2006

Effects of Using Visual Design Principles to Group Response Options in Web Surveys

Jolene D. Smyth
Washington State University, jsmyth2@unl.edu

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

Leah M. Christian
Washington State University

Michael J. Stern
Washington State University, stern-michael@norc.org

International Journal of Internet Science
2006, 1 (1), 6-16

IJIS.NET

Effects of Using Visual Design Principles to Group Response Options in Web Surveys

Jolene D. Smyth, Don A. Dillman, Leah M. Christian, Michael J. Stern

Social Economic Sciences Research Center, Washington State University

Abstract: In this paper, we show that in Web questionnaires verbal and visual languages can be used to create groups and subgroups of information, which influence how respondents process Web questionnaires. Following Schwarz (1996; and also Schwarz, Grayson, & Knäuper, 1998) we argue that respondents act as cooperative communicators who use formal features of the questionnaire to help guide them through the survey conversation. Using data from three Web surveys of random samples of Washington State University undergraduates, we found that when response options are placed in close graphical proximity to each other and separated from other options, respondents perceive visual subgroups of the categories, increasing the likelihood that they select an answer from each subgroup. We also found that graphical proximity creates subgroups with and without the use of category heading to describe the subgroups and that the addition of a verbal instruction to "please select the best answer" encouraged respondents to select one answer from each subgroup instead of overriding the effects of proximity. In addition, the effects of grouping were consistent when the subgroups were positioned either vertically or horizontally in relation to each other. Lastly, we found that the effects of visual grouping are consistent across both opinion- and behavior/fact-based questions, although the effects appear to be greater on opinion-based questions. Our findings contribute to the increasing evidence that both verbal and visual languages influence how respondents process and respond to surveys. Because respondents interpret the verbal and graphical features of survey questionnaires as relevant to answering the survey, inadvertent or stylistic design changes can influence how respondents process and respond to survey questions.

Keywords: Visual grouping, graphical proximity, questionnaire design, survey

Introduction

Each question on a self-administered questionnaire can be viewed as a group of visually presented content that consists of the query, any instructions, response choices (unless open-ended), and spaces where answers are to be marked. Survey researchers have proposed that questionnaires be designed in a manner that clearly shows each question as a distinct group and that each question group be presented in a way that the respondent can infer the order in which questions and their sub-parts are to be processed and answered (Dillman, 2000). Thus, the questionnaire can be viewed as a sequence of information that is divided into interconnected groupings and sub-groupings of question content. However, relatively little research has been done to conceptualize and test

Revision of paper presented at the 2004 Annual Meeting of the American Association for Public Opinion Research, Phoenix, AZ, May 13, 2004. The research reported here was funded in part under Cooperative Agreement 43-3AEU-1-80055 between the SESRC and the United States Department of Agriculture National Agricultural Statistics Service with support from the National Science Foundation Division of Sciences Resource Statistics. We appreciate the encouragement of NSF-SRS staff to undertake this investigation.
Direct all correspondence to Jolene Smyth, Department of Sociology and the Social Economic Sciences Research Center (SESRC) at Washington State University, Pullman, Washington 99164-4014, USA. E-mail: jsmyth@wsu.edu.

how the grouping of information is communicated and how changes in grouping affect measurement or other response behaviors.

Research on how human beings process visual information has shown that people are more likely to perceive particular configurations of stimuli as one group. Changes in spacing, color, brightness, size, orientation, figure-ground, and/or common fate, can each influence whether stimuli appear to subjects as groups or as an undifferentiated series of visual stimuli (Palmer, 1999, p. 257-261). Experimental research has also shown that respondent comprehension of questions, instructions and answers can be significantly influenced by graphical manipulations for both paper (Christian & Dillman, 2004) and Web (Christian, 2003) questionnaires. Some of these effects may result from how the graphical language establishes grouping, thus affecting how respondents process the survey and its questions.

Our purpose in this paper is to report results from a series of experimental manipulations designed to examine whether alternative visual groupings of response options influences respondent answers to survey questions in Web surveys. Specifically, we examine whether or not the use of headings and spacing to form visually distinctive subgroups of response options (as shown in Figure 1) affects responses. The question format subjected to experimental testing was identified as presenting significant difficulties to pretest respondents during the evaluation of a Web prototype of the NSF sponsored National Survey of Earned Doctorates (Altheimer & Dillman, 2002). This research was undertaken to identify possible solutions to those problems as well as to contribute to our understanding of how visual design influences respondent answers through grouping processes.

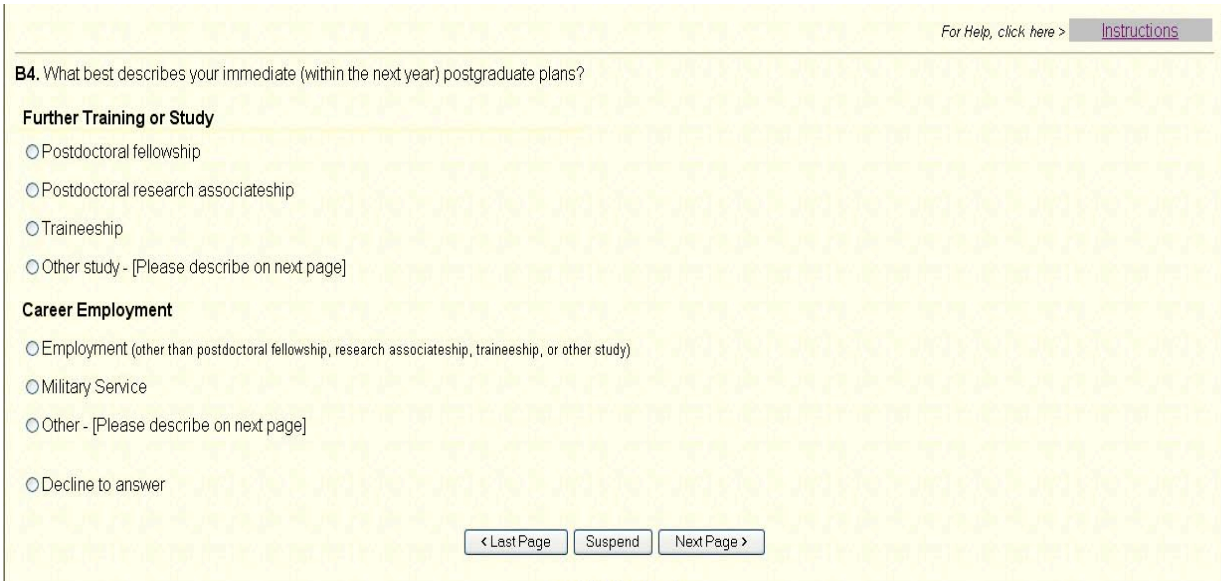


Figure 1. Screenshot of Survey of Earned Doctorates Web Pilot Question.

Theoretical Background

Adherence to the Rules of Communication

Respondents to surveys follow rules or maxims of communication (Schwarz, 1996). In other words, they approach the survey instrument as if they are having a conversation, with the instrument representing the researcher's contribution to the conversation. Within this context Schwarz highlights how apparently formal features (e.g. graphical layout) of the questionnaire from the researcher's perspective are important in the answering process because they communicate to the respondents what is expected of them (Schwarz, Grayson, & Knäuper, 1998). The importance of these formal features is magnified when the respondent is unsure what is being asked of them or how they are expected to answer. In these situations respondents are more likely to take their cues from design features of the questionnaire (Schwarz, 1996). Within the framework of respondents as cooperative communicators grouping is highlighted as an important formal feature of questionnaire construction

that communicates expectations to the respondent. Grouping can help respondents understand the intent of the question, response options, and answer spaces and thus reduce their likelihood of committing errors.

For example Schwarz and Hippler (as cited in Schwarz, 1996) found that respondents gave different answers to two questions (asking about marital and general satisfaction) when the questions were grouped together by placing them within a single box, as opposed to when they were presented separately in two boxes. Specifically, respondents' level of general satisfaction and marital satisfaction were less correlated when both questions were presented together in a box (as a group) than when they were presented separately (in two boxes). This finding indicates that the grouping of the questions affected how respondents interpreted them and subsequently what responses they gave.

Couper, Traugot, and Lamias (2001) report similar findings from an experiment in which they presented items on separate screens compared to several items on one screen in a Web survey. They consistently found that when multiple items were placed on one screen they were correlated more highly than when those same items were presented on separate screens; however, none of their comparisons reached significance. In a similar experiment Tourangeau, Couper, and Conrad (2004) found that eight related agree/disagree items pertaining to diet were more highly intercorrelated when they were presented on one screen as opposed to being presented as groups of four items on two screens or one item per screen. They interpret their results in light of the "near means related" heuristic discussed below.

Visual Processing as a Basis for Grouping

When respondents first look at a questionnaire they use preattentive processing (Neisser, 1967) to quickly take in the whole scene and make sense of the information presented (Jenkins & Dillman, 1997). At this broad level of processing all objects in the field of vision are competing for the respondent's attention (Neisser, 1967). It is during this stage that certain features of the questionnaire (e.g., the number one or a bold sign saying "start here") are likely to capture respondent attention. Alternatively, attentive processing involves respondents choosing a part of the questionnaire to focus on and then shifting their attention to another part, moving through the available information until the survey is completed. According to Neisser (1967), these attentive acts are "carried out in the context of the more global properties already established at the preattentive level" (p. 90). Thus, survey designers can use grouping techniques to help direct the respondent at both the preattentive and attentive processing levels. For example, section headings and boxes encompassing questions are sometimes provided to help respondents group questions at the preattentive processing level (e.g. Dillman, 2000, p. 397), whereas the content of individual items—the query, any instructions, and answer choices—are usually grouped in a consistent format within the question for consideration at the attentive stage of processing. Respondents use these groupings and sub-groupings as tools to navigate through the questionnaire (Dillman, 2000).

The Use of Visual Language to Achieve Grouping

A number of different methods are available for survey designers to use in their efforts to influence how respondents comprehend questionnaires and the individual items they contain. The most obvious is to manipulate the verbal language, or words, used to communicate with the respondent. However, research has shown that manipulating verbal language is only one way to convey information as respondents also rely on nonverbal languages to determine meaning (Redline & Dillman, 2002). Nonverbal languages include graphical language (font size, color, brightness, spacing, etc), numerical language (the use of numbers to suggest order), and/or symbolic language (i.e. the use of culturally prescribed symbols such as arrows to direct movement through questions). Nonverbal languages are used in conjunction with verbal language to communicate certain meanings to the respondent such as where to start, where to proceed next and how to process a specific question.

In addition to its direct effects, graphical language also serves as the conduit through which the other languages are transmitted. In other words, graphical manipulations (changes in the size, shape, location, spatial arrangement, color, brightness, contrast, and figure/ground composition) can influence the way verbal, numeric, or symbolic languages are perceived and interpreted (Jenkins & Dillman, 1997). Thus, they play a crucial role in guiding respondents through the answering process and can be used effectively to create desired groupings and sub-groupings.

Three pattern recognition concepts from Gestalt psychology, described by Palmer (1999) and applied to survey design by Jenkins and Dillman (1997), are relevant here. The first is the *principle of proximity,* which states that we tend to group things together based on the distance of their spatial separation. In other words, we will see items that are close to one another as a group and items that are distant as separate. The second is the *principle of similarity* whereby respondents are more likely to mentally group images that appear alike. Similarity can be established through several means such as font, shape and color (Palmer, 1999). The third, the *principle of*

pragnanz states that figures with simplicity, regularity, and symmetry are easier to perceive and remember (see Palmer, 1999, p. 257-261).

Tourangeau et al. (2004) expand upon these pattern recognition concepts, arguing that respondents make conceptual inferences about response options based on the visual presentation of those options. They describe five interpretive heuristics that respondents use in processing visual stimulus and argue that the use of these simple heuristics can oftentimes lead respondents to misinterpret survey questions. One such heuristic, based on the Gestalt psychology principle of proximity, states that respondents will see items as more conceptually related when they are located in closer proximity to each other on a page or screen ("near means related") (Tourangeau et al., 2004, p. 370). Another, based on the principle of similarity, states that respondents will assume items are more conceptually related to one another when they appear visually similar to one another ("like means close") (Tourangeau et al., 2004, p. 370).

These Gestalt principles and the heuristics derived from them suggest several ways to group information in questionnaires. One way is through the use of space on the instrument. For example, using greater space between questions than between the stem of a query and the accompanying response options creates grouping based on proximity and clarifies the boundaries between questions. Sub-grouping of response options with a question can also be established through the use of proximity. While differences in the proximity of the items can create the impression of grouping, that impression may be reinforced by the assumption of conceptual similarities and differences that it creates through the heuristic "near means related." Another way of establishing grouping is through graphical similarity (or contrast). Through graphical manipulations, one may make all question stems bold or in a larger font, while keeping response options smaller, thus establishing subgrouping. Similarity can also be achieved through orientation. For example, items that are oriented horizontally on the page might appear to belong to one group while items that are oriented vertically appear to belong to another (Palmer, 1999, p. 258).

However, without careful attention to detail, nonverbal language, including graphical manipulations, can work in opposition to verbal language cues and lead the respondent to complete the survey in ways unintended by the researcher, thus introducing response errors. One example is when respondents do not understand and thus fail to follow branching instructions, answering the questions in an incorrect order or failing to respond to all applicable questions.

Research on branching instructions has shown that a significant number of respondents fail to follow the instructions when they are placed on the opposite side of the response option from the answer box. Certain branching errors were reduced by one-third to one-half when the visual prominence of the branching instructions was increased (Redline & Dillman, 2002; Redline, Dillman, Dajani, & Scaggs, 2003). Increasing visual prominence was accomplished by either (1) placing the answer box on the right of the response option with the skip direction only a few spaces to the right of the box (within the respondent's foveal view) or (2) by retaining the original location while making the branching instructions larger and darker, adding arrows, and adding an additional word instruction at the beginning of the next question to indicate that only certain respondents should answer it (e.g. "If no to the previous question"). In the first instance the skip directions were grouped with the answer space so respondents would have them within their foveal view when providing an answer and immediately before moving to the next question. In the second, the size and contrast of the skip instructions was increased to draw attention to their existence (and importance) as part of the grouping and subgrouping of the questions. In both cases, increasing the visual prominence of the instructions increased the likelihood respondents saw it, thereby increasing compliance with the branching instructions.

Subgrouping Response Options in the NSF Earned Doctorate Survey Web Prototype

In 2001 a longstanding paper questionnaire survey conducted by the National Science Foundation, The Survey of Earned Doctorates, was converted to a Web survey format. Attempts are made each year to get every person who finishes a doctoral degree at a U.S. university to complete this survey. Because it was anticipated that this survey would be conducted for a number of years as a mixed-mode survey, and because of the importance of maintaining trend lines, it was deemed important that attention be given to achieving mode comparability. Therefore, cognitive interviews were conducted to evaluate a pilot version of the Web survey (Altheimer & Dillman, 2002). Respondents to the NSF Earned Doctorate Survey were asked to indicate their immediate postgraduate plans by choosing one answer from among seven choices. In the original mail questionnaire these categories were visually separated with four of them being placed under a general heading of "further training or study" and the remaining three categories being placed under the heading of "career employment." This format was retained for the pilot Web survey (see Figure 1) where it was observed that some respondents tried to provide answers under each heading. However, the use of radio buttons (where a respondent may choose only one answer) for this question meant that some respondents unintentionally erased their answer to the choice they

made under the first general category when they chose another under the second general category without being aware of having made the change (Altheimer & Dillman, 2002).

The cognitive interviews revealed the need to better understand how the grouping of information is communicated to respondents and whether some formats are more effective than others in getting respondents to answer questions accurately. In the subsequent surveys we adapted the Earned Doctorate Survey prototype question so that it could be evaluated in Web surveys conducted among undergraduate students. We report here results from a series of three Web surveys, each with up to four comparisons. The experiments contained within these three Web surveys allow us to address the following research questions:

1. Does vertical grouping of response options yield different answers than no grouping?
2. Can the effects of vertical grouping of response options be over-ridden by an instruction encouraging only one answer?
3. Are the effects of grouping similar on different types of questions (opinion-based vs. fact/behavior-based)?
4. Does horizontal grouping of response options have the same effect as vertical grouping?
5. What is the relative importance of headings and spacing in establishing grouping?

Procedures

Over a year and a half period three Web survey experiments were conducted. Survey dates, sample sizes, and response rates can be found in Table 1 along with other descriptive information about each survey. All three surveys were designed to assess the undergraduate experience at Washington State University (WSU) and in the surrounding community, although there was some variation in question topic across the surveys. The respondents for each survey consisted of a randomly selected sample of students registered for classes at the Pullman campus of WSU during the semester the respective survey was fielded.

Table 1
Design and Implementation Details for the Three Surveys of Random Samples of Washington State University Students

Survey Date	Experimental Versions	Number of Questions	Sample Size	Completed Responses	Response Rate
Spring 2003	4	21	3,004	1,591	53%
Fall 2003	4	25	3,045	1,705	56%
Fall 2004	3	25	1,800	1,082	60%

Note: Only three versions from the Spring 2003 survey and two experiments from the Fall 2003 survey are reported in this paper as the others are not relevant to the current topic. All three versions from the Fall 2004 experiment are reported.

All three surveys used a similar design and for each careful programming was undertaken to ensure that the questions would appear similarly across different computer set-ups such as browsers and hardware configurations. Specifically, all screens were constructed with HTML tables using proportional widths in order to maintain the visual aspect of the screen regardless of individual user window sizes. In addition, font size and style were automatically adjusted using Cascading Style Sheets to accommodate differing user screen resolutions.

For each survey the sampled students were contacted by postal mail and asked to go to the Web and complete the questionnaire. With the initial contact letter respondents received a two-dollar incentive. E-Mail follow-ups to provide a hotlink and additional postal contacts were made. To gain entry to the survey instrument students were required to enter their own personal access code, which was provided to them in the first contact letter and all subsequent contacts. Access codes were used to ensure that only individuals in the sample could participate and that they could only complete the survey once. Respondents were randomly assigned one version of the questionnaire to complete.

In all three surveys the questions used in this paper were designed using HTML boxes for answer spaces. This design feature was intended to allow respondents to interpret the questions (including researcher expectations) and answer them without constraints. In addition, it allows us to analyze how respondents processed the questions by looking at which option(s) they marked as well as how many they marked.

The NSF question requesting immediate postgraduate plans was operationalized in our first survey by asking students to indicate which of six options described the benefits of the student recreation center. Four slightly

varying versions of this item were constructed. The three that are relevant to the current research can be seen in Figure 2. These three treatments allow us to examine research questions number one and two:[1]

Version 1: An underlined heading was placed above each of two subsets of three response options arranged in a vertical format. One was labeled "Health Benefits" and the other was labeled "Academic Benefits." In addition, there was one line of space left blank between the two subsets. This version emulates the original NSF question format.

Version 2: The same question and groupings were presented, but a word instruction stating "Please select the best answer" was added to attempt to override any grouping effect.

Version 3: All six choices were placed in a single vertical line with no indication of subgrouping (no headings and no additional spacing between groups).

The experiment included in the second survey provided a different question topic (see Figure 3). Respondents were asked, "Have you received financial support from each of the following sources while attending WSU?" The treatments in this survey allow us to examine the final research question about the relative importance of headings and spacing in establishing grouping. The two treatments relevant to this paper were as follows:[2]

Version 1: Headings reading "Financial Aid" and "Other Sources" were placed above two double banked subgroups of response options.

Version 2: The double banking was retained, but the headings were removed.

The third and final survey included three experimental treatments and used the same question topic, albeit different question wording, as the second survey (see Figure 4). Respondents were asked, "What best describes the financial support you have received while attending WSU?" This final set of experimental treatments allows us to explore both the third and fourth research questions. The treatments were as follows:

Version 1: Headings reading "Financial Aid" and "Other Sources" were placed above two double banked subgroups of response options.

Version 2: The same headings were placed above two vertically aligned subgroups of response options. One line of space was left blank between the subgroups.

Version 3: The subgrouping (via headings and spacing) was removed, leaving one vertically aligned column of response options.

Our general hypothesis is that when response choices are presented to respondents as separate groups, indicated by the use of words and/or graphic separation, the selection of answers from both groups will become more likely, thus altering response distributions.

Results

Does Vertical Grouping of Response Options Yield Different Answers than No Grouping?

Results from the first set of experiments make it clear that the use of the headings and spacing (Figure 2) influenced answer choices, as expected (see Table 2). Respondents to the vertically grouped version (version 1) not only chose more response categories than the respondents to the version with no subgrouping, but were more likely to choose at least one answer from each of the sub-groupings (70.2% vs. 40.9%). These results clearly suggest that the use of headings and accompanying separation to establish two visual sub-groupings influence responses.

[1] The fourth treatment is only tangentially related to the research questions posed here and did not yield significant results; therefore, it is not reported here in detail.

[2] Two additional experimental treatments were designed for this question (for a total of four experimental treatments), but are not relevant to the purpose of this paper.

Can the Effects of Vertical Grouping of Response Options Be Over-ridden by an Instruction Encouraging Only One Answer?

The addition of the instruction to "Please select the best answer" (see Figure 2) in an attempt to override the visual grouping effect had a mixed effect. Respondents to the version with the instruction (version 2) were only slightly less likely to choose an answer in each subgrouping than were those who completed the version without the instruction (version 1), chi square (1, N = 873) = 1.63, p = .20. However, the mean number of options checked was significantly less for those who had the additional instruction than for any of the other treatments. These results indicate that respondents may have been drawing information from two sources, the instruction to select the best answer (the lower mean) and the sub-groupings (the greater tendency to select from both groupings). These two sources of information, working in concert, seemed to have communicated to the respondents that they were expected to "select the best answer" *from each subgroup*. Thus, the evidence presented here indicates that this particular instruction did not over-ride the grouping effects.

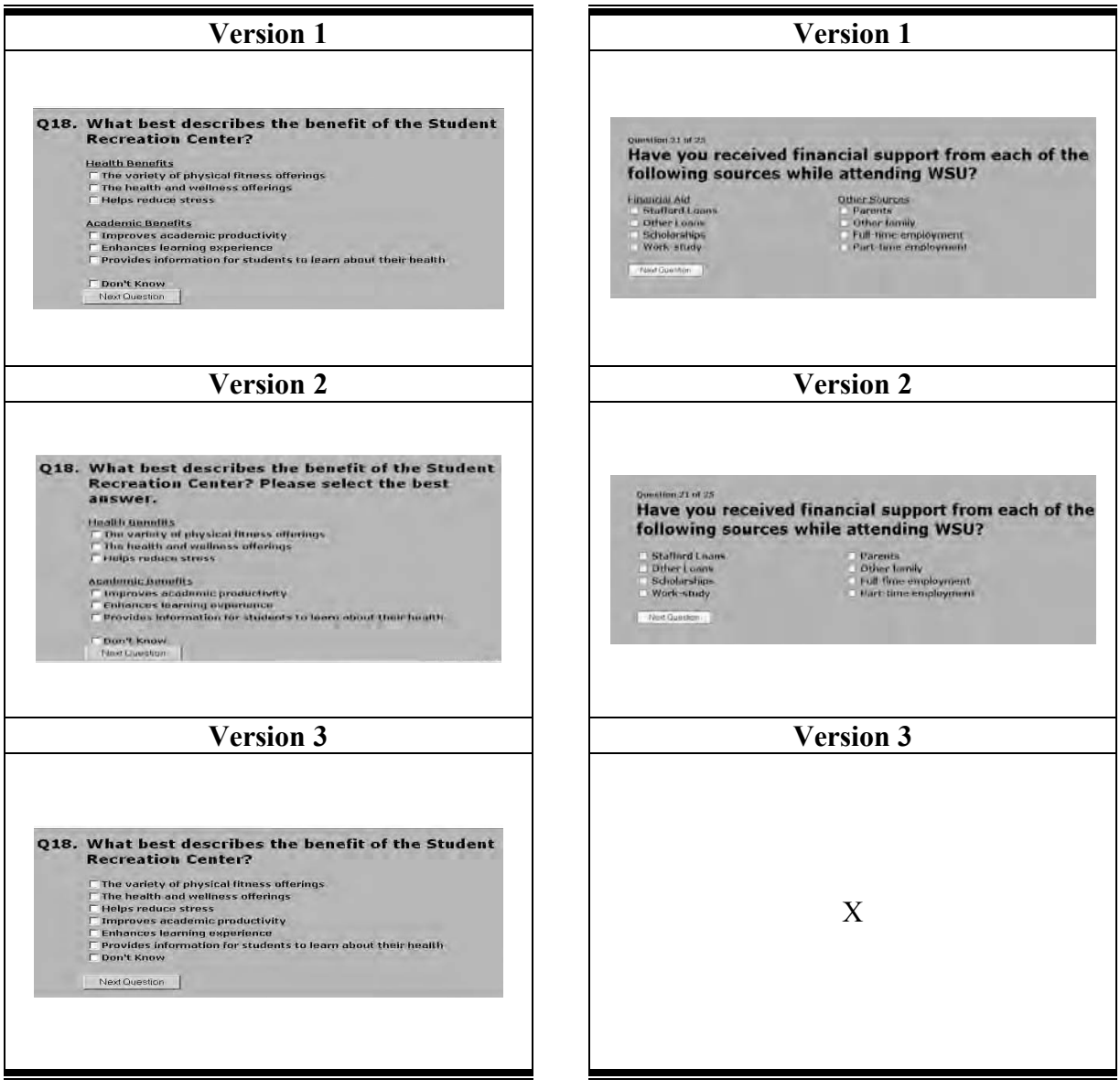


Figure 2. Experimental Treatments for Spring 2003 Survey.

Figure 3. Experimental Treatments for Fall 2003 Survey.

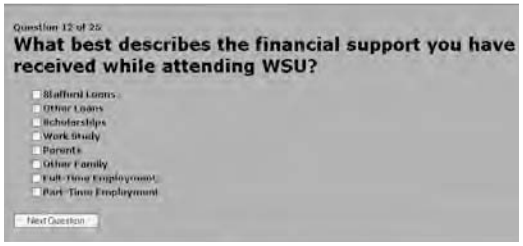
Version 1	Version 2	Version 3
		

Figure 4. Experimental Treatments for Fall 2004 Survey.

Table 2
The Effects of Vertical Grouping
Q18: What best describes the benefit of the Student Recreation Center (Survey #1)

	Version 1 Vertical	Version 2 Vertical/ Instruction	Version 3 No group	Chi-Square Significance Tests (V1 vs. V2)		(V1 vs. V3)		(V2 vs. V3)	
				X^2	*p*	X^2	*p*	X^2	*p*
Marked in 1st Half	99.7	100.0	99.1	1.01	.315	1.45	.228	3.70	.054
Marked in 2nd Half	70.4	66.0	41.8	1.84	.175	59.72	.000*	42.16	.000*
Marked in Both	70.2	66.0	40.9	1.63	.202	62.36	.000*	45.39	.000*
				One-Sided t-Tests					
				t	*p*	*t*	*p*	*t*	*p*
Mean # Checked in 1st Half	1.9	1.5	2.0	5.99	.000*	-1.36	.913	-6.91	.000*
Mean # Checked in 2nd Half	1.1	0.9	0.8	3.51	.000*	4.89	.000*	1.98	.976
Mean # Checked	3.0	2.4	2.7	5.41	.000*	2.38	.009*	-2.39	.009*
N	**435**	**438**	**367**						

Note: All Chi-square tests with df=1.

Are the Effects of Grouping Similar on Different Types of Questions?

To answer this question we turn to the results of the third survey which can be found in Table 3. While the question about the benefits of the student recreation center found in the first survey (Figure 2), is a very opinion-based question, the question in the third survey about financial support for school is much more behavior/fact-based (see Figure 4). Comparing the effects of vertical grouping across these two quite different question types (Tables 2 & 3) shows that while both are prone to grouping effects, the grouping effect in the behavior/fact-based question is substantially less than the grouping effect in the opinion-based question. Vertical grouping in the opinion-based question (Table 2) resulted in 70.2 percent of respondents marking answers in both groups while only 40.9 percent did so in the ungrouped version, for a difference of 29.3 percent. In contrast, for the behavior-fact based question, 64.9 percent of respondents to the grouped version marked answers in both the top and bottom halves of the options while 56.5 percent did so for the non-grouped version, for a difference of only 8.4 percent. Thus, the effects of grouping response options does differ for different types of questions with behavior-fact-based questions being less prone to grouping effects and opinion-based questions being more prone to grouping effects.

Table 3
The Effects of Vertical and Horizontal Grouping

Q12: What best describes the financial support you have received while attending WSU? (Survey #3)

	VERSION 1 Horizontal	VERSION 2 Vertical	VERSION 3 No Group	CHI-SQUARE SIGNIFICANCE TESTS (V1 vs. V2) χ^2	(V1 vs. V2) *p*	(V1 vs. V3) χ^2	(V1 vs. V3) *p*	(V2 vs. V3) χ^2	(V2 vs. V3) *p*
Marked in 1st Half	76.4	75.2	71.9	0.13	.716	1.84	.175	1.04	.309
Marked in 2nd Half	85.8	86.3	81.8	0.04	.838	2.00	.157	2.72	.099
Marked in Both	64.7	64.9	56.5	0.00	.947	4.88	.027*	5.37	.020*
				T-TESTS					
				2-SIDED *t*	*p*	**1-SIDED** *t*	*p*	**1-SIDED** *t*	*p*
Mean # Checked in 1st Half	1.4	1.4	1.4	0.19	.853	0.11	.456	-0.07	.528
Mean # Checked in 2nd Half	1.2	1.3	1.2	-1.52	.129	0.36	.361	1.82	.035*
Mean # Checked	2.6	2.7	2.6	-0.70	.479	0.31	.380	1.01	.157
n	**351**	**379**	**352**						

Note: All Chi-square tests with df=1.

Does Horizontal Grouping of Response Options Have the Same Effect as Vertical Grouping?

Table 3 shows that there is no significant difference in how respondents answered when the sub-groupings were arranged horizontally as opposed to when they were arranged vertically (Figure 4). While the subgrouping of response options has a clear effect, whether these subgroups appear horizontally aligned or vertically aligned in relation to one another appears not to matter.

What is the Relative Importance of Headings and Spacing in Establishing Grouping?

As shown in Table 4, the respondents who were presented with a grouped version of the question including headings did not respond differently from those presented with a grouped version without headings (Figure 3). These findings suggest that grouping can be established inadvertently through simple spacing manipulations (i.e., double banking response options to save space on the survey instrument); the more conscious use of headings is not necessary to establish grouping. This last finding, however, should be considered tenuous and should be subjected to more examination as an alternative explanation is also possible. The difference in question wording in this survey may explain the lack of difference between the two treatments. The check-all-that-apply wording of the question (i.e., each of the following), may have encouraged respondents from both versions to give careful attention to all of the response options, regardless of how the options were presented. In addition, the relative importance of headings and spacing has yet to be examined in the vertical orientation.

Table 4
The Effects of Category Headings on Horizontal Grouping

Q21: Have you received financial support from each of the following while attending WSU? (Survey #2)

	VERSION 1	VERSION 2	CHI-SQUARE SIGNIFICANCE TESTS	
	Headings	**No Headings**	χ^2	*p*
Marked in 1st Half	73.5	78.5	2.57	.109
Marked in 2nd Half	91.9	88.3	1.63	.201
Marked in Both	64.6	66.8	0.38	.538
			TWO-SIDED T-TESTS	
			t	*p*
Mean # Checked in 1st Half	1.4	1.6	-1.72	.087
Mean # Checked in 2nd Half	1.4	1.5	-0.58	.559
Mean # Checked	2.9	3.1	-1.86	.064
n	**393**	**446**		

Note: All Chi-square tests with df=1.

Discussion and Conclusions

Our research demonstrates that the use of visual design principles to create groups and subgroups of information influences how respondents answer Web survey questions. The findings appear to support Schwarz's (1996) argument that respondents gain information about the researcher's expectations from the visual design and layout of the survey instrument and then use that information when answering the questions. When answer categories are separated using graphical language and applying the Gestalt principle of proximity, respondents are more likely to perceive the response options as belonging to separate sub-groups. That perception then guides their response strategy, making them more likely to choose at least one response option from each sub-group. This sub-grouping may be further reinforced if respondents use the heuristic "near means related" to draw inferences about the degree of conceptual similarity or difference among response options based on their proximity to one another.

We also confirm findings from cognitive interviews conducted using the Survey of Earned Doctorates Web prototype where it was originally found that respondents tried to select one option from each subgroup. However, unlike the experiments reported here, in the Survey of Earned Doctorates respondents reported their answer using radio buttons instead of html boxes. This meant that when respondents selected a second category it replaced the selection of the first, creating an additional source of respondent frustration. Thus, our findings suggest that what started out as a helpful design element, the graphical subgrouping of related response options under descriptive headings in the Survey of Earned Doctorates, may have inadvertently created response error and increased respondent frustration in both the paper version of that survey and in subsequent Web versions.

In the Survey of Earned Doctorates prototype, verbal category headings and graphical proximity were used to create subgroups of response categories. We tested these effects separately and found that graphical separation of response options was more influential in establishing grouping than was the use of verbal category headings to define the categories. In a separate test we found that respondents continued to select at least one answer from each subgroup when they were instructed to "select the best answer"; even with the instruction, the effects of the visual subgrouping of response categories still occurred. It appears as if respondents interpreted the instruction to select the best answer in conjunction with the perceived expectations established by the visual grouping. The combined verbal and visual message seemed to be, "please select the best answer from each subgroup."

While the response subgroups were placed vertically, one on top of the other, in the Survey of Earned Doctorates (both the original paper version and the Web prototype), we experimentally tested the effects of placing the subgroups both horizontally (one next to the other) and vertically. We found that the influence of visually subgrouping the response options occurred regardless of the location of the subgroups in relation to each other. In addition, we replicated the Survey of Earned Doctorates question type by testing the effects of grouping using a behavior/fact-based question which asked students about their sources of financial support while attending WSU. However, we also tested the effects of grouping and subgrouping using a question requesting students' opinions.

The effects of grouping were much larger for the opinion- than the behavioral/fact-based question, but remained significant across both types of questions. Both of these findings suggest that the effects of visual grouping and subgrouping seem robust in that they occur regardless of the graphical orientation of the subgroups and they influence the response patterns of various types of survey questions.

The practical implication of these findings is that survey designers must take the grouping and subgrouping of questions and response options into consideration when constructing surveys. Inadvertent graphical changes resulting in the grouping of information can have significant effects on respondent processing and answering. More generally, both verbal and visual languages must be considered throughout the design process as it is becoming increasingly clear that respondents use both types of languages when processing and responding to surveys.

References

Altheimer, I., & Dillman, D. A. (2002). *Results from cognitive interviews of nsf earned doctorate Web survey*. (Social & Economic Sciences Research Center Technical Report #02-30). Pullman, WA.

Christian, L. M. (2003). *The influence of visual layout on scalar questions in Web surveys*. Washington State University, Pullman, WA: Unpublished master's thesis.

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,* 58-81.

Couper, M. P., Traugott, M. W., & Lamias, M. J. (2001). Web survey design and administration. *Public Opinion Quarterly, 65,* 230-254.

Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method* (2nd ed.). New York, NY: Wiley.

Jenkins, C. R., & Dillman, D. A. (1997). Towards a theory of self-administered questionnaire design. In L. E. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 165-198). New York, NY: Wiley.

Neisser, U. (1967). *Cognitive psychology*. New York, NY: Meredith Publishing Company.

Palmer, S. E. (1999). *Vision science photons to phenomenology*. Cambridge, MA: MIT Press.

Redline, C., & Dillman D. A. (2002). The influence of alternative visual designs on respondents' performance with branching instructions in self-administered questionnaires. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.) *Survey nonresponse* (pp. 179-196). New York, NY: Wiley.

Redline, C., Dillman, D. A., Aref, D. N., & Scaggs, M. A. (2003). Improving navigational performance in u.s. census 2000 by altering the visually administered languages of branching instructions. *Journal of Official Statistics. 19,* 403-419.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Mahwah, New Jersey: Lawrence Erlbaum.

Schwarz, N., Grayson, E. G., & Knäuper, B. (1998). Formal measures of rating scales and the interpretation of question meaning. *International Journal of Public Opinion Research, 10,* 177-183.

Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position, and order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly, 68,* 368-393.