Public Opinion Quarterly, Vol. 71, No. 4, Winter 2007, pp. 623–634

VISUAL CONTEXT EFFECTS IN WEB SURVEYS

MICK P. COUPER
FREDERICK G. CONRAD
ROGER TOURANGEAU

Abstract There are many examples of context effects in survey measurement. Responses to survey questions can be shaped by the order of questions, the format of response options, the broader survey environment, and so on. For Web surveys, the inclusion of visual images is a trivial design issue, but may have consequences for the responses obtained because they change the visual context. We report a series of experiments examining how responses may be affected by the use of images in Web surveys. Specifically, we examine the effect that pictures of a healthy woman exercising versus a sick woman in a hospital bed have on self-rated health. We replicated the experiments in three different surveys, varying such factors as the size and placement of the image and the location of the question within the questionnaire. In general, we find that when exposed to a picture of a fit woman, respondents consistently rate their own health lower than when exposed to a picture of a sick woman.

Introduction

"Every picture tells a story" Rod Stewart (1971)

The Web is a rich visual medium. The delivery of full-color images is a trivial matter in HTML, and bandwidth generally poses little constraint on the use of images in Web surveys. Many survey researchers are exploiting this feature in an attempt to enhance the experience for respondents and to improve survey measurement. But is this a good idea? This paper offers a cautionary tale regarding the inclusion of images in Web surveys.

MICK P. COUPER, FREDERICK G. CONRAD AND ROGER TOURANGEAU are with the Institute for Social Research, University of Michigan, P.O. Box 1248, Ann Arbor, MI 48109, USA and with the Joint Program in Survey Methodology, College Park, MD 20742, USA. Address correspondence to Mick P. Couper; (e-mail: mcouper@umich.edu); Frederick G. Conrad; (e-mail: fconrad@umich.edu); or Roger Tourangeau (e-mail: RTourang@survey.umd.edu)

doi:10.1093/poq/nfm044

There are many different uses of images in Web surveys. First, images may be used to replace words in survey questions, providing the visual stimulus that forms the core of the question. Examples include ad recognition studies, scenic preference studies, and so on. Second, images may be used to supplement the survey question, e.g., by clarifying meaning. Many examples exist in Web surveys, ranging from illustrations of objects or products to accompany a description, to explanations of a procedure or task to facilitate respondents' understanding. Third, images may be used to motivate or entertain the respondent, or provide branding for the site or survey. In this case, the images are not intended to influence the responses, but rather to increase participation or reduce breakoffs. Given the ease with which images can be added to Web surveys, designers are often tempted to embellish the surveys with images whose purpose and relation to the question may be unclear.

The difference between the use of images as supplements versus embellishments may be lost on respondents. What a designer may intend as unrelated may be viewed by the respondent as relevant to the question (see Schwarz 1996). Similarly, supplemental images designed to support the survey question may be disregarded by respondents as not relevant to the task if they are not clearly linked to the question. How the images are perceived in relation to the task of answering the survey question is an issue of visual language.

Horn (1998, p. 8) defines visual language as "the integration of words, images, and shapes into a single communication unit." Similarly, Ware (2000, p. 311) notes that most visualizations are not purely graphical, but are composites, combining images with text or spoken language. Images are usually not seen in isolation. When the visual and verbal elements are congruent, they can serve to reinforce each other. But, to quote Horn again, "[What] happens to our thinking process when we place next to one another visual and verbal elements that are not obviously related or not related at all? Almost all of us will struggle valiantly to create meaning out of nothing, to understand the close juxtaposition of the idea conveyed by the visuals and the words. . . . We must be vigilant in visual language, perhaps more so than in prose, because readers will attempt to make meaning out of elements whose proximity may be accidental" (Horn 1998, p. 111).

Given the power of images to command attention, we need to pay attention to the syntax of visual communication, or how the words and images are linked together.

In addition to the purpose of the image and its connection to the survey question, the content and form of the image also plays an important role in its effect on the answers that a respondent may provide. Images range in complexity or richness, from simple line drawings to full-color photographs of complex scenes. The more complex an image is, the greater the variety of meanings it may potentially convey. As the well-known aphorism suggests, "a picture is worth a thousand words." Not only is the content of the image—what information the image conveys—important, but so too is the form of the image.

Features such as the size, placement, vividness, or richness of the image may all serve to draw attention to the image and thereby to the information contained in the image.

Research on Images in Web Surveys

The attention-seeking effects of images are well known. For example, Knobloch et al. (2003) exposed subjects to online news articles with or without pictures. When pictures were included together with the headlines, subjects were more likely to select the story and spent longer on the story than when no pictures were presented. Furthermore, the effect was larger for threatening images than for innocuous images. Sargent (2007) found similar results. In a Web-based survey on political knowledge, Prior (2002) found that adding photographs of political figures significantly improved performance on political knowledge tests.

The inclusion of images can also shape the overall context of the survey, potentially affecting both nonresponse and measurement errors. For example, Witte et al. (2004) asked a series of questions about support for the protection of four threatened and endangered species. The questions were either accompanied by a photograph of the species in question or not. In each case, the presence of the photograph significantly increased support for protection of the species by about 5 percentage points on an average.

Couper, Tourangeau, and Kenyon (2004) looked at the effect of images on behavioral frequency reports. The pictures accompanying the survey question represented low-frequency or high-frequency instances of the behavior in question (shopping, travel, eating out, etc.). For example, when shown a picture of a couple dining in a fine restaurant, respondents reported significantly fewer episodes (eating out in the past month) on average than when they were shown a high-frequency instance of the behavior (eating fast food in a car). Similarly, respondents exposed to a picture of grocery shopping reported more shopping trips in the past month than those exposed to a picture of clothes shopping. The hypothesis was that the images served as cues for the retrieval of relevant incidents from memory, hence affecting the frequency reports.

These studies suggest that the inclusion of images in a Web survey may affect respondents' answers. The decision to use images in a Web survey, and the choice of images used, is consequential for measurement error. The focus of this paper is on visual context effects. Context effects are well known in the survey literature (see Tourangeau, Rips, and Rasinski 2000, Chapter 7, for a review). In most cases, the context of a question is defined as the prior questions in the survey, although Schwarz et al. (1991) have demonstrated that answers can be affected not only by the wording of the question, but also by the visual presentation of the question on the page or screen (see also Tourangeau, Couper, and Conrad 2004). Context studies have found both assimilation

(in which the judgment of the target becomes more like the judgment of the context stimulus) and contrast effects (in which the judgment of the target moves away from the judgment of the context stimulus). According to one theory (Schwarz and Bless 1992), assimilation effects occur when the context brings material to mind that gets incorporated into the respondent's representation of the target; contrast effects result when accessible material that might otherwise affect the judgment is excluded from the representation of the target or is used as a standard to which the target is compared.

There is no reason to think that images or pictures will not have as much impact on the responses as earlier questions do. The study by Couper and his colleagues described earlier represents an assimilation effect produced by images in a Web survey (Couper, Tourangeau, and Kenyon 2004). An earlier study (Manis, Biernat, and Nelson 1991) finds both assimilation and contrast effects with images. We thus extend the notion of context to include visual cues that may appear on the Web page while the respondents are completing a survey, and thereby shape their answers to survey questions. In particular, we are interested in the effects of including photographs in Web survey. We explore both the syntactic and semantic aspects of visual language; that is, both the content of the image and the degree of linkage to the words comprising the survey question.

Study 1

Our first study was designed to explore whether an image could produce contrast effects, by explicitly suggesting a comparison for a judgment of self-rated health. We hypothesized that when exposed to a picture of an obviously healthy person, respondents would tend to rate their own health lower than if they were exposed to a picture of an obviously ill person. In addition to varying the content of the image, we were interested in the proximity of the image to the target question. Specifically, we hypothesized that the closer the image was to the question, the larger the effect of the image content on the ratings.

EXPERIMENTAL DESIGN

The experiment was a 2×3 factorial design, with the following levels of each factor:

1. Image content: a) a sick woman in a hospital bed; and b) a healthy woman jogging.
2. Image position: a) the center of the header, separated from the question by a line; b) the question area, immediately above the question stem on the left of the screen; and c) on a prior screen introducing the series of questions, on the left of the screen.

A no-picture condition was added as a control, though our focus is on the six cells created by crossing image content and position. The image manipulations are illustrated in the Online Appendix. The target question is a global health rating, "How would you rate your health?" The response scale was a 5-point fully labeled scale ranging from "Excellent" to "Poor."[1]

SAMPLE AND DATA COLLECTION

The experiment was embedded in a larger survey containing experiments on several other design issues. All experiments were fully crossed to avoid any possible confounding and to allow us to look for carryover effects. The instrument was programmed in SPSS's mrInterview by staff at Market Strategies, Inc. (MSI). The sample was obtained from Survey Sampling, Inc. (SSI). SSI has two Web sample sources:

- SSI's Survey Spot sample is an opt-in Web panel of almost one million persons who have signed up online to receive survey invitations.
- SSI's eLite is an aggregated list of over seven million records from the opt-in lists of a wide variety of Web sites. It is not a panel in the usual sense, but a large list of persons who have given their permission to be contacted for a variety of purposes, including surveys.

We used the sample selected from both sources. Sample persons were invited to participate by e-mail and nonresponders received one reminder. The URL directed sample persons to the MSI Web site.

The survey was conducted in March/April, 2003. Of the 7838 persons invited in the Survey Spot panel, 1555 began the survey and 1379 completed it, for a response rate of 17.6 percent. For the eLite panel, 31,379 invitations led to 1624 starting the survey and 1343 completing it, for a 4.3 percent response rate. The combined response rate is 6.9 percent and completion rate is 85.6 percent.[2] All analyses are based on the combined set of 2722 persons who completed the survey, but the results do not change if we include the partials. We note that this is not a probability sample. Our goal, however, is randomization rather than representation, to use Kish's (1987) terminology.

ANALYSES

Table 1 shows the mean health ratings by condition. The scale was scored so

1. The study included an experimental manipulation of the response scale. The other version was also a 5-point fully labeled scale ranging from "Extremely good" to "Extremely poor." Similar results were obtained for both scales, and there were no interactions of scale type with the image manipulations, so we present the results for the standard scale here.
2. Response rates reported in this paper are calculated using the RR1 formula from the AAPOR *Standard Definitions*.

Table 1. Means and Standard Errors for Image Content and Position, Study 1.

	Sick		Fit	
Position of image	Mean	S.E.	Mean	S.E.
Introduction	3.29	(0.067)	2.93	(0.066)
Header	3.14	(0.075)	3.29	(0.065)
Question area	3.30	(0.067)	3.05	(0.066)
Total	3.25	(0.040)	3.10	(0.038)

that a high score (5) indicates excellent health while a low score (1) indicates poor health.

As expected, the images did indeed affect the self-ratings of health, producing lower ratings on average for the respondents who got the picture of the healthy woman (mean = 3.10) than those who got the picture of the sick woman (mean = 3.25). The main effect of the picture in the 3×2 experimental design was statistically significant ($F = 7.86$, d.f. = 1, 1155, $p < 0.01$).

There was also a significant interaction between the position of the picture and its content. The overall model is statistically significant ($F = 5.26$, d.f. = 5, 1151, $p < 0.0001$) as is the interaction term ($p < 0.001$). When the picture is in the header, the difference between the two versions is not significant ($p > 0.10$). In both of the other position conditions, the difference between the fit and sick picture is statistically significant ($p < 0.01$). In each case, we find the hypothesized contrast effect, with respondents rating their health lower when exposed to the picture of the healthy woman, and higher when exposed to the picture of the sick woman.

One explanation for the lack of effect of the picture content in the header condition may be "banner blindness." (Benway and Lane 1998; see also (Benway 1998) first identified the phenomenon of banner blindness—the fact that people ignore the banners at the top of Web pages, despite their being designed to be attention getting. However, Bayles (2000) provides counter-evidence, and Pagendarm and Schaumburg (2001) suggest that the attention paid to the banner may depend on the task. They argue that people who are browsing aimlessly are more susceptible to noticing a Web banner, "because the banners' color contrast or animation trigger an orientation reaction that is followed by a bottom-up process of information processing." On the other hand, subjects who are searching for information seem to apply cognitive schemata that suppress a deeper processing of Web banners. Given that our experiment involved the first question in the questionnaire, it is not clear whether respondents are still in browsing mode or are focused fully on the survey question. This issue is worth further exploration.

We were a little surprised by the effect of the picture when placed on the prior screen, expecting it to have a smaller effect than when explicitly linked with the

target question on the same screen. However, the literature on visual priming (e.g., Scharlau 2002) suggests that the effect of the image may be maximal when it appears just before exposure to the target item. In other words, the image may have activated a comparison to the subject in the photograph even before the respondent read the survey question.

In summary, we found support for the hypothesized contrast effect generated by the content of the image. We also found that the effect depended on the visual proximity between the image and the target question. Specifically, when the image appears in the header area, it may either be ignored altogether (not seen) or viewed as irrelevant to the task.

Study 2

The second experiment was designed as a partial replication of the first experiment. We were intrigued by the apparent banner blindness effect and wondered if the nonsignificant assimilation effect we observed was an anomaly. In the current experiment, we placed the image either in the banner or in the question area.

EXPERIMENTAL DESIGN

We used the same pictures as in Study 1, but dropped the prior-screen condition. This results in a 2×2 factorial design, as follows:

1. Image content: a) sick woman in a hospital bed; and b) healthy woman jogging.
2. Image position: a) header; and b) question area.

A 5-point fully labeled scale was used for the same global health rating question, with the responses ranging from 1 = Poor to 5 = Excellent.[3]

SAMPLE AND DATA COLLECTION

This experiment was again embedded in a larger Web survey containing several other design experiments. Randomization for each experiment was done independently to reduce the possibility of contamination from carryover effects. Two sample sources were used:

- SSI's Survey Spot, an opt-in Web panel of 1 million + volunteers.
- AOL's Opinion Place, a "river sampling" or intercept approach using banner advertisements to invite Web users to participate in various Web surveys.

3. Two other scale versions were also tested: 1) a 5-point end-point labeled scale and 2) 10-point end-point labeled scale. Similar results were obtained.

Table 2. Means and Standard Errors for Image Content and Position, Study 2.

	Sick		Fit	
Position of image	Mean	S.E.	Mean	S.E.
Header	3.37	(0.063)	3.23	(0.058)
Question area	3.41	(0.056)	3.25	(0.062)
Total	3.39	(0.042)	3.24	(0.043)

By definition, the latter group had noticed at least one banner—the one inviting them to the survey—immediately prior to starting the survey. The survey was again hosted by MSI, using SPSS's mrInterview software. A total of 13,216 members of the SSI panel were invited to the Web survey, of which 1695 or 12.8 percent responded to the invitation, and 1427 or 10.8 percent completed the survey (for an 84 percent completion rate). The number of invitees from Opinion Place is not known, as invitations are extended on a flow basis until the desired number of respondents (1500 in our case) is obtained. Of the 1500 who started the survey, 1290 completed it, for an 86 percent completion rate.

ANALYSES

We again find a main effect of picture content, in the expected direction ($F = 6.41$, d.f. $= 1, 931$, $p < 0.05$). However, we observed contrast effects regardless of the position of the picture. The interaction between picture content and position is not statistically significant ($p > 0.10$). Table 2 shows the mean health ratings by picture content and position. The overall contrast effect is consistent with that found in the first experiment, but this time we find no effect of the picture position. In other words, we find no support for the banner blindness hypothesis in this partial replication of the image position manipulation.

Study 3

Our third study was designed to further explore the puzzling findings about the placement of the image in the header area. The first experiment provided some support for the banner blindness hypothesis, but the second did not. In both cases, the target item was the first question in the survey. One possible explanation is that the survey format is sufficiently different from a typical Web page that respondents attend to the header but only on the opening few screens. In other words, respondents in our first two studies had not had time to get habituated to the banner over the course of the survey. Thus, in addition to varying the position of the image, we also varied the location of the question in the questionnaire. It could also be that the appearance of an image in the header later in the survey is sufficient change from the usual imageless banner

for respondents to notice the image. We also added a condition where we varied the size of the image in the question area, with the expectation that the larger photograph would command more attention and therefore have a bigger effect on the responses. In the earlier experiments, the images were 73 × 108 pixels. We doubled each dimension to 146 × 216 pixels to produce images four times larger.

EXPERIMENTAL DESIGN

This time we employed a 2 × 3 × 2 factorial design, with the following experimental factors and levels:

1. Image content: a) sick woman in a hospital bed; and b) healthy woman jogging.
2. Image position and size: a) the header, separated from the question by a line; b) small image in the question area; and c) large image in the question area.
3. Location of the experimental questions: a) at the beginning of the survey (Question 1); and b) later in the survey (Question 26).

The pictures used were the same as in Study 1. The large image version is illustrated in the Online Appendix. For this experiment, we used a 10-point unlabeled scale, with responses ranging from 1 = Poor to 10 = Excellent.

SAMPLE AND DATA COLLECTION

We again used samples from two sources—SSI's Survey Spot and AOL's Opinion Place. The survey was again hosted by MSI, using SPSS's mrInterview software. A total of 29,772 invitations were sent to Survey Spot members, resulting in 1498 starting the survey and 1361 completing it for a response rate of 4.6 percent and a completion rate of 90.9 percent. The number of invitees from AOL is unknown, but we obtained 1333 initial responses to the invitation and 1226 completions for a 92.0 percent completion rate.

ANALYSES

The responses to the 10-point scale were first converted to a 5-point scale with a range of 1–5 to permit more direct comparison to the results of the first two studies.[4] There is a significant ($p < 0.001$) main effect of the location of the image experiment, with those getting the questions at the beginning reporting better health than those who got them later in the questionnaire. However, we find no evidence that the presence of the images in the header had a different

4. This was done in 3 steps: 1) subtract 1 from the score; 2) multiply this value by 4/9; and 3) add 1 to the result.

Table 3. Means and Standard Errors for Image Content and Size/Position, Study 3.

	Sick		Fit	
Position and size of image	Mean	S.E.	Mean	S.E.
Header	3.66	(0.041)	3.60	(0.066)
Question area (small)	3.73	(0.040)	3.61	(0.065)
Question area (large)	3.66	(0.046)	3.55	(0.066)
Total	3.68	(0.024)	3.59	(0.025)

effect when they appeared early versus late in the survey. We thus combine the results from the two locations, and present the combined means and standard errors for the content and size/position manipulations in table 3.

Again we find a significant ($F = 7.09$, d.f. $= 1, 2578$, $p < 0.01$) contrast effect produced by the content of the image; those seeing the picture of the sick woman report themselves to be in better health than those exposed to the picture of the fit woman. Contrary to expectation, the effect of the image is no stronger when a larger image is used. In addition, while the effect of the image appears to be smaller when the picture is the header, this interaction fails to reach statistical significance ($p > 0.10$). Thus, we find only modest support for the banner blindness hypothesis. Varying the location of the experiment in the survey also does not shed any further light on the effect of picture placement on the responses to the health question.

General Discussion

Across three experiments, we have found consistent visual context effects. These effects appear to be robust across the size and placement of the image and the position of the question in the questionnaire.

There is some evidence—although not statistically reliable—of banner blindness. In the first and third experiment, we find evidence that the effect of the image is diminished when it appears in the header, but we find no reduction of the effect in the second experiment, and the interaction term does not reach statistical significance in the third experiment. Changing the location of the experimental questions to tease out the possible effects of habituation to the header did not bring clarity to these inconsistencies across experiments. We are left with the explanation offered by Pagendarm and Schaumburg (2001) that attention to the banner or header may depend on the respondent's level of engagement in the task.

The strong and consistent effect of the inclusion of an image on the responses to a straightforward self-assessment of health suggests caution in including images in Web or other kinds of surveys. Many Web surveys include images

for motivational or aesthetic reasons, without considering the consequences for measurement. Sometimes these effects occur even when the image is included in the survey header and not ostensibly linked to the survey questions. While the samples used in these experiments are self-selected, and the images are deliberately selected to test different possible interpretations of "health," there is no reason to expect the effects to differ in other kinds of online surveys.

The use of images is a powerful tool for the survey design to enhance the measurement process, but given the richness of many images—especially photographs—and their power to attract attention, these findings suggest that care should be taken when including images in Web surveys. Images are contextual stimuli and, like prior questions, they can systematically affect responses when their content has relevance to the survey question.

Acknowledgements

The work reported here was supported by grants from the National Science Foundation (SES-0106222) and the National Institute for Child Health and Human Development (R01 HD041386-01A1). We are grateful to Reg Baker and Duston Pope of Market Strategies, Inc., for their help in designing and implementing these studies, and to Mirta Galesic, Andy Peytchev, and Ting Yan for their help with the analyses. We thank the editor and reviewers for their helpful comments on earlier drafts. All errors remain our own.

Supplementary Data

Supplementary data are available online at http://pubopq.oxfordjournals.org/.

References

Bayles, Michelle. 2000. "Just How 'Blind' Are We to Advertising Banners on the Web?" *Usability News*, Summer. Available at http://psychology.wichita.edu/surl/usabilitynews/2S/banners.htm.

Benway, Jan Panero. 1998. "Banner Blindness: The Irony of Attention Grabbing on the World Wide Web." Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting, 463–467. Santa Monica, CA: HFES.

Benway, Jan Panero, and David M. Lane. 1998. "Banner Blindness: Web Searchers Often Miss "Obvious" Links." *Internetworking Newsletter*, ITG 1.3, December. Available at http://www.internettg.org/newsletter/dec98/banner_blindness.html.

Couper, Mick P., Roger Tourangeau, and Kristin Kenyon. 2004. "Picture This! An Analysis of Visual Effects in Web Surveys." *Public Opinion Quarterly* 682:255–266.

Horn, Robert E. 1998. *Visual Language; Global Communication for the 21st Century*. Bainbridge Island, WA: MacroVU, Inc.

Kish, Leslie. 1987. *Statistical Design for Research*. New York: Wiley.

Knobloch, Silvia, Matthias Hastall, Dolf Zillmann, and Coy Callison. 2003. "Imagery Effects on the Selective Reading of Internet Newspapers." *Communication Research* 30(1):3–29.

Manis, Melvin, Monica Biernat, and Thomas F. Nelson. 1991. "Comparison and Expectancy Processes in Human Judgment." *Journal of Personality and Social Psychology* 61(2):203–211.

Pagendarm, Magnus, and Heike Schaumburg. 2001. "Why Are Users Banner-Blind? The Impact of Navigation Style on the Perception of Web Banners." *Journal of Digital Information* 2(1). Available at http://jodi.tamu.edu/Articles/v02/i01/Pagendarm/.

Prior, Markus. 2002. "More Than a Thousand Words? Visual Cues and Visual Knowledge." Paper presented at the annual meeting of the American Political Science Association, Boston, MA, August.

Sargent, Stephanie Lee. 2007. "Image Effects on Selective Exposure to Computer-Mediated News Stories." *Computers in Human Behavior* 23(1):705–726.

Scharlau, Ingrid. 2002. "Leading, but Not Trailing, Primes Influence Temporal Order Perception: Further Evidence for an Attentional Account of Perceptual Latency Priming." *Perception and Psychophysics* 64(8):1346–1360.

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

Schwarz, Norbert, and Hubert Bless. 1992. "Constructing Reality and its Alternatives: Assimilation and Contrasts Effects in Social Judgment. In *The Construction of Social Judgment*, eds. L.L. Martin and A. Tesser, pp. 217–245. Hillsdale, NJ: Erlbaum.

Schwarz, Norbert, Bärbel Knäuper, Hans-J. Hippler, Elisabeth Noelle-Neumann, and Leslie Clark. 1991. "Rating Scales: Numeric Values May Change the Meaning of Scale Labels." *Public Opinion Quarterly* 55:618–630.

Tourangeau, Roger, Mick P. Couper, and Frederick G. Conrad. 2004. "Spacing, Position, and Order: Interpretive Heuristics for Visual Features of Survey Questions." *Public Opinion Quarterly* 68(3):368–393.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge, U.K.: Cambridge University Press.

Ware, Colin. 2000. *Information Visualization: Perception for Design*. San Francisco, CA: Morgan Kaufman.

Witte, James C., Roy P. Pargas, Catherine Mobley, and James Hawdon. 2004. "Instrument Effects of Images in Web Surveys: A Research Note." *Social Science Computer Review* 22(3):363–369.