

INFLUENCE OF PLAIN VS. FANCY DESIGN ON RESPONSE RATES FOR WEB SURVEYS

Don A. Dillman, Washington State University; Robert D. Tortora, Jon Conradt, Gallup Organization;
Dennis Bowker, Washington State University¹
Don A. Dillman, Washington State University, PO Box 644014, Pullman, WA 99164-4014

Key Words: Web surveys, Completion rates, Questionnaire

was printed in black letters on a white screen using a traditional paper questionnaire display.

Web survey designers have taken advantage of more sophisticated software and increased computer power to design increasingly complex web surveys. These surveys utilize html tables, multiple colors, motion and other advanced features (e.g., dynamic html, animation, java-applets, and sound tracks) as a means of getting respondents to provide answers. Also, in an effort to make answering as efficient as possible for computer-literate respondents, questionnaire formats quite different from those typically used in paper questionnaires have been introduced.

The resulting web questionnaire designs may have a paradoxical effect on the conduct of scientific surveys. On the one hand they may increase the attractiveness of volunteering to complete web surveys, even making the task enjoyable so that some who respond do so just to see what sorts of interesting features are used. On the other hand these same advanced features may make questionnaires more difficult for some people to access and complete. Sophisticated survey designs contain more programming information and nontextual information like graphics. As a result, more time is requested for transmission and processing by respondent browsers. The additional information that must be transmitted causes pages to load slower. It may also overload older browsers to the point the questionnaire will not appear on the screen or it crashes. Therefore, advanced formatting using the latest programming features may cause response rates to be lower than those obtained by simpler more conventional designs, the opposite of their intended effect.

This paper reports results of an experiment in which completion rates for a sophisticated or fancy questionnaire are compared to those for a simple questionnaire. The advanced questionnaire used color, graphics for headers, html tables, and a format that departed significantly from that used in most paper questionnaires. In contrast, the simple questionnaire

Background

Interest in the design and implementation of web surveys is increasing dramatically. Cost is a major reason. Collecting responses via the web eliminates printing and mailing costs, and makes near automatic compilation of results possible. The ability to collect thousands of responses at no more cost than collecting dozens of such questionnaires has enticed a culture of web surveys to develop which tends to ignore the scientific underpinnings of surveys. Large numbers of responses are sometimes viewed as indicative of the quality of a survey. Yet from a scientific standpoint such surveys can be considered little more than large convenience samples similar to telephone call-in surveys that allow anyone who wants to respond to do so.

In order for web surveys to be scientifically sound as a basis for generalizing results to a larger population, all members of a carefully defined population need to be given a known chance of being selected to participate. In addition other sources of survey error, including nonresponse error, the extent to which respondents differ from nonrespondents, must be evaluated (Groves, 1989).

There are several interconnected reasons that fancy web questionnaire formats may contribute to the occurrence of nonresponse error. Designs that require greater transmission time, and greater browser power for processing, make it more difficult for some people to receive web surveys. As a result, it may take more time for respondents with less adequate communication lines and slower browsers to provide answers. In the case of self-administered questionnaires it has been shown that longer questionnaires usually produce lower response rates (Heberlein and Baumgartner, 1978; Dillman, Sinclair, and Clark, 1995). In addition, older, less powerful browsers seem more likely to crash, thus requiring respondents to reboot their machines and make a second or even third effort to complete the

¹Don A. Dillman is Professor of Sociology and Rural Sociology, and Deputy Director for Research and Development of the Social and Economic Sciences Research Center at Washington State University; Robert L. Tortora is Chief Methodologist, the Gallup Organizations; Jon Conradt is Application Developer, the Gallup Organization; and Dennis Bowker is Graduate Research Assistant, Department of Sociology, Washington State University. Questions about this paper should be addressed to dillman@wsu.edu.

survey. Or, the questionnaire may appear on the screen with disabled response features. Thus, we hypothesize that fancy questionnaire designs will obtain lower completion rates than simple designs that can be processed more quickly. Inasmuch as such nonresponse may be related to adequacy of telecommunications access and computer equipment, which may in turn be correlated with survey variables of interest to the sponsor, nonresponse error seems highly possible.

It's important that the effect of web questionnaire designs on nonresponse be explored. The fast-changing computer industry encourages a culture of wanting to design on the cutting-edge, using the most advanced software and equipment. Sponsors of web surveys often want their survey to be attractive, with unique qualities that viewers will enjoy. However, it is also more costly to design questionnaires that use more advanced page design features. If there is no benefit to such designs from the standpoint of response accuracy, then a great deal of money might be saved in designing such surveys.

A recent web survey of businesses at the U.S. Bureau of the Census illustrates the problem of programming that was too advanced for the survey population (Nichols and Sedivi, 1998). The desire to use JavaScript limited the universe of potential respondents to those who had Windows 95/NetScape 3.0 + configurations. Although 523 of potential respondents indicated they had web access and were willing to be surveyed on the web, only 73 of those included in this 1996-1997 survey had the capability for doing so. Among those who had adequate browsers and were ultimately offered the opportunity to respond in that way, only 68% successfully completed the web survey. In contrast, 84% of those in a comparison group completed an equivalent paper questionnaire. In addition, over half of those selected for web reporting found it necessary to call the help desk, mostly for help with usernames and ID numbers, but also because of browser configuration problems.

The experiment conducted here examines similar challenges for a much broader population. We compare for several measures of survey completion a questionnaire that used advanced page-layout techniques with a much simpler one on several aspects of completion—from time required for responding to number of questions answered.

The Experiment

A listed sample of purchasers of computer products was obtained for the survey. People on the list were called by the Gallup Organization, who conducted the survey, to determine whether these individuals were at least 18 years of age and if they had used the Internet from home, work or school for at least one application

other than e-mail in the last month. People on the list who met the screening qualification were asked to complete a web questionnaire about which websites they visited in the past, along with questions pertaining to their lifestyles and behaviors. Those who agreed were given the address of the Gallup website over the telephone, with an ID necessary for gaining access. Gallup used a five-call design for the study and the overall telephone refusal rate for people who answered the telephone was 28.7%.

To improve response rates for the web survey, all persons who received IDs for both groups were immediately sent a letter that included a \$2 bill as a token of appreciation. An e-mail with the same text was also sent the next morning. The website address and IDs were repeated in these letters. People who did not log onto the website within a week, or logged on but did not complete the survey, were sent an e-mail reminder. This special e-mail contact was repeated two more times for those who did not respond. An 800 number was provided in these messages for people who experienced problems, but very few calls were received.

The two experimental questionnaires were labeled as "plain" and "fancy". Both questionnaires had the same wording and question order. They shared several design features that Dillman et al. (1998) have proposed as principles for respondent-friendly design.

First, the questionnaire was designed so that the initial question was fully visible on the screen, and could therefore be more easily comprehended and answered by all respondents. Thus, potential first questions with large numbers of items and which required scrolling from one screen to another in order to see the full question were positioned later in the questionnaires.

Second, specific instructions on how to perform needed computer actions were provided at the point those instructions were needed, rather than being grouped together at the beginning of the questionnaire or omitted altogether. Thus, if a question required the use of radio buttons to provide answers (only one choice could be checked) the respondent was informed that to change an answer s/he needed only to click on another answer choice. But, where boxes were used (check-all-that-apply questions) the respondent was informed that changing an answer could be done by double-clicking on the marked answer.

Third, as in the case of mail questionnaires, respondents were allowed to continue to later questions without having to first provide an answer to an earlier question which, for whatever reason, they might not want to answer. This procedure is not necessary, but inasmuch as respondents may encounter questions they wish not to answer, allowing the items to be skipped

would make it possible to obtain answers to most of the remaining questions.

Several specific design features distinguished the plain and fancy questionnaires. First, the fancy questionnaire utilized html tables with sophisticated graphics, a feature requiring considerably more computer memory (Figure 1). In addition, the fancy questionnaire used bright alternating bands of color to help respondents align questionnaire items and answer choices in much the same way this feature of design has often been used in surveys that could be optically scanned (Dillman and Miller, In Press). The alternating bands were purple and pink. Both bands provided sufficient contrast between the black print and colored background that there was no difficulty reading words. These alternating bands (from black letters on pink to black letters on purple, with questions stated in black on white) resulted in a constantly changing figure/ground format. It has been argued elsewhere that frequent changes in figure/ground format makes reading comprehension more difficult (Jenkins and Dillman, 1997). The fancy questionnaire also placed answer boxes on the extreme right side of the screen close to the scroll bar (Figure 1). This presentation method was aimed at minimizing movement of the cursor back and forth across the screen between answer choices and scroll bar. It therefore contrasted with typical paper questionnaires that place answer boxes to the left of the answer categories.

In contrast, the plain version did not use html tables. All words were provided in a common figure/ground format of black letters on white background. The plain version also placed answer choices on the left side of the screen (Figure 2).

Both questionnaires used individual screen formats, with scrolling reserved for individual questions that required more than one screen for viewing. None of the questions were numbered. This choice was made because of the large number of skip questions, which meant that any question numbers that appeared on most people's screens would not have been consecutive, and a potential cause of confusion. No indicator of progress throughout the questionnaire was used, so at no time could the respondent tell how close s/he was to being finished.

Check all that apply questions were used extensively in this questionnaire. From a test standpoint this decision was useful inasmuch as it was expected that "satisficing" would occur if one of the formats was more difficult for respondents than the other format, and could be measured by number of responses given (Krosnick et al., 1996).

In sum, there were three important and practical differences between the plain and fancy versions. The fancy version used bright colors with a constantly changing figure/ground format (vs. black and white),

html tables for constructing the questions, and placement of answer categories to the extreme right of each screen. The practical result was that it took longer to transmit this questionnaire to respondents, and it appeared in a format that was judged less conventional than was the plain questionnaire. The plain version used no graphics and in total represented about 317k. The fancy version included graphics on every page and, assuming that the graphics were cached by the respondent's browser, took 959k. The required time to transmit all of the pages and graphics to a 14.4k modem (the most common at the time of the survey) would be, at least, 225 seconds and 682 seconds, respectively. Each questionnaire consisted of 173 pages, but because of large numbers of skips most respondents would skip over most of the pages.

Besides completion rates by version of the questionnaire, other key measures of experimental outcomes were as follows:

1. Last page completed. The last page (of the 173) which respondents completed was recorded. In theory all respondents should have gotten to the last page of the questionnaire, which applied to everyone.
2. Pages displayed. The total number of pages seen by a respondent was also ascertained. This number is lower if respondents marked fewer of the check-all-that-apply boxes in a way that resulted in them being directed to fewer follow-up questions.
3. Total boxes checked. The total number of boxes checked, excluding "other" boxes where an open-ended response could be provided, provides an indicator of completeness.
4. Other boxes. Several of the items allowed respondents to write in answers. The tendency to add write-in answers gives an indication of the respondent being engaged in responding. It may capture the extent to which each response was carefully evaluated. Write-ins would occur when the respondent could not find exactly what s/he was looking for.
5. Duration. This item measures the total length of time it took to complete the questionnaire. An analysis of outliers was done because of the possibility that some people quit temporarily and did not return to the questionnaire for several days. There were relatively few outliers. Virtually all respondents either completed the questionnaire the day they started or didn't return to it, and the tendency did not differ by form. Consequently, the full data set was retained for this comparison.

6. **Returns.** This item measures the number of times respondents returned to the questionnaire in order to complete it. Completing the questionnaire in one sitting was coded as zero. This measure constitutes a good indicator that the respondent's browser was overloaded, and a second try was required in order to complete the questionnaire.

Findings

There are several relevant measures of completion rates for this experiment. The first rate is the percent of those who were called that completed the screening interview, which was 71.3%. The second rate is the percentage of those that qualified and agreed over the phone to complete the web survey (regardless of whether they actually completed it). For this experiment 81.6% (9,522 individuals) of those in the plain treatment group qualified and were given an ID that would display the questionnaire. A total of 81.5% (2,466 individuals) assigned to the fancy treatment group qualified and were given an ID that would display the fancy questionnaire. These two percentages are, as expected, virtually the same inasmuch as the telephone calls to both groups were the same. Of those who agreed to complete a questionnaire, 75.9% of those assigned to the plain group logged onto that website, compared to 76.1% of those assigned to the fancy questionnaire. Again, as expected, these percentages are virtually the same.

The fourth completion rate is the percent of those who logged on that actually submitted their questionnaire after completing it. Whereas 93.1% of those assigned to the plain questionnaire group submitted their questionnaire, only 82.1% of those assigned to the fancy questionnaire did that, a difference of 11 percentage points, which is a significant difference ($p = .05$).

Thus, the response rate for the plain design was 41.1% (71.3% x 81.6% x 75.9% x 93.1%) compared to 36.29% (71.3% x 81.5% x 76.1% x 82.1%) for the fancy design. This calculation of response rates does not take into account people who could not be contacted by telephone after five contacts, but shows clearly the effect of questionnaire type on overall response.

It can be seen here that the plain version was more likely to be fully completed, with respondents on average answering 166 vs. 156 pages on the fancy form. They also displayed more pages on their screen, 46 vs. 44, but did not check significantly more boxes. Respondents to the plain questionnaire version provided significantly more open-ended answers than respondents to the fancy version. It also took respondents less time to respond to the plain version and they were much more likely to complete the questionnaire in one setting.

Table 1 presents results of the other comparisons.

Table 1. Comparison of Plain vs. Fancy Measures

Variable	Version	Mean	Std. Error	P-Value
Last page completed	Plain	166.15	0.34	0.000
	Fancy	156.07	1.19	
Pages displayed	Plain	46.46	0.11	0.000
	Fancy	44.44	0.34	
Total boxes	Plain	110.40	0.56	0.400
	Fancy	109.11	1.44	
Other boxes	Plain	3.69	0.03	0.000
	Fancy	3.43	0.06	
Duration	Plain	0.56	0.02	0.000
	Fancy	1.29	0.02	
Returns	Plain	0.45	0.05	0.000
	Fancy	0.99	0.05	

Discussion and Conclusion

These results suggest that using a plain questionnaire without color and html tables, which required less transmission time and was done in a more conventional questionnaire format, provided better results than a fancy version of that questionnaire. The plain questionnaire obtained a higher response rate, and was more likely to be fully completed, as measured by last page seen, number of pages completed, and write-in boxes completed. It also took respondents less time to complete the plain questionnaire and they were more likely to do it in one session. These data support the hypothesis that utilization of advanced page layout design features that can be used to create web questionnaires does not translate into higher completion rates or better quality data.

It cannot be concluded exactly why these differences occurred, i.e., whether respondent frustration from browser problems, the greater length of time required for receiving the questionnaire, the unconventional questionnaire format, or other differences such as choice of colors, influenced response. These are questions for future research.

However, the current results bring into question the justification for using sophisticated page-layout techniques on the basis of an expected improvement in response rates. Instead the use of these features is likely to lower response rates. Inasmuch as more programming time is required for building sophisticated page layouts, higher response rates can even be obtained at less cost by keeping page construction simple.

In the future, transmission times are likely to drop significantly and the power of most people's browsers will improve. This means that the response rate penalty associated with using currently available advanced construction techniques should decline. However, as long as development of new possibilities continues, there may continue to be a gap in what skilled web

questionnaire designers can create and what some respondents can easily access and answer.

References

Dillman, Don A. and Kent Miller. "Optical Scanning Technologies, Response Rates and Data Quality in Mail Surveys by Small Survey Organizations," in Couper, Mick et al., Computer-Assisted Survey Information Collection. New York: John Wiley.

Dillman, Don A., Robert Tortora, and Dennis Bowker. 1998 "Respondent-Friendly Principles for the Construction of Web Surveys," unpublished paper.

Dillman, Don A., Michael D. Sinclair, and Jon R. Clark. 1993. "Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys." Public Opinion Quarterly: 57: 289-304.

Heberlein, Thomas A. and Robert M. Baumgartner. 1978. "Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature." American Sociological Review: 43: 447-462.

Groves, Robert M. 1989. Survey Errors and Survey Costs. New York: John Wiley.

Krosnick, Jon, Sowmya Narayan, and Wendy R. Smith. 1996. "Satisficing in Surveys: Initial Evidence," in Advances in Survey Research: New Directions for Evaluation, 70 (Summer). San Francisco: Jossey-Bass.

Nichols, Elizabeth and Barbara Sedivi. 1998. "Economic Data Collection Via the Web: A Census Bureau Case Study." Presented at Annual Meetings of the American Statistical Association.

Figure 1. Example of Fancy Web Questionnaire Design

Figure 2. Example of Plain Questionnaire Design
