

University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Sociology Department, Faculty Publications

Sociology, Department of

2014

The Effect of Answering in a Preferred Versus a Non-Preferred Survey Mode on Measurement

Jolene Smyth

University of Nebraska-Lincoln, jsmyth2@unl.edu


Kristen Olson

University of Nebraska-Lincoln, kolson5@unl.edu

Alian S. Kasabian

University of Nebraska-Lincoln, aliank@unl.edu

Follow this and additional works at: <http://digitalcommons.unl.edu/sociologyfacpub>

 Part of the [Family, Life Course, and Society Commons](#), and the [Social Psychology and Interaction Commons](#)

Smyth, Jolene; Olson, Kristen; and Kasabian, Alian S., "The Effect of Answering in a Preferred Versus a Non-Preferred Survey Mode on Measurement" (2014). *Sociology Department, Faculty Publications*. 254.
<http://digitalcommons.unl.edu/sociologyfacpub/254>

This Article is brought to you for free and open access by the Sociology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Sociology Department, Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

The Effect of Answering in a Preferred Versus a Non-Preferred Survey Mode on Measurement

Jolene D. Smyth
University of Nebraska-Lincoln

Kristen Olson
University of Nebraska-Lincoln

Alian Kasabian
University of Nebraska-Lincoln

Previous research has shown that offering respondents their preferred mode can increase response rates, but the effect of doing so on how respondents process and answer survey questions (i.e., measurement) is unclear. In this paper, we evaluate whether changes in question format have different effects on data quality for those responding in their preferred mode than for those responding in a non-preferred mode for three question types (multiple answer, open-ended, and grid). Respondents were asked about their preferred mode in a 2008 survey and were recontacted in 2009. In the recontact survey, respondents were randomly assigned to one of two modes such that some responded in their preferred mode and others did not. They were also randomly assigned to one of two questionnaire forms in which the format of individual questions was varied. On the multiple answer and open-ended items, those who answered in a non-preferred mode seemed to take advantage of opportunities to satisfice when the question format allowed or encouraged it (e.g., selecting fewer items in the check-all than the forced-choice format and being more likely to skip the open-ended item when it had a larger answer box), while those who answered in a preferred mode did not. There was no difference on a grid formatted item across those who did and did not respond by their preferred mode, but results indicate that a fully labeled grid reduced item missing rates vis-à-vis a grid with only column heading labels. Results provide insight into the effect of tailoring to mode preference on commonly used questionnaire design features.

Keywords: mode preference; data quality; mail survey; web survey; open-ends; check-all; forced-choice; grid questions

1 Introduction

Advances in technology, the deterioration of the frame traditionally used for telephone surveys (Kalsbeek & Agans, 2008) and the introduction of a household-based frame with adequate coverage of the general population (the U.S. Postal Service's Delivery Sequence File; Iannacchione, 2011) have made mixed mode surveys more tenable and popular in recent years (de Leeuw, 2005; Smyth, Dillman, Christian, & O'Neill, 2010). When surveyors use multiple modes to increase response rates, they are assuming that sample members have a pre-existing preference for one mode over another (Groves & Kahn, 1979). Most research on mode preference has focused on how providing someone with their preferred mode will affect unit nonresponse (Olson, Smyth, & Wood, 2012; Shih, 2002). The effect on data quality is

still unknown. In this paper we examine the relationship between responding in one's preferred mode, question format, and data quality. In particular, we examine whether changes in question format have different effects on data quality for those responding in their preferred mode compared to those responding in a mode they did not prefer. If so, this suggests that answering a survey in a preferred mode may affect how respondents process the content and design of a questionnaire.

We first have to clarify what we mean by "mode preference". Previous research generally operationalizes a "preferred mode" as the mode selected when given a choice between two or more modes (e.g., Dillman, Smyth, & Christian, 2009; Shih, 2002). Alternatively, mode preference can be thought of as a pre-existing attitude about survey modes (Olson et al., 2012). These two approaches are related conceptually—a pre-existing attitude certainly should affect which mode one chooses when given a choice. But in practice they are distinct in that mode choice studies often offer a limited set of modes from which the respondent can choose (e.g., web and mail); we cannot tell from these studies

Contact information: Jolene D. Smyth, University of Nebraska-Lincoln, Department of Sociology, 724 Oldfather Hall, Lincoln, NE 68588-0324 (jsmyth2@unl.edu)

whether respondents would have preferred a different mode altogether (e.g., telephone). As a practical matter, measuring mode preference as an attitude prior to administering a survey with multiple modes helps to ensure that mode preference is exogenous to the measurement process in which data quality is evaluated. That is, in this study mode preference is measured separately from the survey we are using to assess data quality, thus minimizing shared method variance as much as possible.

That said, one challenge of conducting research relying on self-reports of mode preference is that those reports themselves are influenced by the mode in which the question is asked (Groves & Kahn, 1979; Millar, O'Neill, & Dillman, 2009; Smyth, Olson, & Richards, 2009). People who are surveyed in a particular mode are more likely to report a preference for that mode. However, they are not guaranteed to report such a preference. As just one example, when Groves and Kahn (1979) asked for mode preference in an in-person survey, about 22% of respondents reported preferring a mode other than in-person. When they asked the same question in a telephone survey, about 60% reported preferring a mode other than telephone. Thus, while this paper is focused on the effects of catering to self-reported mode preference on data quality, a larger question motivating our efforts in this area is whether those self-reports have predictive or explanatory value even though they are influenced by the survey mode in which the questions are asked.

The quality of survey responses is affected by a respondent's motivation, ability, and the burden of the task (Krosnick, 1991). In responding to questionnaires, sample members must expend energy and cognitive effort both interacting with the survey technology (i.e., paper and pencil or computer) and answering the actual questions (Krosnick, Narayan, & Smith, 1996; Tourangeau, Rips, & Rasinski, 2000). Ideal respondents are those who are fully engaged with the survey. Their motivation is high, their ability is equal to the task, and they are not unduly burdened by the survey process. In real life, motivation flags, ability may be low or absent, and surveys are often highly burdensome, making respondents more likely to seek relief by skipping or only minimally engaging in the response process (i.e., satisficing, Krosnick, 1991; Krosnick & Alwin, 1987), resulting in poor data quality. We suspect that respondents prefer modes that they find the least burdensome and the most motivating. In fact, previous research has shown that respondents are highly likely to report preferring modes that they find most convenient, comfortable, and enjoyable, with convenience being the most important factor (Smyth, Olson, & Richards, 2009). In lower burden and more motivating modes, the energy and cognitive effort needed to interact with the survey technology is minimized and more energy and effort can be focused on the questions being asked.

Previous research has not examined the link between re-

ceiving one's preferred mode and data quality. To examine this link, we experimentally assign respondents to one of two versions of each of three question types (multiple answer, open-ended, and grid). The versions differ in how easy they make it for respondents to satisfice (i.e., it is easier to satisfice in a check-all than a forced choice format and when given a small rather than a large open-ended answer box) and/or how difficult they make the response task (i.e., a grid with labels located apart from answer spaces is more difficult than one with labels grouped with answer spaces). In general, we expect respondents who answer in a *non-preferred mode* to be more likely to take advantage of opportunities to satisfice or to make mistakes in difficult formats, resulting in lower data quality. Because their overall burden is higher due to not receiving their preferred mode, these respondents will look to the question format to identify ways to shortcut the response process resulting in the design of the question playing a significant role in their answer.

If respondents are given a survey in their *preferred mode* (i.e., an easier mode for them), we hypothesize they are less likely to shortcut the response process through satisficing behaviors. There are two competing hypotheses about how this process might work. The first is that respondents who participate in their preferred mode will focus primarily on the content of items and thus will provide high quality responses regardless of design features, resulting in smaller question format effects. An alternative hypothesis is that these less burdened respondents will be able to be more attentive to *all* parts of a question, including the design features. Whereas in the first hypothesis, the decreased burden and increased motivation resulting from receiving one's preferred mode is manifest as a focus on content, in this competing hypothesis, it is manifest as an ability and dedication to attend to all the details. In this case, we would expect question format to have an effect. However, inasmuch as those answering in their preferred mode are not expected to seek out opportunities to satisfice, the effect may be different from the effect for those answering in a non-preferred mode. The existing literature on differences in visual design effects across demographic groups shows mixed results, with some types of design features having smaller effects among those with higher cognitive abilities as measured by age and education, some features having larger effects, and still others having no effects at all (Stern, Dillman, & Smyth, 2007).

2 Data and Methods

The data come from two surveys. The first, the 2008 Nebraska Annual Social Indicators Survey (NASIS), used an RDD sample of listed landline telephone numbers in Nebraska. In this telephone survey, 1,811 respondents aged 19 and older (AAPOR RR2 = 31%; see American Association for Public Opinion Research, 2011) indicated their preferred survey mode (see *Appendix* for question wording). A quarter

of the respondents stated a preference for mail (24.6%), almost one-fifth (19.7%) preferred an internet survey, over half (51.3%) preferred an interviewer-administered mode (1.7% in-person, 0.4% cell phone, and 49.2% home phone preference), and 4.5% abstained from giving a preference. At the end of the survey, 1,370 (75.6%) respondents stated their willingness to participate in future social science research (see *Appendix* for question wording).

The following year, 1,229 NASIS respondents who were willing to participate in future research and had complete mailing addresses on the frame were contacted for a follow-up survey, the Quality of Life in a Changing Nebraska (QLCN) survey. Sample members were randomly assigned with equal allocation to one of four mode treatments (mail only, web only, mail then web, or web then mail) so that some received their preferred mode and others did not, and to one of two survey forms that were identical in content (45 questions) but varied in format and design characteristics. All recruitment contacts were made by postal mail. The QLCN had an overall response rate of 46.0% (AAPOR RR2, $n=565$).¹ Six QLCN respondents did not report a mode preference in the NASIS and are removed from these analyses, leaving a final analytic sample size of 559. The web only group had a 25.4% response rate and differed significantly from the other three mode conditions, all of which had response rates of around 50% that did not significantly vary (Olson et al., 2012). Table 1 shows the mode in which respondents completed the QLCN by their reported mode preference from the NASIS. Most of the sample participated via mail (78%). The difference in the distribution of mode of completion across mode preferences is due to differential nonresponse related to mode preference. That is, people's willingness to respond by each mode was related to their mode preference, which is also related to demographic characteristics such as age (Olson et al., 2012).

We operationalized respondents having received their "preferred mode" in two ways. In the first, respondents who received their preferred mode are defined as those who stated a preference for a mode in the NASIS and received it as the first mode assigned in the QLCN, whether or not they actually participated in that mode (See Table A1). Using this measure, which we refer to as the "assigned preferred mode" approach, 23% of the respondents ($n = 126$) received their preferred mode. In the second, respondents who received their preferred mode are defined as those who participated in the QLCN in the mode they previously reported preferring in the NASIS, regardless of the order in which the modes were offered. Using this measure, which we refer to as the "responded in preferred mode" approach, 24% ($n = 136$) of the respondents participated in their preferred mode. Overall, 86.5% of those who were assigned their preferred mode first and 6.2% of those who were not assigned their preferred mode first participated in their preferred mode.

We conducted all analyses with both measures of having received their preferred mode—the assigned and responded measures. In every instance, the results using the assigned mode were in the same direction, but attenuated relative to those using the responded in preferred mode approach. We expect an effect when respondents participate in their preferred mode, regardless of whether it was the first mode to which they were assigned. For this reason and for parsimony, we will report the findings for analyses using the responded in preferred mode approach only; findings based on the assigned approach are available in *Appendix tables A1-A4*.

We evaluate whether the effect of each experimental factor varies across mode preference groups in multivariate models with main effects for the experimental form and mode preference and an interaction effect between form and mode preference. Because responding in one's preferred mode is not an experimentally assigned factor, respondents who participate in their preferred mode may differ on other characteristics that are also related to data quality from those who did not participate in their preferred mode. To address this, we include the respondent characteristics of sex, age, education, employment status, income, an indicator variable for missing data on the income variable, number of children, and urbanicity. These are variables that have previously been identified as predictors of mode preference or of nonresponse (e.g., Millar et al., 2009; Olson et al., 2012; Smyth, Olson, & Richards, 2009). We also include the mode of completion in each model to identify whether the differences across mode preference groups can be accounted for by differential participation in each mode. We tested an interaction term between completion mode and mode preference to see if the effects of answering in a preferred mode versus a non-preferred mode differed for those answering in mail than for those answering in web and found no significant differences in any of the models. We present t-values in the bivariate tables and standard errors in the multivariate tables. Because we are conducting multiple comparisons, we use a stricter $p < 0.01$ value for identifying statistically significant results.

3 Question Design Experiments

Because respondents may be particularly likely to engage in satisficing behaviors when faced with difficult questions (Krosnick, 1991), we choose three high burden question types to examine: multiple answer (i.e., check-all-that-apply) questions (Rasinski, Mingay, & Bradburn, 1994; Smyth, Dillman, Christian, & Stern, 2006), open-ended questions (Smyth, Dillman, Christian, & McBride, 2009), and questions formatted in a grid (Couper, 2008; Kaczmirek, 2011). We then present these questions in formats that make it easier

¹ The cumulative response rate, calculated as the product of the NASIS and QLCN response rates, is $31\% \times 46\% = 14.26\%$.

Table 1
Distribution of mode completion among QLCN respondents by mode preference in the NASIS

Mode Completion in QLCN	Preferred Mode in NASIS							
	Total		Mail		Web		Interviewer	
	N	%	N	%	N	%	N	%
Completed Mail	435	78	93	79	73	62	269	83
Completed Web	124	22	24	21	45	38	55	17
Total	559	100	117	100	118	100	324	100

or more difficult to provide an optimal answer.² We note that satisficing behaviors related to mode preference may occur on other question types and throughout entire questionnaires, but we focus on these three high burden question types as a starting point.

3.1 Experiment 1: Forced-Choice vs. Check-All-That-Apply

Respondents are often asked to report which items within a long list apply to them using either a check-all-that-apply or a forced-choice format (Rasinski et al., 1994; Smyth et al., 2006). Previous research has shown that respondents endorse fewer items and responses exhibit more primacy patterns in the check-all format as this format makes it easier for respondents to satisfice (i.e., it does not encourage deep processing of all of the items in the list Krosnick, 1991; Rasinski et al., 1994; Smyth et al., 2006; Sudman, Bradburn, & Schwarz, 1996; Thomas & Klein, 2006). In contrast, the forced-choice format explicitly requires a response for every item, encouraging deeper processing of the list and leading to more affirmative responses.

We compare responses across the check-all and forced-choice formats for those who did and did not participate in their preferred mode. We focus on a seven-item question that appears as question 29 of 45 (Figure 1) and examine the mean number of items endorsed.³

Consistent with previous literature, fewer items are selected in the check-all than in the forced-choice format for all respondents, although this is not significant at the $p < 0.01$ level (Table 2: $t = 2.01, p = 0.04$). We find the same pattern, with a strong statistically significant effect, when we subset the respondents to only those who did not respond in their preferred mode ($t = 3.39, p = 0.0008$), suggesting that these people took advantage of the opportunity to satisfice that was afforded by the check-all format.

There is no significant difference between the two formats among those who participated in their preferred mode ($t = -1.80, p = 0.07$). We note that the magnitude of the difference for those in their preferred mode is virtually the same as those in their non-preferred mode, but that the direction is opposite existing research on this topic. That is, persons who

29. Thinking about your personal finances, please indicate whether or not you have done each of the following in the past 12 months.

	Yes	No
Delayed/canceled purchasing a home	<input type="checkbox"/>	<input type="checkbox"/>
Delayed/canceled plans to buy a car	<input type="checkbox"/>	<input type="checkbox"/>
Delayed/canceled plans to make a major household purchase	<input type="checkbox"/>	<input type="checkbox"/>
Cut back on vacation spending	<input type="checkbox"/>	<input type="checkbox"/>
Cut back on eating out	<input type="checkbox"/>	<input type="checkbox"/>
Cut back on home internet access	<input type="checkbox"/>	<input type="checkbox"/>
Cut back on home landline telephone services	<input type="checkbox"/>	<input type="checkbox"/>

29. Thinking about your personal finances, which of the following have you done in the past 12 months? Check all that apply.

- Delayed/canceled purchasing a home
- Delayed/canceled plans to buy a car
- Delayed/canceled plans to make a major household purchase
- Cut back on vacation spending
- Cut back on eating out
- Cut back on home internet access
- Cut back on home landline telephone services

Figure 1. Experimental versions of multiple-answer question

participated in their preferred mode endorsed more items in the check-all format than in the forced-choice format. Given the small sample size of those who participated in their preferred mode, we cannot easily isolate why this is happening

² Unified mode design (Dillman et al., 2009) was used across the web and mail modes to eliminate unnecessary mode differences. The question wording and layout were the same across the modes and certain web features such as pop-up reminders were not used because they could not also be used in the mail mode.

³ We did not randomly vary response option order, and thus cannot test primacy effects.

statistically. Regardless, we have evidence that respondents participating in their preferred mode are not satisficing in the check-all format relative to the forced choice format. There is also no significant bivariate difference in the number of endorsed items between those who responded in their preferred mode and those who did not respond in their preferred mode within each of the formats.

In a negative binomial regression model predicting the number of “yes” answers, there is a significant interaction effect of format and responding in a preferred mode with multivariate controls (Table 3), indicating that the effect of the forced choice versus check-all that apply format differs for those who are answering in their preferred mode versus non-preferred mode. Predicted values from the final model in Table 3 indicate that the difference in the number of endorsed items between the two formats shrinks with multivariate controls for those who responded in their preferred mode (0.33 items), but is still roughly the same magnitude for those who did not respond in their preferred mode (0.53 items). Furthermore, when we separate the data into those who participated in their preferred mode and those who did not and rerun the negative binomial regression models, the effect of format is strong and significant ($b = -0.335$, $p < 0.0001$) for those who did not participate in their preferred mode, but is not significant for those who did respond in their preferred mode ($b = 0.193$, $p = 0.29$). That is, the differential effect of mode preference on the selection of “yes” answers cannot be explained by respondent characteristics (i.e., differential nonresponse or differential characteristics related to mode preference) alone, and the effect of format is concentrated among those respondents who did not participate in their preferred mode.

The finding that those answering in a non-preferred mode selected fewer items in the check-all than the forced-choice format is consistent with our hypothesis that this group would take advantage of opportunities to satisfice. The finding of no difference across the forms for those responding in their preferred mode is well explained by the hypothesis that the reduced burden and increased motivation stemming from responding in a preferred mode results in a focus on content and not form (and thus diminished form effects). Thus, it appears that satisficing behaviors are concentrated among those who did not participate in their preferred mode for forced choice versus check-all-that-apply questions.

3.2 Experiment 2: Open-ended Questions

Open-ended questions require respondents to communicate their answer in detail in writing with no preexisting categories to help, making this question type highly burdensome (Bradburn, Sudman, & Wansink, 2004). Since skipping open-ended questions or providing only short answers can ease this burden, open-ended questions are known to produce high item nonresponse rates and low quality answers,

32. All things considered, during the next year, what do you think will be your biggest challenges?

32. All things considered, during the next year, what do you think will be your biggest challenges?

Figure 2. Experimental versions of open-ended question

especially in self-administered modes (Dillman et al., 2009). Previous research has shown that providing larger answer boxes increases the amount of information reported because the size of the box communicates to respondents that more information is needed (Christian & Dillman, 2004; Israel, 2010; Smyth, Dillman, Christian, & McBride, 2009). However, larger answer boxes may also increase the perceived burden of the task, resulting in higher item nonresponse rates.

We compare responses across small and large answer box treatments (Figure 2) of an open-ended item for those who did and did not respond in their preferred mode. The question asked, “All things considered, during the next year, what do you think will be your biggest challenges?” We examine five measures of data quality: item nonresponse, number of words, proportion of respondents providing two or more themes, whether or not the respondent elaborated (provided more information about their initial statement), and the number of elaborations. These items were coded independently by two coders, and differences reconciled (see Smyth and Olson, 2011 for details).

The results indicate that both overall and when analyses are subset only to respondents who responded in their preferred mode, there is no statistically significant difference in the item nonresponse rate across the small and large box treatments (Overall $t = -2.21$, $p = 0.027$; Preferred mode $t = 1.73$, $p = 0.09$). However, among those who did not respond in their preferred mode, the large answer box has a statistically significantly higher item nonresponse rate than the small box ($t = -3.51$, $p = 0.0005$), suggesting that more of these respondents were unwilling to do the work to provide the type of answer suggested by the large answer box.

In multivariate logistic regression models, this translates to significant main effects for the experimental variation in box size ($b = 1.10$, $t = 3.35$, $p = 0.001$), a non-significant (at the $p < 0.01$ level) main effect for responding in a pre-

Table 2
Mean number of items selected in the forced-choice and check-all formats by whether respondents responded in their preferred mode

Responded	Overall		Forced Choice		Check All		Diff.	t
	N	Mean	N	Mean	N	Mean		
Overall	557	1.55	270	1.69	287	1.43	0.25	2.01
Non-preferred mode	422	1.57	205	1.81	217	1.33	0.48	3.39**
Preferred mode	135	1.52	65	1.28	70	1.74	-0.47	-1.80
Difference		0.05		0.54		-0.41		
t		0.32		2.39		-2.19		

* $p < 0.01$ ** $p < 0.001$ *** $p < 0.0001$

Table 3
Negative binomial regression coefficients and standard errors, number of "yes" answers

	Model 1		Model 2		Model 3	
	Coef.	SE	Coef.	SE	Coef.	SE
Check-all	-0.309*	0.095	-0.336***	0.094	-0.339***	0.094
Responded in preferred mode	-0.351	0.145	-0.354	0.143	-0.317	0.144
Check-all × Resp. in preferred mode	0.620*	0.196	0.574*	0.194	0.558*	0.193
Female			0.252*	0.088	0.252*	0.087
Age			-0.008	0.004	-0.009	0.004
Rural			0.111	0.101	0.109	0.101
Income < \$40K			0.286*	0.111	0.263	0.111
Income missing			-0.224	0.158	-0.227	0.157
Number of kids			0.037	0.048	0.031	0.048
Some college			-0.084	0.108	-0.088	0.108
BA or higher			-0.277*	0.104	-0.262	0.104
Married			-0.025	0.101	-0.028	0.100
Employed full-time			0.247	0.115	0.248	0.115
Employed part-time			0.112	0.127	0.132	0.127
Responded in web					-0.184	0.103
Constant	0.596**	0.064	0.794	0.311	0.870*	0.312
Overdispersion parameter	0.309**	0.068	0.208***	0.061	0.202***	0.060
LR chi2	13.77*		60.93***		64.16***	
Pseudo R2 (in %)	0.74		3.35		3.53	
N	557		546		546	

* $p < 0.01$ ** $p < 0.001$ *** $p < 0.0001$

ferred mode ($b = 1.01$, $t = 2.34$, $p = 0.02$) with a significant interaction effect (interaction $b = -2.06$, $t = -3.13$, $p = 0.002$) between box size and responding in a preferred mode that holds with multivariate controls (Table 5). When we estimate the multivariate models separately for those who responded in their preferred mode and those who did not respond in their preferred mode, once again all of the form effect is concentrated among those who did not respond in their preferred mode (not preferred mode: $b = 1.11$, $t = 3.26$, $p = 0.001$; preferred mode: $b = -1.24$, $t = -1.84$, $p = 0.07$). Thus, these results are also consistent with our hypothesis

that those answering in a non-preferred mode will take advantage of opportunities to satisfice introduced by question format and those answering in a preferred mode will be less affected by question format.

We include only persons who provided an answer to the open-ended items to test the effect of box size on the quality of open-ended responses. Consistent with previous research, among all respondents, the large box yielded more words than the small box ($t = -4.78$, $p < 0.0001$). In multivariate negative binomial models, this increase did not significantly vary by whether the respondent responded in their

Table 4
Data quality outcomes for open-ended question by whether respondents responded in their preferred mode

	Overall		Small Box		Large Box		Diff.	t
	N	Mean %	N	Mean %	N	Mean %		
<i>Item Nonresponse Rate</i>								
Overall	559	12.3	272	9.19	287	15.33	-6.14	-2.21
Non-preferred mode	423	12.5	206	6.80	217	17.97	-11.17	-3.51**
Preferred mode	136	11.7	66	16.67	70	7.14	9.52	1.73
Difference		0.7		-9.87		10.83		
t		0.24		-2.43		2.20		
<i>Mean Number of Words</i>								
Overall	490	6.47	247	5.09	243	7.87	-2.78	-4.78***
Non-preferred mode	370	6.52	192	5.20	178	7.93	-2.73	-4.08***
Preferred mode	120	6.33	55	4.71	65	7.71	-3.00	-2.53
Difference		0.18		0.49		0.22		
t		0.26		0.64		0.20		
<i>Percent with 2 or more Themes</i>								
Overall	490	14.90	247	12.55	243	17.28	-4.73	-1.47
Non-preferred mode	370	14.59	192	14.06	178	15.17	-1.11	-0.30
Preferred mode	120	15.83	55	7.27	65	23.08	-15.80	-2.40
Difference		-1.24		6.79		-7.91		
t		-0.33		1.34		-1.44		
<i>Percent Elaborating</i>								
Overall	490	30.61	247	24.70	243	36.63	-11.93	-2.88*
Non-preferred mode	370	32.16	192	25.00	178	39.89	-14.89	-3.09*
Preferred mode	120	25.83	55	23.64	65	27.69	-4.06	-0.50
Difference		6.33		1.36		12.20		
t		1.31		0.21		1.75		
<i>Mean Number of Elaborations</i>								
Overall	150	1.46	61	1.34	89	1.54	-0.20	-1.47
Non-preferred mode	119	1.45	48	1.40	71	1.49	-0.09	-0.62
Preferred mode	31	1.48	13	1.15	18	1.72	-0.57	-2.50
Difference		-0.03		0.24		-0.23		
t		-0.19		0.95		-1.11		

Two-tailed t-tests.

* $p < 0.01$ ** $p < 0.001$ *** $p < 0.0001$

preferred mode (no significant interaction between form and mode preference, $p = 0.58$), nor did this relationship change when accounting for respondent characteristics (results not shown).

For both box sizes, 14% to 15% of respondents who answered in a non-preferred mode provided multiple themes ($t = -0.30$, $p = 0.76$). Among those responding in their

preferred mode, 7.27% of those who received the small box provided multiple themes compared to 23.08% of those who received the large box. This difference is not statistically significant at the $p < 0.01$ level ($t = -2.40$, $p = 0.02$), but the results suggest that the difference may be statistically significant with a larger sample size.⁴

⁴ In multivariate logistic regression models separately estimated

Table 5
Logistic regression coefficients and standard errors predicting item nonresponse to open-ended question (Q32)

	Coef.	SE	Coef.	SE	Coef.	SE
Large box	1.100**	0.328	1.148**	0.338	1.149**	0.339
Responded in preferred mode	1.009	0.431	0.984	0.450	0.980	0.454
Large box × responded in preferred mode	-2.056*	0.658	-2.089*	0.681	-2.088*	0.681
Female			-0.377	0.291	-0.378	0.292
Age			-0.023	0.012	-0.023	0.012
Rural			-0.421	0.378	-0.420	0.378
Income < \$40K			-0.444	0.397	-0.442	0.399
Income missing			0.937	0.400	0.938	0.400
Number of kids			-0.037	0.154	-0.037	0.154
Some college			-0.568	0.371	-0.567	0.371
BA or higher			-0.409	0.339	-0.411	0.339
Married			-0.445	0.323	-0.445	0.323
Employed full time			-0.172	0.367	-0.173	0.368
Employed part time			-0.562	0.440	-0.564	0.442
Responded in web					0.023	0.329
Constant	-2.618***	0.277	-0.266	0.975	-0.274	0.981
LR chi2	15.60*		32.84*		32.86*	
Pseudo R2 (in %)	3.73		7.92		7.92	
N	559		548		548	

* $p < 0.01$ ** $p < 0.001$ *** $p < 0.0001$

Although there are bivariate effects of not participating in one's preferred mode and answer box size on the percent who elaborated on a theme, there are not differential effects of answer box size for those who participated versus did not participate in their preferred mode on the elaboration rate in multivariate logistic regression models controlling for respondent characteristics (multivariate results not shown). We see no significant effect of answer box size on the number of elaborations overall or in multivariate models for any group.

Thus, answer box size had the same effect on the length of responses, providing multiple themes, and the provision of elaborations across those answering in preferred and non-preferred modes, but had differential effects on item non-response for those who responded in their preferred vs. non-preferred mode. As was the case with the forced choice/check-all-that-apply experiments, where we see statistically significant results, they occur among those respondents who did not participate in their preferred mode, which is highly consistent with our hypothesis about this group. However, the findings for those answering in their preferred mode are more mixed. That they were not more likely to skip the question in the large answer box treatment is consistent with the hypothesis that they would be focused on content and less affected by form; however, the finding of shorter answers in the small answer box treatment suggests that those answering in their preferred mode were affected

by the question format, although not more than or in different ways than their counterparts who were answering in a non-preferred mode.

3.3 Experiment 3: Grid Questions

The grid format (i.e., matrices; Dillman et al. 2009) is commonly used to save space by grouping questions that have the same response options. In this format, respondents do not have to process a new question stem or the format of the response options after the first prompt, but they do have to register their answer in the space that aligns with both the row of the item and the column of the desired response option. This requires them to track both horizontally and vertically, which is burdensome. As respondents move to the right and/or downward in the grid the two pieces of information they are trying to connect become further apart, making the task even more difficult (Couper, 2008; Kaczmirek, 2011). As a result, the grid format often results in item missing data and straightlining (Couper, Tourangeau, Conrad, & Zhang, 2013; Tourangeau, Couper, & Conrad, 2004).

One possible way to reduce navigational and response burden is by repeating the response option labels in every
 for the two mode preference groups, the effect of form has a significance level of $p=0.02$ for those who responded in their preferred mode, and a significance level of $p=0.73$ for those who did not respond in their preferred mode.

14. Please indicate how satisfied or dissatisfied you are with the availability of each of the following in your community.

	Very Dissatisfied		Neither Satisfied nor Dissatisfied		Very Satisfied	
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Parks	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bicycle paths	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Outdoor areas to hunt, fish, or hike	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sporting events	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Restaurants	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fine arts (museums and theatres)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cell phone service	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Internet access	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Libraries	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

14. Please indicate how satisfied or dissatisfied you are with the availability of each of the following in your community.

	Very Dissatisfied		Neither Satisfied nor Dissatisfied		Very Satisfied	
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Parks	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bicycle paths	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Outdoor areas to hunt, fish, or hike	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sporting events	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Restaurants	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fine arts (museums and theatres)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cell phone service	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Internet access	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Libraries	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 3. Experimental versions of grid item

row, thereby reducing vertical processing and allowing respondents to focus on horizontal processing (Dillman, 1978). However, doing so may give the grid a dense, cluttered visual appearance, increasing its perceived burden.

In this experiment, respondents received either a top labeled grid with labels only at the column headings or a fully labeled grid with labels next to the answer spaces in every row (Figure 3). Due to a software limitation, the web respondents only saw the top labeled grid so we focus only on mail respondents in the analyses that follow. The data quality measures we use are item nonresponse and straightlining rates within the grid items. Given that all people were exposed to the same number of questions, our measure of item nonresponse is the number of items left blank out of the nine items in the grid (only three respondents skipped the entire grid). Straightlining, or nondifferentiation, occurs when respondents select the same response for all of the items in the grid, and is an indication of lack of cognitive effort (Dillman et al., 2009; Krosnick, 1991). Our measure of straightlining is the proportion of respondents who provided the same answer to all of the items in the grid. We also examined the standard deviation of answers to the grid questions and reached the same conclusions.

Because the fully labeled grid has not, to our knowledge, been empirically evaluated before, we must draw on theory for our predictions of the overall effects. Although the added visual density may increase the *perception of burden*, we expect the full labeling to reduce the *actual burden*, resulting in lower item nonresponse and straightlining rates compared to the top labeled grid.

Consistent with our prediction, the fully labeled grid had significantly fewer items missing than the top labeled grid ($t = 2.62, p = 0.009$; Table 6). While this positive effect of the fully labeled grid held in multivariate analyses (negative binomial model results not shown), it did not vary across those who responded in their preferred mode versus a non-

preferred mode. In addition, straightlining did not significantly differ by the type of grid the mail respondents saw or whether or not they responded in their preferred mode. Thus, we do not have evidence that respondents who answered in a non-preferred mode were more likely to satisfice, nor do we have evidence that respondents who answered in their preferred mode were any more or less likely to pay attention to the grid labels than those who did not answer in their preferred mode. We suspect this lack of effect may be due to the salient nature of the items (i.e., community features/resources) contained in the grid. It is likely that most people can easily comprehend these items and report an opinion about them. Moreover, many respondents were likely actively interested in the availability of these items in their communities. That is, because of the content of the items, motivation may have been high and difficulty low in this grid compared to many other survey grids. Indeed, only three respondents skipped the grid entirely. A less salient and/or more complex set of items might produce different results.

4 Discussion

In this paper we examined whether answering a survey in a preferred mode affects how respondents process the content and design of a questionnaire by looking at whether changes in question format have different effects on data quality for those answering in non-preferred and preferred modes. We hypothesized that those answering in a non-preferred mode would take advantage of the opportunities to satisfice afforded by certain question formats as an effort to reduce burden and would struggle with more burdensome formats, but that those answering in a preferred mode would not take advantage of these opportunities or struggle in the same way. Instead, we hypothesized that the decrease in burden and increase in motivation associated with answering in a preferred mode would allow them to focus on optimizing in their responses. We were unsure whether these gains would result in increased focus on question content alone or increased focus on both question content and information conveyed by design features.

Our paper has two main findings. The first is that those who answered the survey in a non-preferred mode did seem to take advantage of the opportunities to satisfice that arose from question format. They checked fewer answers when a multiple-answer question was presented in a check-all rather than a forced-choice format, were less likely to provide an answer when they received a large open-ended box, and entered shorter and less detailed responses into a small open-ended answer box. In addition, they had higher item nonresponse rates on a grid with labels appearing only at the top of the answer columns than on a grid with labels at each answer space.

Our second main finding is that, in general, those who answered the survey in a preferred mode did not seem to

Table 6
Item nonresponse and straightlining rates for top labeled versus fully labeled grids by whether respondents responded in their preferred mode, mail respondents only

	Overall		Top Labeled		Fully Labeled		Diff.	t
	N	Mean %	N	Mean %	N	Mean %		
<i>Mean Number of Missing Items</i>								
Overall	433	0.31	205	0.46	228	0.18	0.28	2.62*
Non-preferred mode	341	0.30	162	0.41	179	0.19	0.22	2.02
Preferred mode	92	0.37	43	0.63	49	0.14	0.49	1.69
Difference		-0.07		-0.21		0.05		
t		-0.56		-0.90		0.39		
<i>Percent Straightlining</i>								
Overall	430	8.37	203	8.37	227	7.05	1.33	0.51
Non-preferred mode	339	7.96	161	8.70	178	7.30	1.39	0.47
Preferred mode	91	6.59	42	7.14	49	6.12	1.02	0.19
Difference		1.37		1.55		1.18		
t		0.44		0.32		0.28		

Mail cases only; Two tailed t-tests:

* $p < 0.01$ ** $p < 0.001$ *** $p < 0.0001$

take advantage of the opportunities to satisfice afforded by question format. There were few differences in the various data quality outcomes across the different formats for those who responded in their preferred mode. Overall, the lack of differences is consistent with the hypothesis that those who received their preferred mode were able to focus on the content of the question and were less affected by formatting differences than those who received a non-preferred mode. Where there were differences across formats for this group (e.g., more themes in the larger answer box and less item nonresponse in the fully labeled grid), the differences were similar to those that occurred among respondents answering in a non-preferred mode but also in a way that suggested more motivation (less item nonresponse, more themes).

Like all studies, this one has limitations. One possible limitation is that all of the QLCN respondents were previous survey participants who were willing to participate in future studies. Thus, this is a more cooperative sample than the general population. If this cooperative nature translates into higher motivation while responding to the survey, we would expect an attenuation of our effects as even our most unmotivated respondents (those responding in a non-preferred mode) are fairly highly motivated compared to the general population. Thus, our comparisons of the effect of form across those who did and did not complete the survey in their preferred mode may be somewhat conservative. That said, the form effects that we found overall are consistent with those reported in previous research. For example, we found that overall the forced-choice format yielded an additional 0.25 items selected while previous studies have found that it

yielded increases ranging from 0.17 to 0.52 items selected in similar behavior-type questions⁵ (Smyth et al., 2008, 2006; Thomas & Klein, 2006). We also found that the large answer box yielded an additional 2.78 words over the small box compared to increases ranging from 1.8 to 7.9 reported in previous studies where the answer box size is doubled (Christian & Dillman, 2004; Israel, 2010). One exception is that we found that the large open-ended box increased item nonresponse by about six percentage points while previous studies have not found increases in item nonresponse as a result of answer box size; however, the previous studies were conducted via the web and were limited to specialized populations such as college students and Florida Extension users (Israel, 2010; Smyth, Dillman, Christian, & McBride, 2009).

While the use of multiple surveys likely resulted in a more cooperative sample, it is also a unique strength of this study in that it allowed us to measure mode preference separately from and prior to the collection of the data assessed for quality. Thus, we avoid the endogeneity problem that arises when these outcomes are measured simultaneously. Moreover, inasmuch as measuring mode preference and then catering to it in this way is only possible with multiple survey contacts, this limitation may be endemic to all work of this nature.

A second possible limitation is that our hypotheses assume that responding in one's preferred mode will lower the

⁵ We limit the comparisons here to behavior and event questions similar to the item in the QLCN because Smyth, Dillman, and Christian (2008, 2006) found that attitude and opinion questions yielded even larger differences.

likelihood of a respondent satisficing. We think our rationale here is strong, but note that it is also possible that some portion of sample members may prefer modes that help them satisfice. In this case, those who respond in a preferred mode will behave similarly to those who respond in a non-preferred mode, which would attenuate our results. That we found that those responding in a preferred mode did not seem to be taking advantage of opportunities to satisfice suggests strong support for our assertion that responding in a preferred mode reduces burden and/or increases motivation.

A third limitation of this study is that we were unable to include an interviewer administered mode in the QLCN. This raises two problems. First, those who previously reported preferring an interviewer-administered mode were always assigned a non-preferred mode of mail or web (i.e., they were always in the non-preferred group). To ensure that this group was not driving our findings, we re-ran all analyses excluding them (results not shown). Doing so drastically reduced the sample size and statistical power; as a result, many of the differences across experimental formats lost significance. However, for most of our outcomes, excluding them either had no effect on the size of the difference across question formats or increased the size of the difference. These results suggest that, if anything, the inclusion of these respondents in the results shown here was suppressing the form effects, not magnifying them. That is, the results reported here were not driven by this group. The second problem raised by not having an interviewer-administered mode is that we cannot know what would happen if those who reported preferring the web or mail had been assigned to an interviewer-administered mode. This is an area in need of future research, and it will require different question experiments that are not as guided by visual design.

A fourth limitation of this study is that the sample sizes got quite small in some of the subgroup analyses. For example, only 135 people responded in their preferred mode, 65 in one experimental form and 70 in the other. In some cases, the lack of statistical significance at a $p < .01$ level in this study may be due to this small sample size, a suspicion confirmed by posthoc power analyses. We have reported t-tests in the tables for readers interested in what conclusions would be without adjusting the p-value for multiple comparisons. We believe that the primary comparison which would be statistically significant with a larger sample size is the number of words and themes in the open-ended item, in which those who responded in their preferred mode provided more words and themes with a large box than with a small box (these analyses would be significant with a $p < 0.05$ significance level).⁶ The small sample size also limited our ability to report results separately for those who preferred the mail and web modes except for on the grid item (where our analyses were limited to the mail only due to limitations in the web survey software). While additional analyses indicate the re-

sults are similar across these preferences, our sample sizes are small enough that we cannot be confident in these null results.

When we first began this research, we were perplexed by a contradiction between previous research findings and current research practices. On the one hand, research suggested that measures of mode preference were strongly biased toward the mode in which respondents were being surveyed (Groves & Kahn, 1979; Millar et al., 2009; Smyth, Olson, & Richards, 2009); that is, that self-report mode preference may be an artifact of the measurement context. On the other hand, survey researchers across various industries were clamoring to incorporate the web mode alongside mail and telephone surveys because they believed that certain subgroups, especially the young, would prefer it. So while empirical evidence suggested mode preferences are not real, important survey design decisions were increasingly being made under the assumption that people had mode preferences and that catering to them could reduce nonresponse and improve measurement. We set out to identify whether reported mode preferences are related to nonresponse and/or measurement or if they are simply measurement artifacts as suggested by the work of Groves and Kahn (1979) and others. In other words, we wanted to know if measures of mode preference could be useful to surveyors. The answer is mixed. Our findings echo the findings of others in that people were more likely to report preferring the mode they were being surveyed in when asked the mode preference question. However, about half of respondents reported preferring a different mode. Moreover, that receiving the mode one previously reported preferring is related to their propensity to respond (Olson et al., 2012) and, as shown in this paper, to how people process questions, diminishing the effects of format, suggests that reports of mode preference are more than simply an artifact of the measurement situation in which they are gathered. These reports do have some explanatory or predictive utility, even though they are affected by the survey mode in which they are asked.

The practical implications of this research are mixed. When it comes to nonresponse, the findings from Olson et al., 2012 are clear that catering to mode preference can increase response rates, but that those given the web mode will need to be offered another mode as well to further encourage participation (Smyth et al., 2010). Thus, mode preference can be used as part of a tailored design strategy to try to increase response either by asking for mode preference in an early survey in a longitudinal design or using demographic or geographic information from a rich frame to assign modes based on characteristics that are associated with mode preferences (Millar et al., 2009; Smyth, Olson, & Richards, 2009)

⁶ The conclusions for the number of elaborations are sensitive to the distributional form assumed for the significance test, so we are more cautious about these results.

for a cross sectional design.

When it comes to measurement, more research is needed before firm recommendations can be made. This research should examine the effects of answering in a preferred versus a non-preferred mode on different types of questions (i.e., attitude, behavior, etc.); across different question formats (i.e., rating scales, list-style open-ends, etc.); across low and high burden questions; across socially desirable and undesirable questions, and in telephone and face-to-face modes. This paper is a first examination of whether there is a relationship between answering in a preferred mode and the quality of survey responses. That it shows such a relationship does exist should be a catalyst for further research on this topic.

Acknowledgments

The authors thank the Bureau of Sociological Research at the University of Nebraska-Lincoln for access to the NA-SIS data and Dan Hoyt for support in collecting the initial mode preference questions. Additional data collection was supported by a University of Nebraska-Lincoln Office of Research Layman Award to K.O. (PI). Analysis of these data was supported by Cooperative Agreement #58-3AEU-0-0020 between the University of Nebraska-Lincoln (Pis Smyth and Olson) and the U.S. Department of Agriculture National Agricultural Statistics Service, supported by the National Center for Science and Engineering Statistics of the National Science Foundation. The authors also thank the anonymous reviewers for their careful and thoughtful comments.

References

- American Association for Public Opinion Research. (2011). Standard definitions: final dispositions of case codes and outcome rates for surveys. Retrieved from <http://aapor.org/Content/NavigationMenu/AboutAAPOR/StandardsampEthics/StandardDefinitions/StandardDefinitions2011.pdf>
- Bradburn, N., Sudman, S., & Wansink, B. (2004). *Asking questions: the definitive guide to questionnaire design—for market research, political polls, and social and health questionnaires*. San Francisco, CA: Jossey-Bass.
- Christian, L. M. & Dillman, D. A. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Research*, 68, 57–80.
- Couper, M. P. (2008). *Designing effective web surveys*. New York: Cambridge University Press.
- Couper, M. P., Tourangeau, R., Conrad, F. G., & Zhang, C. (2013). The design of grids in web surveys. *Social Science Computer Review*, 31, 322–345.
- de Leeuw, E. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21, 233–255.
- Dillman, D. A. (1978). *Mail and telephone surveys: the total design method*. New York: John Wiley & Sons.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail and mixed-mode surveys: the tailored design method* (3rd ed.). New York: John Wiley & Sons.
- Groves, R. M. & Kahn, R. L. (1979). *Surveys by telephone*. New York: John Wiley & Sons.
- Iannacchione, V. G. (2011). The changing role of address-based sampling in survey research. *Public Opinion Quarterly*, 75, 556–575.
- Israel, G. D. (2010). Effects of answer space size on responses to open-ended questions in mail surveys. *Journal of Official Statistics*, 26, 271–285.
- Kaczmirek, L. (2011). Attention and usability in Internet surveys: effects of visual feedback in grid questions. In M. Das, P. Ester, & L. Kaczmirek (Eds.), *Social and behavioral research and the Internet: advances in applied methods and research strategies* (pp. 191–214). New York: Routledge.
- Kalsbeek, W. D. & Agans, R. P. (2008). Sampling and weighting in household telephone surveys. In J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japac, P. J. Lavrakas, . . . R. L. Sangster (Eds.), *Advances in telephone survey methodology* (pp. 29–55). Hoboken, NJ: John Wiley & Sons.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5, 213–236.
- Krosnick, J. A. & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, 51, 201–219.
- Krosnick, J. A., Narayan, S. S., & Smith, W. R. (1996). Satisficing in surveys: initial evidence. In M. T. Braverman & J. K. Slater (Eds.), *Advances in survey research* (pp. 29–44). San Francisco: Jossey-Bass.
- Millar, M. M., O'Neill, A. C., & Dillman, D. A. (2009). *Are mode preferences real?* Technical Report 09-003 of the Social and Economic Sciences Research Center, Washington State University, Pullman, WA.
- Olson, K., Smyth, J. D., & Wood, H. (2012). Does giving people their preferred survey mode actually increase survey participation? an experimental examination. *Public Opinion Quarterly*, 76, 611–635.
- Rasinski, K. A., Mingay, D., & Bradburn, N. M. (1994). Do respondents really 'mark all that apply' on self-administered questions? *Public Opinion Quarterly*, 58, 400–408.
- Shih, X., T. Fan. (2002). Response rates and mode preferences in web-mail mixed-mode surveys: a meta-

analysis. *International Journal of Internet Science*, 2, 59–82.

Smyth, J. D., Dillman, D. A., & Christian, L. M. (2008). Does 'yes or no' on the telephone mean the same as 'check-all-that-apply' on the web? *Public Opinion Quarterly*, 72, 103–113.

Smyth, J. D., Dillman, D. A., Christian, L. M., & McBride, M. (2009). Open-ended questions in web surveys: can increasing the size of answer boxes and providing extra verbal instructions improve response quality? *Public Opinion Quarterly*, 73, 325–337.

Smyth, J. D., Dillman, D. A., Christian, L. M., & O'Neill, A. C. (2010). Using the Internet to survey small towns and communities: limitations and possibilities in the early 21 century. *American Behavioral Scientist*, 53, 1423–1448.

Smyth, J. D., Dillman, D. A., Christian, L. M., & Stern, M. J. (2006). Comparing check-all and forced-choice question formats in web surveys. *Public Opinion Quarterly*, 70, 66–77.

Smyth, J. D. & Olson, K. (2011). *Comparing numeric and text open-end responses in mail and web surveys*. Paper presented at the Annual Meeting of the American Association of Public Opinion Research, May 11–15, Phoenix, AZ.

Smyth, J. D., Olson, K., & Richards, A. (2009). *Unraveling mode preference*. Presented at the annual meeting of the American Association for Public Opinion Research, Hollywood, FL. May ,14–17, 2009.

Stern, M. J., Dillman, D. A., & Smyth, J. D. (2007). Visual design, order effects and respondent characteristics in a self-administered survey. *Survey Research Methods*, 1, 121–138.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: the application of cognitive processes to survey methodology*. San Francisco, CA: Jossey-Bass.

Thomas, R. K. & Klein, J. D. (2006). Merely incidental? effects of response format on self-reported behavior. *Journal of Official Statistics*, 22, 221–244.

Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position, and order: interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly*, 68, 368–393.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York: Cambridge University Press.

Appendix
Question Wording

Mode Preference If you received a request to do another survey like this one, would you prefer to participate in an in-person interview at your home, an interview on your home phone, an interview on your cell phone, a paper survey sent by mail, or a survey on the Internet?

Willingness We are in the process of putting together a list of people who are willing to be contacted in the future to be a part of additional social research projects. Would you be willing to be contacted again by researchers at the University of Nebraska-Lincoln to participate in a similar research study?

Table A1
Distribution of mode assignment among QLCN respondents by mode preference in the NASIS

Mode Assignment in QLCN	Preferred Mode in NASIS							
	Total		Mail		Web		Interviewer	
	N	%	N	%	N	%	N	%
Assigned Mail First	315	56	68	58	58	49	189	58
Assigned Web First	244	44	49	42	60	51	133	42

Table A2
Mean number of items selected in the forced-choice and check-all formats by whether respondents were assigned to their preferred mode

	Forced Choice	Check All	Difference	t
Overall	1.69	1.43	0.25	2.01
Assigned non-preferred mode	1.74	1.37	0.37	2.60*
Assigned preferred mode	1.47	1.63	-0.16	-0.60

Two-tailed t-tests

* $p < 0.01$ ** $p < 0.001$ *** $p < 0.0001$

Table A3
Data quality outcomes for open-ended question by whether respondents were assigned to their preferred mode

	Small Box	Large Box	Difference	t
<i>Item Nonresponse Rate</i>				
Overall	9.19	15.33	-6.14	-2.21
Assigned Non-preferred mode	7.01	16.44	-9.42	-3.07*
Assigned preferred mode	17.24	11.76	5.48	0.87
<i>Mean Number of Words</i>				
Overall	5.09	7.87	-2.78	-4.78**
Assigned Non-preferred mode	5.15	7.81	-2.66	-4.08**
Assigned preferred mode	4.85	8.07	-3.21	-2.46
<i>Percent with 2 or more Themes</i>				
Overall	12.55	17.28	-4.73	-1.47
Assigned Non-preferred mode	14.07	16.39	-2.32	-0.63
Assigned preferred mode	6.25	20.00	-13.75	-2.08
<i>Percent Elaborating</i>				
Overall	24.70	36.63	-11.92	-2.88*
Assigned Non-preferred mode	23.12	39.89	-16.77	-3.59**
Assigned preferred mode	31.25	26.67	4.58	0.52
<i>Mean Number of Elaborations</i>				
Overall	1.34	1.54	-0.20	-1.47
Assigned Non-preferred mode	1.39	1.48	-0.09	-0.57
Assigned preferred mode	1.20	1.81	-0.61	-2.56

Two-tailed t-tests

* $p < 0.01$ ** $p < 0.001$ *** $p < 0.0001$

Table A4
Item nonresponse and straightlining rates for top labeled versus fully labeled grids by whether respondents were assigned to their preferred mode

	Top Labeled	Fully Labeled	Difference	t
<i>Mean Number of Missing Items</i>				
Overall	0.46	0.18	0.28	2.62*
Assigned Non-preferred mode	0.42	0.22	0.19	1.68
Assigned preferred mode	0.68	0.02	0.66	2.40
<i>Percent Straightlining</i>				
Overall	8.37	7.05	1.33	0.51
Assigned Non-preferred mode	8.23	7.26	0.97	0.34
Assigned preferred mode	9.09	6.25	2.84	0.47

Two-tailed t-tests.

* $p < 0.01$ ** $p < 0.001$ *** $p < 0.0001$