Impact of Mode on Open-ended Responses

Jennifer Edgar

Abstract

Research has found that the mode of administration can impact survey response (Epstein, Barker & Kroutil, 2001; Dillman & Christian, 2003; de Leeuw, 2005). This phenomenon has often been demonstrated when comparing Web-based surveys to other modes (Bates, 2001; Carini et al., 2003; Dillman et al, 2001). Findings have revealed that Web-based surveys have higher item-nonresponse than paper versions of the same survey (Bates, 2001; Burr, Famolaro, Levin, 2002; Vehovar, Lozar, Manfreda, 2002). The majority of these studies focused on close-ended questions, differences between modes for any open-ended questions included in the surveys were not isolated.

The level of respondent burden experienced when answering open-ended questions may vary depending on mode. For example, it may be physically easier for some respondents to type an answer than to write it on paper. On the other hand, as Hansen and Haas (1988) suggest, the computer may have a negative effect on respondents, since they must rely on the mouse and/or keyboard to provide their answers, thus respondents with poor keyboarding skills might be expected to abbreviate their response.

This paper examines the results of two studies that investigated mode differences in responses to open-ended questions. The first was a laboratory experiment designed to test for differences in response to open-ended questions based on mode. The second offered respondents a survey either on paper or on a website. Results reveal
minimal differences between mode, but interesting differences emerge between groups with different characteristics, including education and computer experience.

Background

Historically, research has found that the mode of administration can impact survey response (Epstein, Barker & Kroutil, 2001; Dillman & Christian, 2003; de Leeuw, 2005). This phenomenon has often been demonstrated when comparing Web-based surveys to other modes (Bates, 2001; Carini et al., 2003; Dillman et al, 2001). Specifically, a survey of customers of a federal government agency using both Web and paper administration found that significantly more paper respondents skipped the open-ended questions than did Web respondents (Burr, Famolaro, Levin, 2002). Other research has also found that Web-based surveys have higher item-nonresponse than paper versions of the same survey (Bates, 2001; Burr, Famolaro, Levin, 2002; Vehovar, Lozar, Manfreda, 2002). These studies focused on close-ended questions, differences between modes for any open-ended questions included in the surveys were not isolated.

The level of respondent burden experienced when answering open-ended questions may vary depending on mode. For example, it may be physically easier for a respondent to type an answer than to write it on paper. On the other hand, as Hansen and Haas (1988) suggest the computer may have a negative effect on respondents, since they must rely on the mouse and/or keyboard to provide their answers, thus respondents with poor keyboarding skills might be expected to abbreviate their response. Research has shown that the amount of answer space provided for open–ended questions can impact responses (Christian & Dillman, 2004), and it seems logical that mode may also have an effect.

Methodology

Study 1. The first study was a survey sent to users of Consumer Expenditure Survey data. They were selected from a list of known users, obtained from purchase records, contact histories, and identified by staff. Users were divided into two groups, based on the contact information available. One group had users with email addresses available (n=208), the other had users without email addresses and users with invalid email addresses (n=360). The available contact information determined which mode of the survey the user was sent, either a Web survey or a paper survey. It should be noted that a low response rate was expected, as many respondents were likely not the actual data users, rather support staff who orders the data, and would not have the information required to respond to the survey.

The content of the survey form was as similar as possible for both modes. The question wording, order and placement were the same on the Web and paper. The Web survey was created using Microsoft Word, saved as an html file on the local drive, and presented to the participants using a Web browser. The paper

version was also designed using Microsoft Word, with the same visual layout. The space provided for participants to write their answers to the open-ended questions was the same size in both versions. The follow-up procedures were also the same for both groups; two weeks after the original contact, a follow-up letter or email was sent containing the survey information. Three weeks after the original contact, a final letter or email was sent to respondents who had not returned a survey.

Study 2. The second was a laboratory experiment designed to test for differences in response to open-ended questions based on mode. Forty participants were recruited through a newspaper advertisement and screened before selection to provide a distribution along predetermined demographic variables, including gender, race, education and computer experience. Each laboratory session lasted about 30 minutes, and participants were compensated $35 for their time and travel expenses.

The content of the questions was taken from the Survey of Consumer Attitudes, a survey conducted by the University of Michigan. The purpose of this survey is to measure changes in consumer attitudes and expectations, to understand why these changes occur, and to evaluate how they relate to consumer decisions to save, borrow, or make discretionary purchases. Twenty-one questions were selected based on their expected relevance for all participants.

Included in the questions were seven open-ended questions. These were spread throughout the survey, typically following a close-ended question. For example, one question asked "Generally speaking, do you think now is a good time or a bad time to buy a house?" with "Good," "Bad," and "Don't Know" response options. The next was an open-ended question asking "Why do you say so?" The last question of the survey was an open-ended question asking for "Additional Comments?" as commonly found on survey forms.

Participants

Participants were randomly assigned a paper survey (n=20) or a Web survey (n=19) to complete. Survey questions were identical in both modes. Background information on each participant was collected. This allowed for analysis of mode differences by demographic and behavioral variables. Participants were randomly assigned to modes. This resulted in a distribution of computer usage that was approximately equal across groups. To obtain this information, participants were asked how often they use the computer. The definition of 'computer use' was left up to the participant.

Findings

Study 1.

Response Rate. Since this study was a field survey, response rates were calculated. There was a small, non-significant difference in the response rates between modes, with the paper survey having a 26% response rate, and the Web survey having a 29% response rate (calculated using AAPOR formula 1).

Item Nonresponse. There were four open-ended questions included in the study, and for all but one (number 4, the general question included at the end of the survey), the paper version had higher item nonresponse than did the Web version, as shown in Table 1. However, no differences were found to be significant using a t-test (and controlling for the multiple comparisons, $p>.01$). With all the questions combined, the item nonresponse was higher on the Web version, but this difference wasn't significant ($p>.01$) either.

Table 1 *Percent Non-Response on Open-Ended Questions*

Question	Paper Version	Web Version
1	54.8%	46.9%
2	75.3%	71.4%
3	74.0%	67.3%
4	67.1%	73.5%
Combined	68.0%	64.8%

Word Count. Overall, respondents using the Web wrote more words when answering open-ended questions than those using paper (Table 2). However, a t-test showed no significant differences between modes in the number of words respondents used to answer the open-ended questions ($p>.01$). As no significant differences were found for individual items, they were combined and compared between modes. The word count for the combined questions on the Web form was significantly higher than that of the paper form ($p<.01$).

Table 2. *Average Word Count on Open-Ended Questions*

Question	Paper Version	Web Version
1	18.6	28.7
2	23.4	28.9
3	21.7	31.6
4	31.6	47.2
Combined*	291	196

* statistically significant difference, p<.01

No significant differences were found in the minimum or maximum (Table 3) number of words respondents used to answer each question (p>.01). In addition to averaging more words per open-ended response on the Web, the maximum number of words used by Web respondents was higher than the maximum number of words used by paper respondents across all questions.

Respondents who provided an answer to the open-ended questions tended to have the same minimum word count (Table 3), regardless of forms. This may suggest the minimum number of words they believe are needed to provide a complete response, which doesn't seem to be impacted by mode.

Table 3. *Minimum and Maximum Word Count on Open-Ended Questions*

Question	Paper Version		Web Version	
	Min.	Max.	Min.	Max
1	4	48	5	99
2	5	53	4	87
3	5	50	7	69
4	5	125	6	187

Study 2.

Item Nonresponse. For Study 2, participants completed the surveys in a laboratory setting, and so were expected to answer most questions and have minimal item-nonresponse. This was found to be true. Except for the last question, which asked for "general comments," the majority of participants gave answers to the open-ended questions (Table 4). Item non-response was higher for most of the questions on the Web version, and this difference was significant for three of the questions after adjusting for the multiple comparisons; 1b, 16 and 18, (p<.006). After combining the questions, and comparing the item nonresponse, it was found that the paper respondents were significantly more likely to skip questions (p>.006).

Table 4. *Percent Non-Response on Open-Ended Questions*

Question	Paper Version	Web Version
1b	0%	10.5%*
6	0%	5.3%
14	10.0%	10.5%
16	0%	15.8%*
18	0%	10.5%*
20	15.0%	21.1%
21	40.0%	52.6%
Combined	9.4%	18.0%*

* p< 0.006

Word Count. There were no significant differences (p>.006) found between the Web and paper versions in the number of words participants used to answer the open-ended questions (Table 5). Web participants used increasingly more words as they went through the survey, with the exception of Questions 16 and 21. This may be explained by the fact that Question 21 was a general "any other comments" question, which may explain the lower word counts. Question 16 was not conceptually different than the others however; it also asked participants to explain their answer to the previous question. Combined across questions, respondents using the paper version used only slightly more words, and this difference wasn't statistically significant (p>.006).

Table 5. *Mean Word Count on Open-Ended Questions*

Question	Paper Version	Web Version
1b	15.1	10.7
6	13.1	13.3
14	14.9	16.4
16	15.7	11.8
18	13.6	15.0
20	8.3	14.4
21	8.6	6.3
Combined	12.8	12.6

Frequency of Computer Use.

Frequency of computer use may influence the impact of mode on response to open-ended questions. To investigate this possibility, participants who used the Web form were isolated, and word counts were compared by frequency of computer use, as shown in Figure 1. A trend can be seen; the more frequently a participant uses a computer, the more words they tended to use. This pattern was found in five of seven questions, and with the

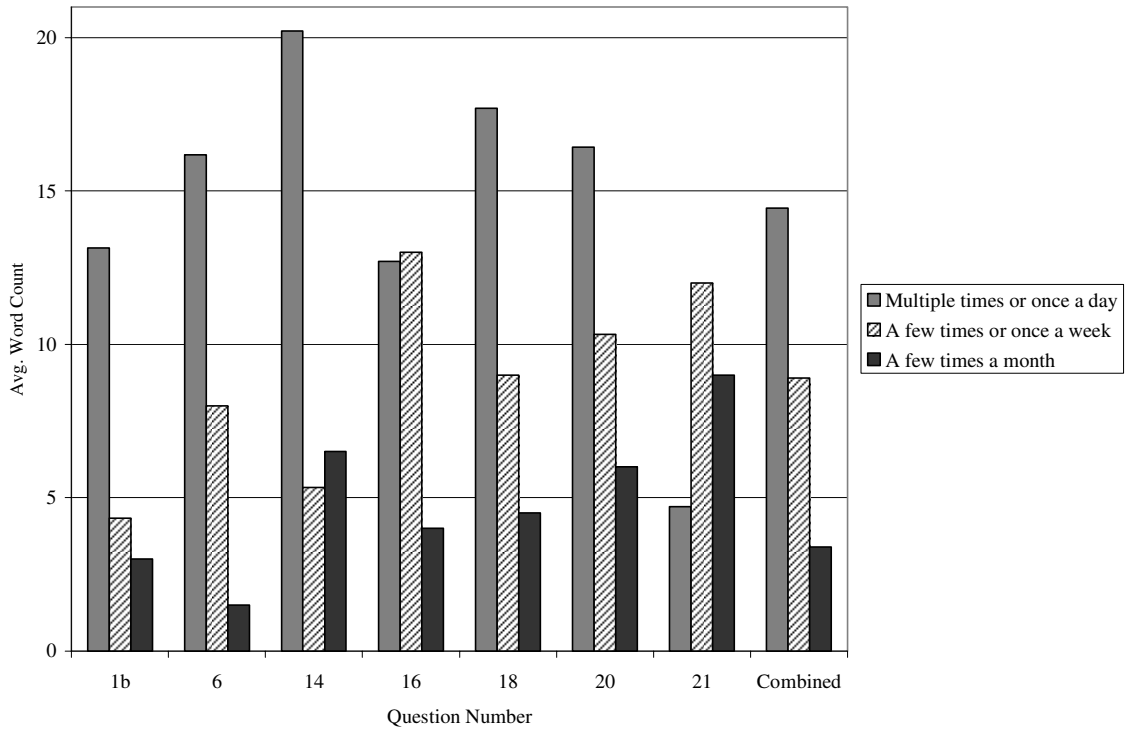

Figure 1. Average Word Count by Computer Use (Web-participants only)

questions combined; however an ANOVA revealed that none of the differences were significant (p>.006). It should be noted that most (63%) of the participants use the computer at least once a day.

Demographics. In addition to collecting information on the frequency of computer use, other demographic information was collected; including gender, age, education, income and tenure. Exploratory analyses were done to see if word counts differed by these variables, as shown in Tables 6 through 9. Initial comparisons were done with mode groups separated, but no mode differences were found so the groups were combined for reported comparisons. ANOVAs were used to test each of the differences, and no significant differences were found (p>.006), which was not surprising given the small sample size.

Table 6. *Average Word Count by Gender*

Question	Male	Female
N	20	19
1b	8.2	18.0
6	7.8	19.0
14	10.2	21.4
16	11.0	16.8
18	12.0	16.6
20	10.4	12.2
21	7.0	7.8
Combined*	9.5	16.0

* p< 0.006

As shown in Table 6, on an individual question level, there were no significant differences between males and females on the number of words they used to answer open-ended questions (p>.006), when the questions were combined, females used significantly more words than did the males (p<.006).

Table 7. *Average Word Count by Income*

Question	<15k	15-30k	30-45k	45-60k	60-75k	>75
N[a]	4	5	9	9	4	7
1b	11.0	15.2	11.6	10.1	26.8	11.6
6	5.8	14.8	13.9	13.2	17.8	14.3
14	10.0	11.0	15.1	21.1	15.5	17.9
16	12.0	14.4	13.3	18.8	12.0	11.4
18	11.3	15.2	18.9	16.4	8.8	11.6
20	9.0	9.0	13.1	12.4	8.3	13.6
21	6.3	17.8	3.6	12.8	1.3	3.4
Combined	9.3	13.9	12.8	14.9	12.9	11.9

[a] one participant did not provide this information

Although no significant differences were found between groups (p>.006), an interesting trend was seen, the number of words participants who earn between 60 and 75 thousand dollars a year used to answer the questions declined with each question (Table 7). They used more words for the earlier questions than they did the later ones. No other consistent patterns were found when comparing word count by income group.

Table 8. *Average Word Count by Education*

Question	<HS & HS	Some College	Associates & College Degree	Advanced Degree
N[a]	7	18	9	4
1b	18.7	9.4	19.4	6.8
6*	14.9	9.4	23.8	6.0
14	12.7	12.3	29.1	8.8
16	8.6	13.3	21.1	11.8
18	12.3	17.1	13.8	9.3
20	12.1	10.8	11.6	13.8
21	9.1	6.7	11.7	0.3
Combined*	12.6	11.3	16.2	10.8

[a] one participant did not provide this information
* p< 0.006

Another surprising pattern was found when comparing groups by education level (Table 8). Participants with an associates or college degree used more words than other participants for all five of the seven questions. Those with an advanced degree used fewer words in all but one question. Only the difference for question 6 was statistically significant (p<.006). A LSD post hoc test revealed that participants with associates or college degrees used significantly more words than those participants with an advanced degree (p<.006). When the questions were combined, the rests were consistent; participants with associate or college degrees used statistically more words than those with advanced degrees (p<.006).

Table 9. *Average Word Count by Age*

Question	18-5	35-45	45-55	55+
N[a]	6	15	9	8
1b	12.5	11.8	8.1	22.3
6	16.8	10.9	9.8	19.9
14	24.3	11.3	15.8	18.8
16	15.7	12.2	16.4	14.0
18	15.5	10.1	12.3	25.0
20	18.2	7.6	8.9	17.0
21	3.3	10.0	6.7	7.5
Combined*	15.2	10.6	11.1	17.8

[a] one participant did not provide this information
* statistically significant difference, p<.006

For three of the eight questions, participants over the age of 55 used more words than did the other groups (Table 9), but these differences were not statistically significant (p>.006). When the questions were combined, a significant difference was found (p<.006). A LSD post hoc test revealed that the 55 and older group used significantly more words than did the 35 to 45 year old group, and the 45 to 55 year old group.

The final analysis done to examine differences by demographic was linear regression (Table 10). Gender, age, education, income and computer use were used to predict total word count (all questions combined). Variables were run individually, and only gender was found to be a statistically significant predictor of word count (p<.001, R2 = .067). When combined with gender, computer use, income and age all created significant models (p<.001) but the R2 value was very low, none higher than .064. Finally, the Web participants were isolated, and computer use was used to predict word count to. This model was not statistically significant (p>.001).

Table 10. *Regression Analysis to Predict Combined Word Count*

Variables in Model	R2
Gender*	.067
Age	.010
Education	.006
Income	.009
Computer Use (Web participants only)	.045
Gender & Education*	.063
Gender & Computer Use*	.061
Gender & Income*	.058
Gender & Age*	.059
Gender, Computer Use & Education*	.064
Gender, Computer Use & Income*	.061
Gender, Computer Use & Age*	.061
Gender, Education & Income*	.063
Gender, Education & Age*	.063
Gender, Income & Age*	.059

* p< 0.001

Discussion & Conclusions

Study 1 found that Web respondents tended to be more likely to answer open-ended questions (a non-statistically significant finding), but Study 2 found that Web participants were more likely to skip open-ended questions. The conflicting results may be attributable to the field versus lab setting or to the random assignment to groups in Study 2 or the non-random choice of modes in Study 1. These results suggest that additional research is necessary, respondents who choose to answer a survey on the Web may be more inclined to provide answers to the open-ended questions, whereas those that are forced to use the Web, may not be comfortable using the Web , and this may cause them to skip open-ended questions that require typing rather than 'point and click' answers.

Neither study found a statistically significant in the number of words respondents used to answer questions between modes. Study 1 found that respondents tended to use more words on the Web form; however, this finding wasn't statistically significant. Study 2 found consistent results favoring Web participants, also not statistically significant. It was also consistent across studies that Web respondents used fewer words in their open-ended responses to the last question, which asked for 'additional comments.' One possible explanation is that they included their 'additional comments' in the earlier questions, where they used more words than the paper respondents did.

As no evidence was found to support the impact of mode on open-ended responses, comparisons were done to determine if different types of people were more likely to use more words when answering open-ended questions. The sample size was very small (n=39), and no differences found were statistically significant.

There were differences in word count among types of participants. With both paper and Web participant groups combined, it was found that participants with an advanced degree used the least words, participants with an associates or college degree used the most words, and participants over 55 used more words.

Finally, a series of regressions showed that although gender is a statistically significant predictor of word count, the practical significance was limited, suggesting that gender, age, computer use, education or income are not effective predictors of the number of words a respondent will use to answer an open-ended question.

Future research should be done to provide insight into the impact of mode choice on item non-response. Although these results do not show an interaction between computer use and impact of mode, it may exist. Having objective measures of computer use, specifically how and how often respondents use the computer, may provide additional information needed to identify this effect. In addition, the differences in word counts found between various demographic groups may serve as a starting point for research examining the value of open-ended questions across groups. It may be the case that open-ended questions don't collect the same type or quality of data across all respondents.

References

Bates, N. 2001. "Internet versus Mail as a Data Collection Methodology from a High-Coverage Population." Paper presented at the American Association for Public Opinion Research (AAPOR), Montreal Canada.

Burr, M.A., Famolaro, T., & Levin, K.Y. (2002). Examining Item Non-Response in a Mixed Mode Customer Satisfaction Study. The American Association for Public Opinion Research (AAPOR) 57th Annual Conference, 2002

Carini, R.M., Hayek, J.H., Kuh, G.D., Kennedy, J.M.,. and Ouimet, J.A. (2003). "College student responses to web and paper surveys: Does mode matter?" *Research in Higher Education*, 44, 1, 1-19.

Christian, L.M. & Dillman, D.A. 2004. The Influence of Graphical and Symbolic

Language Manipulations on Responses to Self-Administered Questions. Public Opinion Quarterly,68: 57 - 80.

de Leeuw, E. D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys. The Journal of Official Statistics, 21(2), 233-255.

Dillman, D. A. and Christian, L. M. (2003). Survey Mode as a Source of Instability in Responses across Surveys. Revised version of a paper presented at the Workshop on Stability of Methods for Collecting, Analyzing and Managing Panel Data, American Academy of Arts and Sciences, Cambridge, MA March 27, 2003. Forthcoming in the journal, Field Methods.

Dillman, D. A., Phelps, G., Tortora, R, Swift, K., Kohrell, J. and Berck, J. 2001. Response Rate and Measurement Differences in Mixed Mode Surveys Using Mail, Telephone,Interactive Voice Response and the Internet. Draft Paper. http://survey.sesrc.wsu.edu/dillman/papers.htm

Epstein, J.F., Barker, P.R., & Kroutil, L.A. (2001). Mode Effects in Self-Reported Mental Health Data. Public Opinion Quarterly (65), 529-549.

Hansen, W.J. and Haas, C. 1988, Reading and writing with computers: A framework for explaining differences in performance, Communications of the ACM, 31, 1080-1089.

Greer, J.G. (1988). What do open-ended questions measure? *The Public Opinion Quarterly, 52,*3, 365-371.

Vehovar, Lozar, Batagelj, Zaletel, (2002). Nonresponse in Web Surveys. In: R. Groves et all (Eds.): Survey Nonresponse. Wiley.