Previous Finding	6.2% (n=855)	8.4% (n=843)	0.0783
Filtered on Speeders*			
Current Study	9.4% (n=892)	12.8% (n=870)	0.0254
Previous Finding	8.14% (n=381)	10.8% (n=360)	0.2097

* Conditional on Rs who deserved/received at least one prompt

37

Previous Findings: no effects on break-offs

- Rs might find prompts irritating and abandon questionnaire
 - if so, prompting might trade improved data quality for fewer completed questionnaires
- In fact, no differences in break-off rates between prompting condition

	Prompt Every Time	No Prompt
% Break-offs during the experiment	1.85	2.33

38

Effects of Prompting and Rs' Educational Levels - Straightlining

- “Some College & Associate” group also straightline less in subsequent grid questions when prompted in earlier numeracy questions

% ANY Straightlining in the Two Grids	Prompt Every Time	No Prompt	P-value of CHISQ
HS or Less	17.9%	18.0%	0.9736
Some College /Associate	6.9%	12.5%	0.0159
BA	6.7%	5.8%	0.7099
Master or More	8.4%	12.1%	0.3396

* Conditional on Rs who deserved/received at least one prompt.

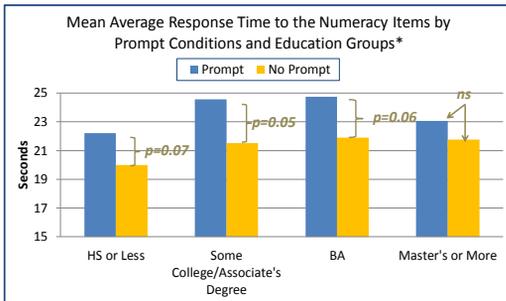
39

Effects of Prompting and Rs' Educational Levels - Response Accuracy

- Response accuracy of the numeracy questions
 - more correct response as education increaeases
 - Prompts IMPROVE accuracy for one education subgroup, "Some College & Associate"
 - For other education groups, prompts have no significant impact on the accuracy of their responses

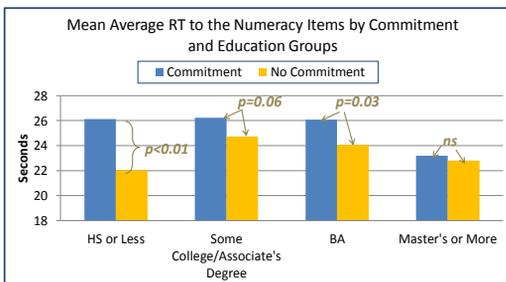
40

Response Time: Prompting



41

Response Time: Commitment



42

Effects of Prompting and Rs' Educational Levels

- Why are the effects on accuracy inconsistent, and even in the opposite direction?
- We speculate that Rs' educational levels may play a role in their responses to those numeracy questions.
- Four educational categories examined: **HS or less** (19%), **Some college/associate** (40%), **BA** (26%), **Master or more** (15%)
- It turns out the effects of prompting on **Response Times**, **Accuracy**, and **Straightlining** in the later grid questions, ALL depends on Rs educational levels.

43

Effects of Prompting and Rs' Educational Levels - Response Times

- Response times of the numeracy questions
 - Regression analysis of average RT (log-transformed)
 - Findings:
 - 1) Prompts slow down all but those with a Master's degree or more;
 - 2) The amount of slowdown does not differ significantly among the other three education groups

44

Speculation About Why Education Matters

- Most educated Rs (Masters or More) may
 - read fast so get prompted unnecessarily
 - find the numeracy questions very easy, so fast answers do not indicate reduced quality
- Least educated Rs (HS or Less) may
 - treat prompts as feedback that answers incorrect
 - lose confidence so choose answers they don't really endorse

45

Commitment & Prompt

- Some evidence that “Masters or More” slow down due to the prompts only when they are committed

Average Response Time	Committed		Not Committed	
	Prompt	No prompt	Prompt	No prompt
Master or More	26.3 (n=71)	23.3 (n=78)	24.8 (n=93)	23.6 (n=65)

46

Commitment & Prompt

- The effects of prompts on accuracy also DO NOT depend on commitment

Average # of Correct Answers*	Committed		Not Committed		
	Prompt	No prompt	Prompt	No prompt	
HS or Less	2.3	2.6	2.3	2.5	$p=0.20$
Some College & Associate	3.3	2.9	3.3	3.1	$p=0.03$
BA	3.9	4.0	3.6	3.5	$p=0.01$
Master or More	4.3	4.3	3.8	3.9	$p=0.03$

*Conditional on Rs who deserved/received at least one prompt; The first numeracy item is excluded in the count of correct answers.

47

Speed and Accuracy

- Most educated Rs (Masters) do not slow down when prompted or when committed but are more accurate when committed
- Least educated Rs (High School) slow down when prompted and committed but are not more accurate in either case
 - May lack aptitude to improve despite increased effort
- Moderately well educated Rs

48

Interactivity of the Web

- Web *Rs*' actions might indicate they are at risk of providing data with reduced quality
 - long periods of inactivity
 - advancing without answering
 - straightlining
- it is possible to provide information and motivation as soon as these actions are detected
- Makes it possible to address the problem before *Rs* start the next question
- Here, we explore prompting *Rs* when they *speed*, i.e., answer too quickly to think much before answering

49
