Web Surveys: The Questionnaire Design Challenge

Mick P. Couper
Survey Research Center, University of Michigan
P.O. Box 1248
Ann Arbor, MI, USA
mcouper@umich.edu

1. Introduction

Web surveys have burst onto the scene with great fanfare in the past few years. But, like other new methods of data collection, this new innovation has been greeted both with a great deal of hype and no small measure of skepticism. Now that the initial flurry of rhetoric is settling down, it remains to understand how to maximize the potential benefits of this new approach and minimize the possible drawbacks. While Web surveys present challenges in many areas, particularly sampling, coverage and nonresponse, the focus of this paper is on measurement error related to the design of the survey instrument.

In some ways Web surveys are like other methods of survey data collection, but in other ways they are unique. What are the essential characteristics of Web surveys that present unique challenges for questionnaire or instrument design? First, Web surveys are self-administered. In interviewer-administered surveys, well-trained interviewers can often explain unclear terms to respondents, keep them motivated, reassure them of the confidentiality of their answers, probe incomplete or inadequate responses, and so on. In other words, they serve as intermediaries between the researcher and the respondent. In self-administered surveys there is no such intermediary, and the survey instrument itself serves to convey the researcher's questions and expectations to the respondent. Similarly, the respondents' only means of communication with the researcher is through the medium of the instrument, and their answers and intentions must be taken at face value. This places greater focus on the self-administered survey instrument. The instrument must be easy to understand and to complete, must be designed to keep respondents motivated to provide optimal answers, and must serve to reassure respondents regarding the confidentiality of their responses.

Second, Web surveys are computerized. In contrast to mail surveys which are static instruments, Web surveys can make use of the full power of computer-assisted interviewing (CAI) methods. These include automated branching or skipping, tailored fills or question wording, range and edit checks, feedback to respondents, and so on. While not all Web survey design make use of these features, they nonetheless are powerful tools for interacting with respondents, assisting them in the completion of the task, and motivating them to continue with the survey.

In these two respects, Web surveys are like computer assisted self-interviewing (CASI) methods. However, Web surveys differ in that there is no trained interviewer present, and the survey organization does not provide the equipment. This latter point means that respondents are completing the survey on a variety of different hardware platforms and using a variety of software systems, with the potential that the instrument does not look and act in an identical manner for all respondents. This again puts more burden on design.

Finally, Web surveys have the power to extend the visual elements of presentation beyond what is usually feasible in paper surveys. Web survey instruments no longer only (or primarily) consist of verbal features (words and numbers) but can also make use of rich visual features. These visual enhancements include still and moving images, animation, line drawings, pictures, color, shapes, etc., not to mention true multimedia which includes both sound (aural) and pictures (visual) features. The graphical nature of the Web frees the survey designer from the traditional constraints of paper-based questionnaires (order, font,

color, etc.). It's not that these design features could not be used before, but they were expensive and time-consuming to develop and duplicate, and were thus used sparingly. In contrast, embedding a color photograph in a Web page is a trivial task. Similarly, changing the background color or layout of a Web survey page is virtually costless.

All these elements combine to make the Web a unique medium for the presentation of survey questions and for the elicitation of responses. In fact, the Web permits the extension beyond the traditional survey "question" to include a host of stimulus material.

While the Web provides a wonderful opportunity to "think out of the box" and expand the variety of ways information can be presented to respondents, this freedom may come at a price. For example, while images are often used in Web sites and Web surveys to enhance the user experience and motivate respondents to continue with the survey, the addition of images may have unintended consequences for the survey questions and the responses being elicited. Even when the image is explicitly designed to supplement the question text, the effect may be different than that desired. Thus, with the increased range of tools comes the possibility of inadvertently introducing measurement error.

In this paper I present several examples of the possible effects of screen layout and the use of visual images on survey responses and speculate about likely other effects. Many of the examples are serendipitous, but together with Roger Tourangeau, Fred Conrad and others, I am embarked on a program of research to explore systematically the ways in which the visual and interactive features of Web surveys may affect the answers that respondents provide.

The idea that visual features of an instrument affect respondent answers is not new to the Web. For example, Smith (1995), has some delightful examples where unintended printing variations in paper questionnaires produce large shifts in responses in both interviewer- and self-administered surveys. Similarly, Redline and Dillman's work on mail surveys (e.g., Redline & Dillman, 1999; Dillman, Redline & Carley-Baxter, 1999) has shown that the design of navigation and other visual elements can affect the quality of data obtained in mail questionnaires. Couper (2000; see also Couper & Hansen, 2001) demonstrates that poor design can affect answers in computer-assisted surveys administered by interviewers. The difference is that the Web vastly expand the range of design opportunities and that the skills brought to the design of Web surveys focus more on programming and general Web design than on Web *survey* design.

2. The Effect of Design on Measurement Error

In this section I briefly present several examples of findings on the effect of screen design on measurement error to illustrate the importance of these issues. In a Web survey of University of Michigan students, the width of a numeric entry box was inadvertently varied. A random subset of respondents received a longer box than was necessary for the task (entering a number from 0 to 9). All other features of the question were identical. The length of the box had a significant effect on the answers received. In the long box version, 20.7% of respondents made an invalid entry on one or more of the five boxes in the question, compared to 11.3% for the short box version (see Couper, Traugott, & Lamias, 2001).

A similar result was obtained in an unpublished study conducted by deRouvray, Dennis, and Couper in 2000. Using the Knowledge Networks panel, we compared short and long versions of text boxes for three open-ended questions. In each case the average number of characters typed in response to the question with the longer entry box was significantly higher than that for the shorter entry box.

Another example comes from the annual Web surveys conducted by the Georgia Institute of Technology's Graphics, Visualization and Usability (GVU) Center in 1997 and 1998. An apparently minor change in question format between the two years yielded different distributions on age. The 1997 version employed a drop box, with the "rather not say" option listed first, and individual ages below, with "under 5," "5," "6" and "7" being visible. The 1998 version used two columns of radio button categories,

starting with "under 5," "5-10," and so on, with "rather not say" being the last option. In 1997, 96 respondents (or about 1% of the sample) claimed to be between 5 and 10 years old, of whom almost all chose one of the visible age options (5, 6 or 7). Only 5 respondents chose this age range in 1998. Similarly, the placement of the "not say" option changed the percentage choosing this option from 3.1% in 1997 to 1.1% in 1998.

In an unpublished experiment, in a different Web survey of University of Michigan students conducted in 1999, Couper and Baker varied the presentation of a list of 17 reasons (plus "other," "don't know," and "refused") for drinking alcohol. In one version the 20 items were displayed in a single long column on the Web page; in the other version the items were presented in two columns. We found that significantly more items were endorsed in the top half of the single-column version and in the left column of the two-column version. For one particular item, which appeared on the lower half of the single column and on the left of the two-column version, the percent endorsing this item varied from 15.8% in the former version to 25.5% in the latter.

Finally, Couper, Tourangeau and Steiger (2001) experimentally varied the use of pictures of a male researcher, a female researcher or a neutral survey title ("Survey of Attitudes and Lifestyles") on a variety of sensitive items and gender-related attitudes. We expected respondents of both sexes to report the most pro-feminist attitudes when the program displayed pictures and messages from the female investigator and the least pro-feminist attitudes when the program displayed the pictures and messages from the male investigator. We expected the group who got the survey logo to fall in between the other two. This pattern was apparent, and reached statistical significance (F=5.52, df=1,3028, p<.05). Together, these results provide evidence that simple variations in layout affect the answers provided in Web surveys.

3. Development of Theoretical Understanding of Measurement Error Effects

Demonstrating that visual elements of a Web survey affect measurement error is only the first step. If they do so, the next step is to explore why and how these effects occur. To this end, we need to develop a theoretical understanding of the processes or mechanisms underlying these effects. In this section I offer some brief observations to facilitate the process of developing such theoretical propositions.

Following Ware (2000) one can draw a distinction between verbal or textual information, on the one hand, and visual information, on the other. In surveys, the verbal information encompasses question wording, instructions, and certain auxiliary cues, such as the numbers used to label the scale points. The visual information includes standard features present in all questionnaires such as type font and size, color, position on the page, layout, and graphical symbols; it also encompasses such additional features as images, line drawings, and animation. Furthermore, the psychological evidence is that verbal and visual information are stored and processed in parallel (e.g., see Paivio's, 1986, dual coding model, which argues that both working and long-term memory use the two types of representation).

Given this, there are many ways that the visual and verbal features can interact to shape respondents' understanding of the questions and influence their responses. We believe that there are three potential problems emerging from this interaction:

- The visual information, intended as an embellishment to the text, in fact draws attention away from more important verbal information;
- The visual information is the basis for unintended inferences that lead to misinterpretations of the questions or the navigational path;
- The visual information conflicts with the verbal information, slowing respondents down and leading to incorrect responses.

But these effects presuppose a relevance or link between the visual and verbal information. The likely effect of the visual information on the verbal task (responding to the survey question) may depend on the degree of coupling or linkage between the two. For example, banner ads are often viewed as irrelevant to many

Web-based tasks, and are consequently ignored. In addition, interference is likely to occur only if the visual and verbal elements convey *different* information. To the extent that the two types of information are congruent, the likely effect is one of accentuation; however, if the information is incongruent, then interference effects are more likely.

The use of images (photographs or drawings) is only one way to enhance the visual appeal of a Web questionnaire. A variety of other design tools–textured or colored backgrounds, font variations, use of lines and boxes, etc., in addition to the layout effects presented earlier–are common features of Web surveys. These too may affect the way respondents answer survey questions on the Web. Studies by Schwarz and his colleagues (e.g., Schwarz, Grayson, & Knäuper, 1998) demonstrate that potentially *all* information can be seen by respondents as relevant to the task of completing the survey. Space precludes a detailed discussion of examples at this time, but we have several projects currently underway to test a variety of propositions about various layout features in Web surveys.

In summary, the Web provides wonderful opportunities for extending the range of questionnaire features available to the survey designer. However, such features can and do appear to affect measurement error. It is thus important to understand when and why these effects occur in order to most effectively use the rich visual features available in Web surveys. Our goal should be to enrich the survey experience for respondents and extend the range of measurements than can be used in surveys without inadvertently compromising data quality.

REFERENCES

Couper, M.P. (2000), "Usability Evaluation of Computer Assisted Survey Instruments." *Social Science Computer Review*, 18 (4): 384-396.

Couper, M.P. & Hansen, S.E. (2001), "Computer Assisted Interviewing." In J.F. Gubrium and J. Holstein (eds.), *Handbook of Interviewing*. Thousand Oaks, CA: Sage, forthcoming.

Couper, M.P., Tourangeau, R., & Steiger, D.M. (2001), "Social Presence in Web Surveys." Paper presented at CHI ' 01, Conference on Human Factors in Computing Systems, Seattle, April.

Couper, M.P., Traugott, M., & Lamias, M. (2001), "Web Survey Design and Administration." *Public Opinion Quarterly*, forthcoming.

Dillman, D.A., Redline, C.D., & Carley-Baxter, L.R. (1999), "Influence of Type of Question on Skip Pattern Compliance in Self-Administered Questionnaires." *Proceedings of the American Statistical Association, Survey Research Methods Section*.

Redline, C.D. & Dillman, D.A. (1999), "The Influence of Symbolic, Numeric and Verbal Languages on Navigational Compliance in Self-Administered Questionnaires." Paper presented at the International Conference on Survey Nonresponse, Portland, OR, November.

Schwarz, N., Grayson, C.E., & Knäuper, B. (1998), "Formal Features of Rating Scales and the Interpretation of Question Meaning." *International Journal of Public Opinion Research*, 10 (2): 177-183.

Smith, T.W. (1995), "Little Things Matter; A Sampler of How Differences in Questionnaire Format Can Affect Survey Responses." *Proceedings of the American Statistical Association, Survey Research Methods Section*, pp. 1046-1051.

Ware, C. (2000), *Information Visualization: Perception for Design*. San Francisco: Morgan Kaufman.

RESUME

This paper focuses on the effect of visual features on Web surveys on measurement error. Visual elements of a Web survey include the design and layout of the screen as well as the use of images such as pictures or drawings. Given the wide array of such features available to the Web designer, this paper

explores the impact of such non-verbal elements on the answers that respondents give, and why they may do so.