Summer 2009

Open-Ended Questions in Web Surveys: Can Increasing the Size of Answer Boxes and Providing Extra Verbal Instructions Improve Response Quality?

Jolene D. Smyth
University of Nebraska-Lincoln, jsmyth2@unl.edu

Don A. Dillman
Washington State University, dillman@wsu.edu

Leah Melani Christian
Pew Research Center for the People and the Press

Mallory McBride
University of Montana

Published in *Public Opinion Quarterly* 73:2 (Summer 2009), pp. 325–337: doi: 10.1093/poq/nfp029

Published online May 28, 2009.

Open-Ended Questions in Web Surveys: Can Increasing the Size of Answer Boxes and Providing Extra Verbal Instructions Improve Response Quality?

Jolene D. Smyth,[1] Don A. Dillman,[2] Leah Melani Christian,[3]

and Mallory McBride[4]

1. Survey Research and Methodology Program and the Department of Sociology, University of Nebraska–Lincoln, Lincoln, Nebraska, USA
2. Washington State University, Pullman, Washington, USA
3. Pew Research Center for the People and the Press, Washington, DC, USA
4. University of Montana, Missoula, Montana, USA

Corresponding author – Jolene D. Smyth, email jsmyth2@unl.edu

Abstract

Previous research has revealed techniques to improve response quality in open-ended questions in both paper and interviewer-administered survey modes. The purpose of this paper is to test the effectiveness of similar techniques in web surveys. Using data from a series of three random sample web surveys of Washington State University undergraduates, we examine the effects of visual and verbal answer-box manipulations (i.e., altering the size of the answer box and including an explanation that answers could exceed the size of the box) and the inclusion of clarifying and motivating introductions in the question stem. We gauge response quality by the amount and type of information contained in responses as well as response time and item nonresponse. The results indicate that increasing the size of the answer box has little effect on early responders to the survey but substantially improved response quality among late responders. Including any sort of explanation or introduction that made response quality and length salient also improved response quality for both

early and late responders. In addition to discussing these techniques, we also address the potential of the web survey mode to revitalize the use of open-ended questions in self-administered surveys.

Introduction

Our purpose in this paper is to report a series of web survey experiments that test question design techniques for improving responses to open-ended questions. In doing so, we take a holistic view of survey questions, meaning that we consider every component of the question (i.e., stem, answer spaces, answer choices, and instructions) important for obtaining quality responses. Ultimately, the goal is to produce an open-ended question in which all verbal and visual components are working together to produce the desired end product: thick, rich, descriptive responses.

Previous research suggests that both answer spaces and question stems may be important for obtaining high-quality responses in open-ended questions. Several studies have shown that the visual manipulation of increasing answer-box size for interviewers (Smith 1993) or in mail surveys (Christian and Dillman 2004; Israel 2006) leads to longer responses containing more themes. Other researchers have shown that providing clarifying and motivating instructions in interviews increases the amount of information provided in response to open-ended questions (Miller and Cannell 1982). Neither of these strategies, however, has been tested in internet surveys. Given these two strands of research, we pose three questions in this paper:

1. Does increasing the size of open-ended answer boxes increase response quality in web surveys?
2. Does adding an explanation that "you are not limited in the length of your response by the size of the box" increase response quality in web surveys?
3. Does emphasizing the importance of the open-ended response to the research project increase response quality?

Beyond these specific experiments, the larger issue that this paper speaks to is whether or not the internet provides a better venue for collecting open-ended data than mail surveys where response quality has typically been rather poor.

Methods and Analytic Procedures

The data for this paper are from three web surveys, the details of which can be found in Table 1. The surveys asked random samples of Washington State University (WSU) undergraduate students who were enrolled in classes on the main campus about their student experience. In each survey, students were randomly assigned to an experimental version, each of which contained the same questions in the same order. Students were first contacted with a personalized letter via postal mail and provided with a two-dollar incentive and individual identification number needed to access the survey. Additional reminders (up to six) were sent to all nonrespondents by postal mail and by e-mail for those for whom

an e-mail address was available (≈ 66 percent). Screenshots of the questions reported here can be seen in Figure 1 (questions appeared one per screen).

Table 1. Details of Surveys

Survey number	Date	Number of versions	Number of questions	AAPOR 1 response rate
1	Fall 2003	4	25	1,528 of 3,045 50%
2	Fall 2004	3	25	1,054 of 1,800 59%
3	Spring 2006	4	27	1,369 of 2,400 57%

Our major concern in this paper is increasing open-ended response quality which we gauge by response length, number of themes reported, elaboration on themes, response time, and item nonresponse. We are not referring to known validity inasmuch as validity is difficult, if not impossible, to ascertain for a true open-ended question. A theme was defined as a concept or subject that answered the question and was independent of all other concepts within the response. A respondent was coded as elaborating if (s)he provided additional descriptive information or explanation about a theme without introducing a new theme. The distinction between themes and elaboration has not, to our knowledge, previously been made by researchers studying open-ended questions but is considered important here because themes and elaboration are very different types of information with different uses for survey sponsors. Response time was measured from when the page loaded to when the respondent clicked the "next question" button and item-nonresponse rates were calculated.[1]

Coding proceeded in three steps. First, three of the authors each coded 10 percent of the responses to develop a list of coding rules that clearly established what was to be coded as themes and elaboration. Second, one of the authors used the established rules to code the remaining responses. Third, in a verification step, another of the authors independently coded 10 percent of the responses. The main and verification coders were in agreement about themes on 79 percent of responses and about elaboration on 83 percent of responses.

As illustration of the coding procedure described above, we provide the following response to a question from the second survey that asked why the respondent chose to attend WSU (Q15). This particular respondent answered, "I chose to attend WSU because of the people. From the first visit until now, there is no other school that matches the spirit of the Cougs. Just walk to the class and say, 'Go Cougs!!' to anyone and they will indeed say it right back." This answer was coded as having 45 words, one theme, and elaboration. The theme that was counted was "the people" and the rest of the response was considered elaboration on that theme as the respondent simply went on to explain what about the people—their school spirit—drew him/her to WSU without introducing a new, unique topic. The response time for this respondent was 50.02 seconds and he/she was assigned a value of zero for item nonresponse.

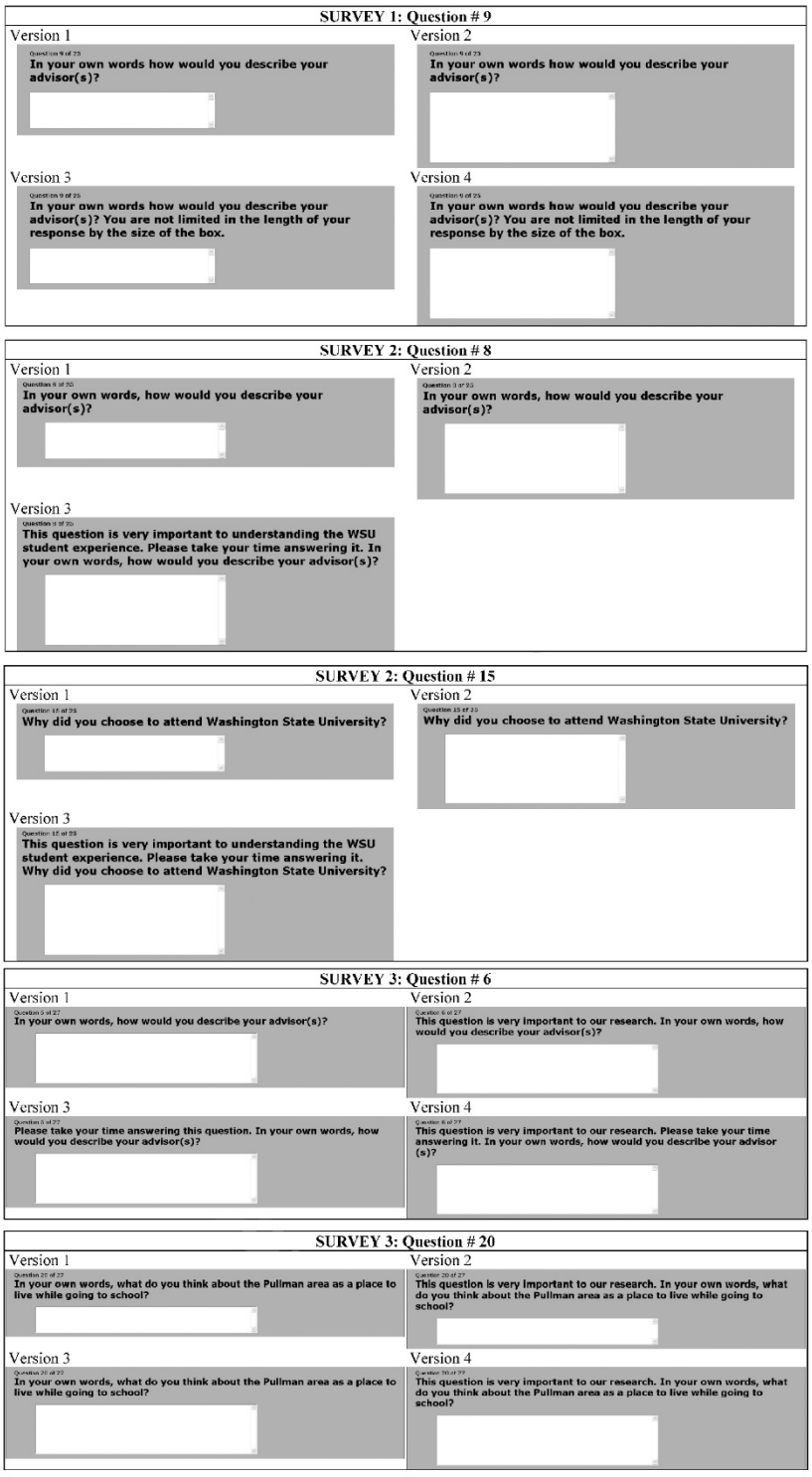

Figure 1. Open-Ended Screenshots from All Surveys

Inasmuch as respondents who reply late in the fielding period are likely less motivated overall than those who respond earlier, the steps taken here to motivate longer responses may have different effects on them than on early responders. As a result, the analyses that follow are done separately for these two response groups. Respondents who replied on or before the median response date for each survey are coded as early respondents while those who responded after this date are coded as late respondents. In all analyses, significance is gauged by a probability value less than or equal to .05.

Visual and Verbal Answer-Box Manipulations

Couper, Traugott, and Lamias (2001) and Christian, Dillman, and Smyth (2007) show that when an answer box is designed inappropriately for the response task, respondents are more likely to make errors like providing verbal explanation in an answer box that was only intended to collect numeric responses. This finding is consistent with visual design theory which asserts that respondents actively utilize the visual aspects of the questionnaire to guide them in the answering process (Schwarz 1996; Jenkins and Dillman 1997). The implication of this theory for open-ended answer boxes is that they should be appropriately sized for the type and length of answer desired. A number of studies have supported this assertion. In the first such study, Smith (1993) found that interviewers recorded significantly longer responses when an unintentional printing error resulted in larger answer spaces. Christian and Dillman (2004) experimentally tested this finding by altering box sizes for three open-ended questions in a paper survey and found that the larger boxes produced significantly longer answers with more themes. More recently, Israel (2006) examined responses to open-ended questions with answer boxes of seven different heights between 0.28 and 1.68 inches and found that response length increased linearly with box size. These studies show that larger answer boxes in paper surveys encourage longer responses. Given their similar visual nature, we might expect the same in web surveys.

It is possible, however, that space carries a different meaning on a computer screen. Whereas a piece of paper has real boundaries, computer-mediated or virtual space is seemingly limitless (i.e., we can scroll up and down in a box). This difference may make answer-box size less salient on web surveys, eliminating this type of visual design as a tool for motivating respondents. We first test the effects of answer-box size using four comparisons that were administered in multiple versions of three web surveys. In each comparison, a small box and a large box that is twice the size of the small box are compared. The specific questions on which tests were conducted can be seen in figure 1 and the versions compared to examine the effects of answer-box size can be seen in table 2.

Table 2 contains results only for late responders because among early respondents there was no significant difference between those receiving the small- and those receiving the large-box versions in the length of responses, the number of themes reported, or the percent who elaborated. Moreover, response time was only significantly different in the expected direction in one of these comparisons and there was no significant difference in item nonresponse in any of them.[2] As shown in table 2, however, late responders were influenced by box size, with those receiving the small box giving significantly shorter responses than those receiving the long box in all four comparisons (three to seven words

shorter). The shorter response length for the late respondents to the small-box treatment was reflected in significantly fewer themes in two of the comparisons and a significantly lower percentage of respondents elaborating in two comparisons. There was no significant difference in response time or item nonresponse among the late responders.

Table 2. Effects of Box Size for Late Responders

	Mean number of words	Mean number of themes	Percent elaborated	Mean response time (seconds)
Survey 1				
Q9: Small (1)	**26.1**	**2.0**	**48.2**	67.2
Q9: Large (2)	**32.6**	**2.3**	**61.5**	71.1
Difference	**6.5**	**0.3**	**13.3**	3.9
Sig. tests	**−2.29**	**−1.82**	**6.04**	−0.68
Survey 2				
Q8: Small (1)	**17.2**	1.8	58.2	51.4
Q8: Large (2)	**20.1**	1.9	66.9	60.3
Difference	**2.9**	0.1	8.7	8.9
Sig. tests	**−1.77**	−0.40	2.09	−1.57
Q15: Small (1)	**15.9**	1.9	**24.2**	52.3
Q15: Large (2)	**18.6**	2.0	**40.8**	54.1
Difference	**2.7**	0.1	**16.6**	1.8
Sig. tests	**−2.05**	−0.34	**8.64**	−0.35
Survey 3				
Q20: Small (1)	**28.8**	**2.6**	53.6	69.8
Q20: Large (3)	**35.4**	**3.1**	58.9	72.1
Difference	**6.6**	**0.5**	5.3	2.3
Sig. tests	**−2.22**	**−2.20**	0.76	−0.34

Notes: (#) The numbers in parenthesis correspond to version numbers from Figure 1. Bold values indicate $p \leq .050$. Significance tests are one-sided *t*-tests for the number of words, themes, and response time and chi-squared tests for percent elaborated.

We next test the effects of adding an explanation stating "you are not limited in the length of your response by the size of the box" to the question stem. This explanation is intended to make explicit that the boundaries of the box on the web screen are not as limiting as they may appear and that responses can exceed the space shown. As shown in table 3, the addition of the explanation did significantly increase response quality for both early and late respondents in two comparisons. Within both groups, respondents who received the version with the explanation gave significantly longer responses (between 6 and 19 words). In the first comparison, the difference in response length was due to a significant increase in the number of themes reported and the percent of respondents elaborating among late respondents. In the second comparison, the difference in response length is reflected in significant increases in both themes and percent elaborating among early responders. Finally, the presence of the explanation significantly increased response time among late responders in the first comparison and among both early and late responders

in the second comparison. There were no significant differences in item nonresponse in either comparison. This additional explanation most likely increased response quality by drawing attention to the length of responses and by implying that answers were expected to exceed the size of the box.

Table 3. Effects of Box-Size Explanation

	Mean number of words		Mean number of themes		Percent elaborated		Mean response time (seconds)	
	Early	Late	Early	Late	Early	Late	Early	Late
Survey 1								
Q9: No exp. (1)	**32.1**	**26.1**	2.4	**2.0**	55.7	**48.2**	83.5	**67.2**
Q9: Exp. (3)	**38.2**	**44.9**	2.2	**2.4**	65.4	**74.5**	87.3	**103.4**
Difference	**6.1**	**18.8**	−0.2	**0.4**	9.7	**26.3**	3.8	**36.2**
Sig. tests	**−1.81**	**−5.04**	1.18	**−2.40**	3.69	**23.7**	−0.60	**−4.84**
Q9: No exp. (2)	**28.5**	**32.6**	**2.1**	2.3	**50.5**	61.5	**74.7**	**71.1**
Q9: Exp. (4)	**46.0**	**42.1**	**2.3**	2.2	**70.0**	65.9	**101.9**	**91.1**
Difference	**17.5**	**9.5**	**0.2**	−0.1	**19.5**	4.4	**27.2**	**20.0**
Sig. tests	**−5.21**	**−2.13**	**−1.85**	0.52	**16.39**	0.69	**−3.96**	**−2.84**

Notes: The box size explanation stated, "You are not limited in the length of your response by the size of the box." (#) denotes version numbers from Figure 1. Bold values indicate $p \leq .050$. Significance tests are one-sided *t*-tests for the number of words, themes, and response time and chi-squared tests for percent elaborated.

Providing Clarifying and Motivating Instructions

Miller and Cannell (1982) propose that interviewing procedures designed to increase telephone respondents' commitment to the response tasks and provide them with clarifying and motivating instructions and feedback can improve response quality (i.e., complete and honest responses). The results of their experimentation show that telephone respondents to open-ended questions gave more information when they had agreed to a statement that they would "think carefully about each question in order to give accurate information" and when they were provided with clarifying instructions specific to questions (Miller and Cannell 1982, p. 254).

In the second survey, we tested web procedures similar to Miller and Cannell's (1982) telephone interviewing procedures by adding an introduction stating "This question is very important to understanding the WSU student experience. Please take your time answering it" to both the question asking students to describe their advisors and the question asking them why they chose to attend WSU. The introduction was designed to serve a motivating role by emphasizing the importance of the question to the study and a clarifying role by establishing the expectation for a well thought out, detailed response. In addition, it was designed to invite students to take the necessary time to provide a quality answer. Research on memory and response time generally shows that longer response times increase the accuracy of responses by allowing for a more thorough use of memory and recall strategies (Tourangeau, Rips, and Rasinski 2002).

As shown in table 4, for both questions in Survey 2 when the introduction was present, early and late respondents provided responses that were three to six words longer than those given when the introduction was not present. There was no significant difference in the mean number of themes provided in either comparison, but in the first there were significant increases of 15 percentage points among early responders and 20 percentage points among late responders in the number of respondents elaborating. Moreover, the added introduction increased mean response time from 14 to 20 seconds for both early and late responders. The introduction, however, seemed to have a negative effect on item nonresponse, increasing it from four to six percentage points for both early and late responders in both comparisons.

Table 4. Effects of Clarifying and Motivating Introductions

	Mean number of words		Mean number of themes		Percent elaborated		Mean response time (seconds)	
	Early	Late	Early	Late	Early	Late	Early	Late
Survey 2								
Q8: No intro. (2)	**21.3**	**19.9**	1.9	1.9	**70.4**	**65.5**	**69.9**	**59.8**
Q8: Intro. (3)	**26.6**	**25.9**	2.0	1.9	**85.1**	**85.1**	**89.9**	**79.1**
Difference	**5.3**	**6.0**	0.1	0.0	**14.7**	**19.6**	**20.0**	**19.3**
Sig. tests	**−3.16**	**−3.49**	−0.68	0.08	**7.67**	**12.04**	**−2.34**	**−2.65**
Q15: No intro. (2)	**19.2**	**18.6**	2.1	2.0	40.1	40.5	**54.8**	**53.9**
Q15: Intro. (3)	**23.9**	**21.6**	2.2	1.9	46.5	43.7	**71.8**	**68.3**
Difference	**4.7**	**3.0**	0.1	−0.1	6.4	3.2	**17.0**	**14.4**
Sig. tests	**−3.20**	**−2.04**	−0.95	0.34	1.16	0.27	**−2.56**	**−2.37**

Notes: (#) denotes version numbers from Figure 1. Bold values indicate $p \leq .050$. Significance tests are one-sided *t*-tests for the number of words, themes, and response time and chi-squared tests for percent elaborated.

The mixed positive and negative results of including the introduction led us to question whether informing respondents that a question was very important might have encouraged some to give more in-depth answers while asking them to take their time answering caused others to skip the question. As a result, we designed four experimental treatments for question 6 in Survey 3 to test the effects of each part of the introduction independently and then to test both parts together (see Figure 1). Because we expected the "important" introduction to produce higher quality responses, we designed two additional tests of the effectiveness of this introduction in question 20 of Survey 3. The results can be seen in tables 5 and 6.

The results for question 6 indicate that the presence of an introduction compared to no introduction increased the response length for both early and late responders regardless of its contents. Further results indicate that the "important" introduction compared to no introduction significantly increased both themes and elaboration among late but not early responders while the "take your time" introduction had no effect on themes but increased elaboration among all respondents. Using both introductions together significantly increased themes among late responders and elaboration among both early and late responders. Moreover, providing either introduction alone significantly increased the response time

from 19 to 25 seconds over the no introduction treatment while providing them both together increased the response time from 27 to 34 seconds. Finally, compared to the no introduction treatment, the presence of both the "important" introduction and both introductions together significantly decreased item nonresponse (−5.0 and −6.2 percentage points respectively) among early responders. None of the introductions significantly affected item nonresponse rates among late responders.

Table 5. Effects of Content of Clarifying and Motivating Introductions

	Mean number of words		Mean number of themes		Percent elaborated		Mean response time (seconds)	
	Early	Late	Early	Late	Early	Late	Early	Late
Survey 3								
Q6: No intro. (1)	28.1	23.0	2.6	2.2	46.2	38.0	**58.7**	52.0
Q6: "Important" (2)	37.3	37.4	2.8	2.8	53.1	56.7	**77.7**	77.4
Q6: "Take time" (3)	37.8	31.2	3.0	2.3	60.1	61.4	**80.2**	71.4
Q6: Both intros. (4)	42.5	37.5	3.0	2.7	63.6	58.9	**85.6**	85.5
No intro versus "important"								
Difference	**9.3**	**14.4**	0.2	**0.6**	6.9	**18.7**	**19.0**	**25.4**
Sig. tests	**−2.87**	**−4.02**	−0.96	**−3.60**	1.79	**10.29**	**−3.77**	**−4.23**
No intro versus "take time"								
Difference	**9.8**	**8.2**	0.4	0.1	**13.9**	**23.4**	**21.5**	**19.4**
Sig. tests	**−3.14**	**−2.76**	−1.92	−0.65	**7.48**	**14.51**	**−3.86**	**−3.17**
No intro versus both								
Difference	**14.5**	**14.5**	0.4	**0.5**	**17.4**	**20.9**	**26.9**	**33.5**
Sig. tests	**−4.54**	**−4.36**	−1.97	**−3.27**	**12.13**	**12.16**	**−4.98**	**−5.11**
"Important" versus "take time"								
Difference	0.5	−6.2	0.2	**−0.5**	7.0	4.7	2.5	−6.0
Sig. tests*	−0.14	1.63	−1.09	**2.99**	1.85	0.65	−0.41	0.91
Important versus both								
Difference	5.2	0.1	0.2	−0.1	10.5	2.2	7.9	8.1
Sig. tests*	−1.37	−0.03	−1.08	0.28	4.32	0.14	−1.35	−1.15
"Take time" versus both								
Difference	4.7	6.3	0.0	**0.4**	3.5	−2.5	5.4	14.1
Sig. tests*	−1.26	−1.77	0.07	**−2.67**	0.51	0.18	−0.87	−1.98

Notes: (#) denotes version numbers from Figure 1. A Bonferroni correction was made to adjust for multiple comparisons. With this correction, for the six comparisons made here, a p-value of ≤ .010 is needed to reach significance. Bold values are used in the table to designate comparisons that reached this level of significance. Significance tests are one-sided t-tests for the number of words, themes, and response time (except where denoted by an asterisk (∗)—two-sided) and chi-squared tests for percent elaborated. Significance tests examine differences across experimental treatments within the early and late groups.

While including an introduction versus no introduction increases response quality, comparing the introductions to each other directly reveals that most of the differences between them are not significant. For example, the only significant improvement the "important" introduction offers over the "take your time" introduction is an increase in themes among late responders. The same is true for both introductions compared to the "take your time" introduction. Overall, these findings suggest that using both introductions together

is slightly more effective, but if including such a long introduction is prohibitive, a short introduction will produce comparable gains in response quality.

The effectiveness of the "important" introduction finds mixed support among the two comparisons in question 20. In the first comparison (small box), the "important" introduction results in significant increases in the response length of 8 and 12 words for early and late responders respectively. The introduction also significantly increases the mean number of themes among both early and late responders (+0.4 themes) and the percent of early responders (but not late) who elaborate (+11 percentage points). Response time is also significantly increased among early responders (+17 seconds). In the second comparison (large box), however, the "important" introduction does not significantly increase words, themes, or elaboration for either early or late responders. The only significant effect in this comparison is a 14-second increase in response time among early responders.

Table 6. Two Additional Tests of the Effect of the "Important" Introduction on Response Quality

	Mean number of words		Mean number of themes		Percent elaborated		Mean response time (seconds)	
	Early	Late	Early	Late	Early	Late	Early	Late
Survey 3								
Q20: No intro. (1)	**33.7**	**28.8**	**3.0**	**2.6**	**55.6**	53.6	**67.7**	69.8
Q20: "Imp." (2)	**41.6**	**40.7**	**3.4**	**3.0**	**66.3**	60.8	**85.0**	77.9
Difference	**7.9**	**11.9**	**0.4**	**0.4**	**10.7**	7.2	**17.3**	8.1
Sig. tests	**−2.34**	**−3.10**	**−2.11**	**−1.96**	**4.43**	1.53	**−3.12**	−1.21
Q20: No intro. (3)	38.5	35.4	3.3	3.1	61.3	58.9	**74.7**	72.1
Q20: "Imp." (4)	43.5	38.0	3.5	3.2	63.2	62.7	**89.0**	80.8
Difference	5.0	2.6	0.2	0.1	1.9	3.8	**14.3**	8.7
Sig. tests	−1.51	−0.76	−1.13	–0.66	0.15	0.40	**−2.27**	−1.32

Notes: (#) denotes version numbers from Figure 1. Bold values indicate $p \leq .050$. Significance tests are one-sided *t*-tests for the number of words, themes, and response time and chi-squared tests for percent elaborated.

The Future of Open-Ended Questions in Self-Administered Surveys

Our purpose in this paper was to report the results of a series of web survey tests of techniques for improving responses to open-ended questions that had previously been shown to be effective in other modes. The first technique, increasing the size of the answer space, was not effective for early responders but did significantly increase response quality among late responders. This finding is encouraging because it indicates that a visual design manipulation can stimulate less motivated respondents to give more complete responses. The second and related technique of drawing attention to the flexibility of answer-box size and thereby response length improved response quality among both early and late responders.

The third technique, providing clarifying and motivating instructions, was effective at improving response quality among all respondents. The findings reported here suggest that using an introduction that emphasizes the importance of responses to the research increases response length, number of themes, elaboration, and response time and reduces

item nonresponse while taking up less screen space and reading effort than a similarly performing introduction that also asked respondents to take their time answering.

In addition to testing specific techniques for improving response quality to open-ended questions in web surveys, we can also comment on the performance of open-ended questions in web compared to paper modes. A comparison of two web questions presented here (describe your advisor and why you attend WSU) to the results from the same two questions presented by Christian and Dillman (2004) in a paper survey (also with a WSU undergraduate sample) shows that web respondents gave longer answers (advisor: +15 words; attend: +11 words) with more themes (advisor: +0.4; attend: +0.3) and more elaboration (advisor: +34.6 percentage points; attend +9.3 percentage points).[3] While these are not direct experimental comparisons, they do parallel an experimental comparison reported by Schaefer and Dillman (1998) who found that e-mail respondents used 30 more words on average to answer open-ended questions than mail respondents. This evidence suggests that web surveys produce higher quality answers to open-ended questions than do paper modes.

In the recent past, the high costs of administering open-ended questions along with their reputation for poor response quality has led some to minimize or even avoid using this question format. The findings reported in this paper, however, suggest that open-ended questions may reemerge in web self-administered surveys as an effective format for collecting thick, rich, descriptive information from respondents. The web mode by itself seems to have a higher capacity for collecting quality open-ended responses than other self-administered modes. Moreover, verbal and visual changes to the components of web open-ended questions can improve the quality of responses, similar to interviewers in telephone and face-to-face surveys, but without the high costs associated with training and maintaining quality interviewers. Overall, our findings suggest that high quality responses to open-ended questions are obtainable in web surveys.

One of the most important steps for further experimentation is to test these techniques for improving response quality in surveys of the general population. Web surveys are suitable for a college student population because almost all students have access to the internet, but college students are also a highly educated population and may be particularly attuned to written instructions. It is possible, therefore, that the results reported here are larger than we might expect in other populations. In the meantime, however, inasmuch as no negative effects of increasing answer-box size or including introductions were apparent, the potential gains of utilizing these techniques in web surveys of all populations seem to outweigh the low costs.

Acknowledgments – This paper is a revised version of a paper presented at the 2006 World Association for Public Opinion Research Meeting held during May 16–21 in Montreal, Quebec, Canada. Research reported here was conducted by the Washington State University Social and Economic Sciences Research Center (SESRC) with support from the USDA-National Agricultural Statistics Service and the National Science Foundation Division of Science Resources Statistics under cooperative agreement no. 43-3AEU-1-80055 (Don A. Dillman, Principal Investigator). Additional support for data collection was provided by the SESRC and the Gallup Organization. The authors wish to thank the reviewers for their helpful comments and suggestions on a previous draft of this paper.

Notes

1. On all time analyses, outliers were removed at two standard deviations above the mean. When respondents returned to a page multiple times, the response times were summed (most respondents only visited each page once).
2. Findings for early responders and nonresponse findings are available from the authors upon request.
3. Before making this comparison we recoded and analyzed the data reported by Christian and Dillman to verify that our new coding scheme didn't artificially create the differences. The small-box treatment from the paper and web experiments is compared here. The advisor question was asked in all three web surveys so the web results reflect averages across them.

References

Christian, Leah Melani, and Don A. Dillman. 2004. "The Influence of Graphical and Symbolic Language Manipulations on Responses to Self-Administered Questions." *Public Opinion Quarterly* 68: 57–80.

Christian, Leah Melani, Don A. Dillman, and Jolene D. Smyth. 2007. "Helping Respondents Get it Right the First Time: The Influence of Words, Symbols, and Graphics in Web Surveys." *Public Opinion Quarterly* 71: 113–25.

Couper, Mick P., Michael W. Traugott, and Mark J. Lamias. 2001. "Web Survey Design and Administration." *Public Opinion Quarterly* 65: 230–54.

Israel, Glen D. 2006. "Visual Cues and Response Format Effects in Mail Surveys." Paper presented at the Annual Meeting of the Southern Rural Sociological Association, February 7, 2006. Orlando, FL.

Jenkins, Cleo R., and Don A. Dillman. 1997. "Towards a Theory of Self-Administered Questionnaire Design." In *Survey Measurement and Process Quality*, eds. L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin. New York: Wiley.

Miller, Peter V., and Charles F. Cannell. 1982. "A Study of Experimental Techniques for Telephone Interviewing." *Public Opinion Quarterly* 46: 250–69.

Schaefer, David R., and Don A. Dillman. 1998. "Development of a Standard E-mail Methodology: Results of an Experiment." *Public Opinion Quarterly* 62: 378–97.

Schwarz, Norbert. 1996. *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Mahwah, NJ: Lawrence Erlbaum.

Smith, Tom W. 1993. "Little Things Matter: A Sampler of How Differences in Questionnaire Format Can Affect Survey Responses." GSS Methodological Report No. 78.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2002. *The Psychology of Survey Response*. New York: Cambridge University Press.