

How Visual Grouping Influences Answers to Internet Surveys¹

By

Jolene D. Smyth, Don A. Dillman, Leah Melani Christian and Michael J. Stern
Washington State University

ABSTRACT

A questionnaire can be viewed as a sequence of information divided into interconnected groupings and sub-groupings of question content that communicate important information about expectations to the respondent. However, relatively little research has tested how grouping information is communicated, and how changes in grouping might affect measurement or other response behaviors. Our purpose in this paper is to report results from a series of experimental manipulations for three types of questions to examine whether visual grouping influences respondent answers to survey questions. They include: (1) the effects of providing headings for subgrouping of response options, (2) alternative procedures for presenting long lists of answer choices to respondents, and (3) the differential location and use of symbols, spacing and words to define requests for dates. The three question formats subjected to experimental testing were identified as presenting significant difficulties to pretest respondents during the evaluation of a web prototype of the NSF sponsored National Survey of Earned Doctorates. This research was undertaken to identify possible solutions to those problems as well as contribute to our understanding of how visual design decisions influence respondent answers through grouping processes. Results regarding heading use indicate that the use of headings when response options are arranged in a vertical listing greatly influences answer choices by encouraging respondents to provide at least one response under each heading instead of simply providing one response overall. With respect to how best to group a long list of answer choices, the findings indicate that allowing the respondents to type in their answer is the most efficient method in terms of response time while the two-stage approach of listing the specific options (student majors) under more general headings (general area) is least efficient. Finally, when asking respondents to report a date in a specific format, results indicate that grouping the symbols that communicate the desired format (MM YYYY) with the appropriate answer boxes (by locating them directly below their respective boxes as opposed to off to the right of both boxes) results in higher use of the desired format. Taken together, these results indicate that grouping and subgrouping of information has significant effects on both the efficiency with which respondents are able to answer questions and the way they answer them. In addition, they suggest the need for additional research exploring the effects of specific visual design features to establish grouping.

¹ Revision of paper presented at the 2004 Annual Meeting of the American Association for Public Opinion Research, Phoenix, AZ, May 13, 2004. Jolene Smyth, Leah Christian and Michael Stern are graduate assistants in the Department of Sociology and the Social Economic Sciences Research Center (SESRC), and Don A. Dillman is Regents Professor in the Department of Sociology and Department of Community and Rural Sociology, and Deputy Director of the SESRC, all at Washington State University, Pullman, Washington 99164-4014. 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. This is SESRC Technical Report # 04-023

TABLE OF CONTENTS

ABSTRACT	i
INTRODUCTION	1
BACKGROUND	2
Visual Processing as a Basis for Grouping	2
The Use of Visual Language to Achieve Grouping	2
Adherence to the Rules of Communication	3
Web Surveys	4
Problem Question Formats in the Earned Doctorate Survey Web Prototype	5
PROCEDURES	6
THE INFLUENCE OF CATEGORY HEADINGS ON GROUPING	7
Results for Category Heading Effects	9
SELECTING A FIELD OF STUDY	11
Results for Selecting a Field of Study	13
REQUESTS FOR MONTH AND YEAR OF AN ACTIVITY	15
Results for Month and Year of an Activity	17
DISCUSSION AND CONCLUSIONS	20
APPENDIX A: EXPERIMENTAL TREATMENTS	23
REFERENCES	31

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 grouping distinct from all other questions 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 how grouping information is communicated and whether 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. For example, 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 1999:257-261). Experimental research has also shown that respondent comprehension of questions, instructions and answers can be significantly influenced by graphical manipulations of this nature for both paper (Christian and Dillman, 2004) and web (Christian, 2003) questionnaires. Some of these effects may result from how information is grouped for processing.

Web surveys are of particular interest because they provide far more potential than paper questionnaires for manipulating information in ways that visually change how it is interpreted by respondents. On the Internet it is easy and inexpensive to manipulate the visual aspects of pages in various ways. In addition, new question formats (e.g. drop down menus) and prescribed uses of symbols (radio buttons vs. html boxes), produce visual variations that may contribute to the occurrence of unintended grouping effects which may in turn affect answers.

Our purpose in this paper is to report results from a series of experimental manipulations for three types of questions to examine whether alternative visual grouping and related changes influence respondent answers to survey questions. They include: 1) the effects of providing headings for subgrouping of response options, 2) alternative procedures for presenting long lists of answer choices to respondents, and 3) the differential location and use of symbols, spacing and words to define requests for dates. The three question formats subjected to experimental testing were identified as presenting significant difficulties to pretest respondents during the evaluation of a web prototype of the NSF sponsored National Survey of Earned Doctorates (Alzheimer and Dillman, 2002). This research was undertaken to identify possible solutions to those problems as well as contribute to our understanding of how visual design decisions influence respondent answers through grouping processes.

BACKGROUND

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 and 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. 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 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: 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 and 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 affects, 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 and Dillman 1997). Thus, they play a crucial role in guiding respondents through the answering process. However, without careful attention to detail nonverbal language 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 incorrect orders or failing to respond to applicable questions.

Nonverbal language can be used effectively to create desired groupings and sub-groupings to simplify the answering process. Three relevant pattern recognition concepts from Gestalt psychology have been described by Palmer (1999) and applied to survey design by Jenkins and Dillman (1997). 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: 257-261).

These Gestalt principles 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. Another way of establishing grouping is through 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: 258).

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 in 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 respondent what is expected of them. 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 than when the researcher's expectations are more clear (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. It 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 (1992) 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, the 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.

Similarly, 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 and Dillman 2002, Redline, Dillman, Dajani and 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.

Web Surveys

As web surveys grow in popularity and frequency it is important to understand how grouping and subgrouping can be conveyed in them as well as some of the additional abilities and challenges that web survey designers confront. To the extent that they rely on visual presentation of material, web surveys may be very much like paper surveys—correct completion of them depends on the graphical presentation of verbal as well as nonverbal languages. Experiments have shown that varying the layout of multiple answer questions (Dillman, Smyth, Christian and Stern, 2003) and ordinal scale questions (Christian, 2003) produces similar differences in paper and web surveys. However, web surveys also differ from paper surveys in a number of important ways.

One fundamental difference is the ability to affect grouping via the number of questions included on any one screen (Best and Krueger, 2004). For example, whereas it is not plausible to construct a paper survey with one question per page to designate grouping, it is easy to construct a web survey in this manner, referred to as page-by-page construction. This format allows the designer to clearly define the components of each question. Another option is to construct the survey instrument using scrolling. With

scrolling numerous questions can be included on one screen and the respondent has to scroll up and down the page to see them (For an examination of the effects of page-by-page construction versus scrolling see Couper, Traugott, and Lamias 2001). This construction style requires the designer to use other methods to designate groups and subgroups.

Another difference has to do with orientation. Whereas paper surveys constructed with 8.5 by 11 inch sheet of paper are vertically oriented, the screens of computers tend to be oriented horizontally. This difference in orientation gives the survey designer a different type of space to work with and may encourage the designer to orient the material on the screen in a more horizontal manner. Changes in how information is grouped visually may result.

Finally, web surveys provide the opportunity to arrange response options in a number of different ways, some similar and some different from paper surveys. For example, the decision to arrange response options in columns as opposed to rows must be made for both web and paper surveys, but the ability to organize response options in drop down menus is unique to the internet and may affect responses (Couper, Tourangeau, and Conrad 2004). Similarly, on web surveys the format of the answer space can convey information to the respondent as certain kinds of answer spaces have been reserved for certain types of responses. For example, radio buttons have been reserved for single-answer questions while the use of HTML check boxes has been reserved for multiple-answer questions. However, respondents may not be aware of this convention and the fact that one can provide only one answer when radio buttons are provided while several answers can be marked when HTML boxes are provided.

The avoidance of errors is important for the sake of measurement, but it is especially important on web surveys where oftentimes respondents will be informed of their error and asked to correct it through an error message. Although error messages might help the respondent provide an acceptable answer, they may also promote a certain level of frustration that could result in respondents failing to complete the survey (Alzheimer and Dillman 2002). This is especially true if respondents receive several such messages and still cannot figure out what is expected of them. Finally, clear grouping can indirectly affect respondents' frustration levels by reducing the amount of time it takes to answer questions and thus shortening the perceived length and personal costs of completing the survey.

Problem Question Formats in the Earned Doctorate Survey Web Prototype

In 2001 a longstanding paper questionnaire survey conducted by the National Science Foundation, The Earned Doctorate Survey, 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. These revealed significant problems on three question formats. For all three questions, problems appeared to stem from grouping and subgrouping issues (Alzheimer and Dillman 2002).

Alternative versions of these questions are experimentally compared in this study to better understand how grouping processes may influence respondent answers. The first question included a set of response categories that appeared to be visually divided into two groups through the insertion of descriptive headings, but respondents were expected to choose only one answer. In the cognitive interviews respondents exhibited a tendency to try to choose at least one answer from each group despite the fact that the radio button format allowed only one answer to be selected.

The second question was a request for respondents to choose a “field of study” from a list of fields divided into categories and subcategories. On the web survey this display, taken directly from the paper questionnaire, required the use of two screens. Respondents had difficulty toggling between the two screens to find their appropriate field of study.

The third question was simply a request to report month and year that the respondent started graduate school. To emphasize that two digits were to be used to report month and four to report year, the symbol “MM YYYY” was placed after the year box. However, respondents made various errors including among others the use of letters for the month or trying to place two digits in each box.

Respondent difficulties to each of these pilot study questions led to error screens that produced additional respondent frustration (Alzheimer and Dillman 2002). The cognitive interviews revealed the need to better understand how the grouping of information is communicated to respondents and whether some versions are more effective than others in getting respondents to answer questions accurately. In the present study we adapted each of these questions to topics and formats that could be evaluated in web surveys conducted of undergraduate students. We report here results from a series of two web surveys, each with up to four comparisons. Examples of the adapted questions and formats can be seen in Figures 1 through 7.

PROCEDURES

Two web survey experiments were conducted, with results from the first experiment being used to guide the design of the second one. The first web survey, which consisted of 21 questions, was conducted in the spring of 2003. The survey was designed to assess the undergraduate experience at Washington State University (WSU). It used a page-by-page design and questions were presented in black text against a colored background with answer spaces appearing in white so as to provide contrast between the answer spaces and the background. Careful programming was undertaken to ensure that the questions would appear similarly across different computer set-ups such as browsers and hardware configurations. The sample consisted of 3,004 randomly selected WSU undergraduate

students who were registered for courses on the Pullman campus during the spring 2003 semester. Of the 3,004 students in the sample 1,591 completed the survey for a response rate of 53 percent.

The second web survey, consisting of 25 questions, was conducted in the fall of 2003. This survey was designed to assess students' experiences both on and off campus in the Pullman, Washington area. Similarly to the first survey, this one was designed using page-by-page construction with questions presented in black text against a colored background. Answer spaces appeared in white. Similar precautions were taken to ensure that the appearance of the questions would not be manipulated in important ways by different computer set-ups. The sample consisted of 3,045 randomly selected WSU undergraduate students who were registered for courses on the Pullman campus during the fall of 2003. The response rate for this survey was 56 percent or 1,705 students.

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 two 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. A random number generator was used to assign one of the versions to each respondent when they entered the survey.

Where appropriate chi square statistics are calculated to compare responses across versions. In other instances t-tests of differences in means are calculated. For both surveys client side paradata were collected (Heerwegh and Loosveldt 2002; Stern et al. 2004). Of specific value to the questions being explored in this paper is the time respondents spent on each question.

THE INFLUENCE OF CATEGORY HEADINGS ON GROUPING

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 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 where it was observed that some respondents tried to provide answers under each heading. However, the use of radio buttons 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. The grouping issue addressed by this question is whether spacing differences in combination with the use of words (expressed as headings) influence people's answer choices. Our general hypothesis is that when response choices are presented to respondents as separate groups they are more likely to

choose answers from both groups. 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 four treatments were as follows and can be seen in Figure 1 of Appendix A:

Version 1: An underlined heading was placed above each of two subsets of three response options. 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 sub-grouping (no headings and no additional spacing between groups).

Version 4: This treatment was the same as version one except that the headings were slightly indented so that the answer spaces appeared in an uninterrupted, vertical line.

The experiment included in the second survey provided a different question topic and somewhat different approaches to grouping, as well as efforts to override any effects. The topic of this operationalization was sources of financial support. The subgrouping of the eight response options included the presentation of half of the options in a second column (i.e. a double-banking format). In other words, the sub-groups were aligned with each other horizontally instead of vertically as in the first survey. The four treatments were as follows and can be seen in Figure 2 of Appendix A:

Version 1: Headings reading “Financial Aid” and “Other Sources” were placed above the two double banked sub-groups of response options.

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

Version 3: The response options were converted to a yes/no forced-choice format. The headings were included as they were in version one.

Version 4: The response options were converted to a yes/no forced-choice format, but the headings were not included.

We expect the headings to have less effect in the second survey; the groupings are less visually prominent and are overshadowed by the use of the two column format. In addition, it has been suggested that double-banking leads to a tendency for the second column not to be seen as it is often placed outside of the foveal view of a respondent.

The forced-choice format has been proposed as a procedure that encourages respondents to process items more thoroughly than when a check-all format is used (Rasinski et al. 1994; Dillman et al. 2003). This change in format is an attempt to measure the influence of headings under another condition—using a format that has been found to increase the number of response options selected for other questions.

Results for Category Heading Effects

Results from the first set of experiments make it clear that the use of the headings influenced answer choices, as expected (see Table 1). Respondents to version 1 (headings and separation) of the first survey not only chose more response categories than the respondents to version 3 (no headings or separation), 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.

Table 1: The Effects of Category Headings on Vertical Grouping

Q18: What best describes the benefit of the Student Recreation Center (Survey #1)

	VERSION 1	VERSION 2	VERSION 3	VERSION 4
	Headings	Headings/ Instruction	No Headings	Indented Headings
	%	%	%	%
Marked in Top	99.7	100.0	99.1	100.0
Marked in Bottom	70.4	66.0	41.8	67.2
Marked in Both	70.2	66.0	40.9	67.2
Mean # Checked	3.0	2.4	2.7	2.9
Mean # Checked in Top	1.9	1.5	2.0	1.9
Mean # Checked in Bottom	1.1	0.9	0.8	1.0
MEAN TIME (seconds)	13.6	13.5	14.3	15.3
n	435	438	367	351

	SIGNIFICANCE TESTS		
	Headings vs. No Headings (V1 vs. V3)	Headings vs. Headings/Instruction (V1 vs. V2)	Headings vs. Indented Headings (V1 vs. V4)
Marked in Top	$X^2 = 1.45, p = .228$	$X^2 = 1.012, p = .315$	$X^2 = 0.80, p = .371$
Marked in Bottom	$X^2 = 59.72, p = .000*$	$X^2 = 1.836, p = .175$	$X^2 = 0.88, p = .348$
Marked in Both	$X^2 = 62.36, p = .000*$	$X^2 = 1.630, p = .202$	$X^2 = 0.75, p = .387$
Mean # Checked	$t = 2.38, p = .018*$	$t = 5.41, p = .000*$	$t = 1.06, p = .290$
Mean # Checked in Top	$t = -1.36, p = .175$	$t = 5.99, p = .000*$	$t = 0.24, p = .813$
Mean # Checked in Bottom	$t = 4.89, p = .000*$	$t = 3.51, p = .001*$	$t = 1.49, p = .137$
MEAN TIME (seconds)	$t = -1.25, p = .212$	$t = 0.04, p = .966$	$t = -2.37, p = .018*$

Notes: There was no significant difference in the use of the Don't Know option across versions, nor was there a significant difference in missing data across versions. Percents exclude missing data as well as those who marked the "Don't Know" option. * $p \leq .05$

The addition of the instruction to “Please select the best answer” in an attempt to override the visual grouping effect (version 2) had a mixed effect. Respondents to the version with the instruction (version 2) were only slightly less likely to choose an answer in each sub-grouping than were those who completed the version without the instruction (version 1) (chi square = 1.630, p = .202). However, the mean number of options checked was significantly less for those who had the additional instruction (version 2) than for any of the other treatments. These results intimate 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 in concert seemed to have indicated to the respondents that they were expected to select the best answer from each sub-group.

Table 2: 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	VERSION 3	VERSION 4
	Headings	No Headings	Headings	No Headings
	%	%	Yes/No	Yes/No
Marked in Left	73.5	78.5	78.1	77.3
Marked in Right	91.9	88.3	90.2	89.5
Marked in Both	64.6	66.8	69.9	68.8
Mean # Checked	2.9	3.1	3.2	3.2
Mean # Checked in Left	1.4	1.6	1.7	1.6
Mean # Checked in Right	1.4	1.5	1.5	1.6
MEAN TIME (seconds)	20.7	20.0	25.8	25.7
n	393	446	423	443

	SIGNIFICANCE TESTS		
	Headings vs. No Headings (V1 vs. V2)	Headings Y/N vs. No Headings Y/N (V3 vs. V4)	Headings vs. Headings Y/N (V1 vs. V3)
Marked in Left	$X^2 = 2.57, p = .109$	$X^2 = 0.07, p = .792$	$X^2 = 2.10, p = .148$
Marked in Right	$X^2 = 1.63, p = .201$	$X^2 = 0.11, p = .736$	$X^2 = 0.16, p = .692$
Marked in Both	$X^2 = 0.38, p = .538$	$X^2 = 0.12, p = .729$	$X^2 = 2.35, p = .125$
Mean # Checked	$t = -1.86, p = .064$	$t = 0.29, p = .775$	$t = -3.15, p = .002^*$
Mean # Checked in Left	$t = -1.72, p = .087$	$t = 0.84, p = .400$	$t = -2.63, p = .009^*$
Mean # Checked in Right	$t = -0.58, p = .559$	$t = -0.74, p = .462$	$t = -1.36, p = .176$
MEAN TIME (seconds)	$t = 0.87, p = .384$	$t = 0.11, p = .911$	$t = -6.22, p = .000^*$

Notes: There was no significant difference in missing data across versions. All four versions were grouped the same with respect to response options. Mean checked refers to the mean number marked “yes”. It ignores how many responses were checked “no”. Respondents taking over 2 minutes were removed from calculations of mean times. * $p \leq .05$

The results from the second survey (Table 2) are somewhat different than those from the first. There was no significant difference between the percent of respondents who marked responses in the left, right, or both left and right groups across the versions with and without the headings. In addition, there was no significant difference in the mean number of response options marked in the left, right, or overall across the version with the heading and the version without. These findings are consistent with our expectation that the groupings are overshadowed by the use of the double-banking format.

Converting the answer options to a forced choice format had mixed effects. When the headings were present (Version 1 vs. Version 3) conversion to the forced choice format did not significantly affect the percent of respondents choosing responses in the left, right, or both sides, but it did significantly affect the mean number of options marked (affirmatively). When the response options were in the forced choice format respondents selected significantly more options affirmatively overall (3.2 vs. 2.9; $t = -3.15$, $p = .002$) and in the left group (1.7 vs. 1.4; $t = -2.63$, $p = .009$) than when the response options were not in the forced choice format. However, when the headings were not present (Version 2 vs. Version 4) conversion to the forced choice format did not have any significant effects.

Overall, these findings suggest that the headings and spatial separation play a larger role when the groups are aligned vertically as opposed to when they are aligned horizontally. They support the hypothesis that the grouping is overshadowed by the double-banking of the response options. In addition, they suggest that grouping via the use of headings and spatial separation may not have the same results across different question formats. Alternative question formats may moderate the message that such grouping sends.

SELECTING A FIELD OF STUDY

The second problematic question in the NSF Earned Doctorate Internet Survey requested that respondents choose their Ph.D. Field of Study from a long list of possible majors, which because of the number of options (282 fields organized under 13 headings and 16 sub-headings) had to be displayed on two different screens. On the paper questionnaire these options were presented on a single page at the end of the survey and did not appear to cause significant problems for respondents. However, several difficulties were encountered by respondents in the web prototype. For example, when their desired fields were not available some respondents repeatedly toggled between screens trying to find an appropriate response option. The toggling was inefficient and frustrating to a number of these respondents, appearing to be a significant problem for them.

Several grouping issues emerged from these observations. One concern was to find a more efficient method of locating the desired response category. Drop-down menus emerged as a possibility, but seemed to present a potential problem if they were to be constructed in the three level manner used for the paper questionnaire (general area, sub-area, and specific field). For example, if a respondent entered a general and then sub area field only to find that their specific field was not included he/she would have to initiate

another hierarchical search through another path. Observations in cognitive interviews indicated that this was a realistic scenario (Alzheimer and Dillman 2002). In fact, respondents searching for fields such as environmental science, criminal justice, and statistics searched under various general areas until they found the best or perhaps least objectionable answer. Another possible solution to the grouping issues was to present the groups of fields as one completely alphabetized list (one step) rather than as alphabetized fields contained under subheadings (multiple steps). A final possibility was to pose the field of study question as an open-ended question, thus avoiding the grouping process altogether. These alternative possibilities posed two main difficulties: (1) making provision for the possibility that students had more than one field of study and (2) determining whether or not the structure of the question would encourage or discourage the reporting of multiple majors.

A two-step experimental design was used to address these issues. In the first experiment, respondents were asked to report their major. The following are the treatments and can be seen in Figure 3 of Appendix A:

Version 1: The question was followed by an open-ended answer space large enough for reporting more than one major. In this treatment grouped information was not presented to the respondent. Instead, determining and reporting an appropriate answer was left entirely to the respondent.

Version 2: A two-step drop down menu was used. The first menu listed general areas as used on the University web page. When a general area was clicked specific majors would appear in the second drop down menu. It was impractical to test the three stage drop down menus approach with our population of undergraduate students.

Version 3: An alphabetized list of all university majors was presented. Respondents used the scroll bar to proceed through the list. Radio buttons were used so that only one major could be selected.

Version 4: An alphabetized list organized by general areas (categorized in the same way as in version 2) was presented. Radio buttons were used so that only one major could be selected.

The second experiment differed from the first in that it was constructed to allow for the possibility of multiple majors. The treatments were the same as in the first experiment save for accommodations made to allow the second or subsequent majors to be reported (See Figure 4 of Appendix A). For example, the wording of the question was changed to ask the respondents to report their “major or majors.” In addition, two sets of drop down menus were provided for the second version of this experiment. Answer spaces for versions three and four consisted of html boxes rather than radio buttons to allow for multiple selections as well. An inadvertent programming error resulted in one of the questions (version 3) retaining the original question wording (asking for “major” instead of “major or majors”).

The test of these formats across the two experiments embraces several grouping issues. One of them is the efficiency of the search process, and whether respondents are able to locate and mark their answer(s) more quickly in some formats than others. Another involves grouping based upon one step versus a multiple-step process. In addition there is the issue of providing a complete group from which answers are to be selected versus depending upon respondent recall to provide an answer in an open-ended fashion. Thus, this experiment involved different methods of grouping verbal information (completely alphabetical vs. alphabetical within general areas) as well as graphical layout.

Results for Selecting a Field of Study

The most important issue with this question is which format for grouping the response options provides for the most efficient processing of the question and response options. Through the use of paradata the mean times spent on each version were calculated (See Table 3). Outliers were removed at two minutes. For the first survey the open ended version was completed the quickest with a mean response time of 9.6 seconds. This was followed by the alphabetical list (14.2 seconds) and then the alphabetical list under general headings (21.1 seconds). Due to a programming error no time data was available for the drop down version on this first survey. That data was available on the second survey, however.

Table 3: Effects of Grouping on Field of Study Question in First Survey

Q20: What is your major? (Survey #1)				
	VERSION 1	VERSION 2	VERSION 3	VERSION 4
	<u>Open-Ended Box</u>	<u>Drop Down Menus</u>	<u>Alphabetical List</u>	<u>Grouped/ Alphabetic</u>
	%	%	%	%
Other	9.7	3.2	2.2	5.1
Undecided	3.0	5.5	0.8	4.0
Missing	1.4	0.0	1.4	0.6
Double Majors	9.4	----	----	----
Triple Majors	0.2	----	----	----
Gave Extra Detail	10.3	----	----	----
Used Abbreviations	4.8	----	----	----
TIME (seconds)	9.6	----	14.2	21.1
n	435	438	367	351

TESTS OF SIGNIFICANCE		
	t	p
Open-Ended vs. Alphabetical (V1 vs. V3)	-5.43	.000*
Open-Ended vs. Grouped/Alphabetical (V1 vs. V4)	-12.31	.000*
Alphabetical vs. Grouped Alphabetical (V3 vs. V4)	-6.06	.000*

Notes: In version one those majors that were not on the list used for versions 2 through 4 are counted as “other.” Respondents taking over 2 minutes were removed from calculations of mean times. Due to a programming error the data for the mean time was not available for Version 2. * p ≤ .05

On the second survey the mean time it took to complete the open ended version was lowest at 15.3 seconds (See Table 4)². This grouping scheme was followed by the drop down menus version which took on average about 34.4 seconds to complete. Responding to the alphabetical list version took a mean time of 50.0 seconds. When the response options were grouped alphabetically under general areas the mean response time was 57.3 seconds. The format that took the longest to respond to was the grouped/alphabetic format which most closely resembles the NSF format. In terms of time, the open ended version was the most efficient while the grouped/alphabetic version was the least efficient.

Table 4: Effects of Grouping on Field of Study Question in Second Survey

Q23: What is your major or Majors? (Survey #2)				
	VERSION 1	VERSION 2	VERSION 3	VERSION 4
	<u>Open-Ended Box</u>	<u>Drop Down Menus</u>	<u>Alphabetical List</u>	<u>Grouped/ Alphabetic</u>
	%	%	%	%
Other	3.1	----	1.9	2.3
Undecided	3.3	4.3	4.3	4.7
Missing	7.4	9.6	13.7	11.7
Double Majors	11.5	16.4	12.5	18.1
Triple Majors	0.5	----	2.1	2.3
Gave Extra Detail	10.2	----	----	----
TIME (seconds)	15.3	34.4	50.0	57.3
n	393	446	423	443

TESTS OF SIGNIFICANCE		
	t	p
Open Ended vs. Drop Down (V1 vs. V2)	-15.17	.000*
Open Ended vs. Alphabetical (V1 vs. V3)	-27.12	.000*
Open Ended vs. Grouped/Alphabetic (V1 vs. V4)	-32.41	.000*
Drop Down vs. Alphabetical (V2 vs. V3)	-10.13	.000*
Drop Down vs. Grouped/Alphabetic (V2 vs. V4)	-14.85	.000*
Alphabetical vs. Grouped/Alphabetic (V3 vs. V4)	-4.63	.000*

Notes: In version one those majors that were not on the list used for versions 2 through 4 are counted as "other." In version 3 one person checked 4 options and another checked 10 options. In version 4 one person checked four options and one person checked six options. Respondents taking over two minutes were removed from calculations. * $p \leq .05$

² The mean times for the second survey are consistently greater than for the first because they are page-elapse times whereas the mean times for the first refer to how long it took to check an answer independently of the time it took the page to load. The page-elapse times are used for the second survey so that the drop down version could be included. The important thing here is not the exact time it took for each version, but the consistent pattern of times relative to one another across the two surveys.

However, the open ended version also required coding by the researchers and introduced the ability for the respondent to use unclear abbreviations, provide unsolicited explanation, and list as many majors as space allowed. About 4.8 percent of respondents listed their major using abbreviations that could not be coded in the first survey. Ten percent provided extra information such as providing their major but then explaining that they are about to change it or explaining the emphasis they have within the major (i.e. sociology, family emphasis). Another nine percent provided double majors, and one person listed three majors on the first survey. Comparable figures can be found in Table 4 for the open ended version on the second survey.

Another important finding regards the reporting of double majors. Whereas the first survey did not allow respondents to report double majors, except in the open ended version, the second was designed specifically to do so. Results indicate that on average across the four versions about 15 percent of respondents reported double majors, suggesting that the inability to do so in the first survey may have inconvenienced a significant number of respondents. As a matter of fact, paradata analysis indicates that only one percent of respondents to the open-ended version of the first survey changed their answer. In contrast, 10.5 percent of respondents to the alphabetical listing and 13.9 percent of respondents to the grouped/alphabetical version changed their answers. Of those changes on the alphabetical and grouped/alphabetical versions 63.6 percent appeared to be due to the lack of ability to report double majors. These respondents marked an answer and then attempted to mark another answer, but apparently upon realizing that their second answer voided their first they then went back and re-marked their original selection.

REQUEST FOR MONTH AND YEAR OF AN ACTIVITY

The third question format evaluated in this study was a simple request for respondents to record the month and year they completed their degree, using two digits for the month and four digits for the year. In the NSF web survey prototype several respondents completed this question inaccurately—either entering letters for a month or only two digits for the year—and received an error message. As a result of these difficulties the NSF question that asked for the month and year of degree completion was converted to a question asking undergraduates for the month and year that they began their studies at WSU. The following treatments were designed to ascertain how best to get respondents to use the correct reporting format (See Figure 5 in Appendix A):

Version 1: This version emulates the original NSF format with the symbols “MM YYYY” located to the right of both answer boxes.

Version 2: The same answer boxes are used in this version, but the symbols are moved directly below the boxes such that the “MM” is located below the month box and the “YYYY” is located below the year box.

Version 3: Again, the same answer boxes are used, but the symbols for month and year are replaced with the words “Month” and “Year” placed directly under their appropriate answer boxes.

Version 4: In this version the “Month” and “Year” designation was maintained under the answer boxes, but the size of the box for reporting the month was reduced. This treatment was intended to give the impression through graphical manipulation that only two digits were desired for the month while the larger year box required four digits.

These specific treatments were designed because it was hypothesized that the abovementioned errors might result from the grouping of the month and year boxes together, followed by the grouping of the symbols “MM YYYY” together after the boxes (as shown in Version 1 of Figure 5 in Appendix A). Thus, the symbol for the month was placed after the answer box for the year, separating the instruction from the answer space to which it was intended to apply. Similarly, the symbol for the year was placed after the symbol for the month, creating interruption between it and the answer space to which it was intended to apply. In addition, the “YYYY” may fall outside the respondents’ foveal view. We replicated this design with version one to provide a comparison. Version 2 was designed using the principle of proximity. We expect the closer grouping of the specific symbols with their respective answer boxes to result in a more clear understanding of expectations among respondents and as a result more respondents using the desired format. We expect the third version to encourage the use of letters to report the month rather than numbers as the word “Month” does not carry the same symbolic meaning about the number of digits that the symbol “MM” conveys in that context. Finally, we expect the shortening of the month answer box to result in less use of letters and more use of numbers as the expectations communicated by the graphical shrinking of the box in concert with the expectations communicated by the word “Month” communicate to the respondent how he/she should format the response. Thus, the experiments in the first survey evaluate multiple issues related to how individuals group information—the use of symbols versus words, the location of the symbols, and the graphical aspect of size.

The follow-up experiment used the treatment from the first set of experiments that most successfully encouraged the desired two-digit month/four-digit year response format and extended it in three ways (See Figures 6 and 7 in Appendix A).

Version 1: The symbols “MM” and “YYYY” were located directly below the appropriate answer boxes.

Version 2: The format of the answer boxes and symbols was retained from version 1, but the stem of the question was changed such that it asked “what month and year did you begin your studies . . .” as opposed to the previously used “When did you begin your studies . . .” wording.

Version 3: This version contained the stem change as in version two, but the symbols were replaced with the words “Month” and “Year” under the appropriate answer boxes.

Version 4: For this version two drop down menus were used. In the first the respondent reported the month and in the second the respondent reported the year. The question stem referred to both month and year in this version as well.

The stem change was implemented to determine whether such additional specification would encourage the desired response and/or affect the efficiency with which it was provided. In addition, a drop down menu format was used as a further evaluation of the efficiency of the possible formats.

Results for Month and Year of an Activity

Respondents who answered the date questions using a two-digit month and a four-digit year were regarded as using the “desired format.” Results from the first survey indicate that locating the symbols below the answer spaces as opposed locating them to the right of the answer spaces resulted in a significantly larger proportion of respondents using the four-digit format for reporting the year on this question (94% vs. 89.9%). As such, the version in which the symbols were placed below as opposed to at the right of the answer box produced more use of the desired format (90.6% vs. 88.5%), although the difference was not significant. Using words below the answer space instead of symbols significantly decreased the percent of respondents reporting the date in the desired format (55.3%). However, with words below the answer spaces, reducing the size of the answer space for the month such that it appeared to be about half the size of the answer space for the year led to significant increases in those reporting the date using a four-digit year (76.9% vs. 67.3%) and as such those using the desired format (63.3% vs. 55.3%).

The findings regarding the difference between the use of symbols and words to indicate expectations were replicated twice in the second survey. When students were asked the date at which they began their studies 87.2 percent used the desired format when symbols were located below the answer spaces while only 45.4 percent did so when words were located below the answer spaces. When asked when they expected to complete their studies 70.9 percent used the desired format to report the expected date when symbols were located below the answer space while 35 percent did so when words were located below the answer space.

Paradata time analysis is not available for the set of experiments in the first survey, but is available for both questions in the second survey. There were very few significant time differences between the versions. In the first question, asking when the students began their studies the drop down version was completed the quickest on average. It was significantly faster than both the version with the symbols below the answer spaces (Version 2) and the version with words below the answer space (Version 3). In the

second question the drop down menus version was not the quickest version, but instead the version with the symbols below the answer spaces was. However, none of the mean times varied significantly from each other.

Table 5: Results for Month and Year of an Activity from the First Survey

Q2: When did you first begin your studies at WSU?				
	VERSION 1	VERSION 2	VERSION 3	VERSION 4
	Symbols Right	Symbols Below	Words Below	Words Below, Sm. Mo. Box
MONTH	%	%	%	%
1 digit	2.5	3.4	18.3	20.5
2 digits	97.5	96.1	80.4	78.9
Alpha	0.0	0.22	1.10	0.6
Other	0.0	0.22	0.27	0.0
YEAR				
2 digits	9.9	5.5	32.4	21.7
4 digits	89.9	94.0	67.3	76.9
Other	0.2	0.5	0.3	1.4
Desired Format	88.5	90.6	55.3	63.3
n	435	438	367	351

TESTS OF SIGNIFICANCE		
	Chi Squared	p
DESIRED FORMAT		
Symbols Right vs. Symbols Below (V1 vs. V2)	1.06	.302
Symbols Below vs. Words Below (V2 vs. V3)	131.19	.000*
Words Below vs. Words Below Sm. Mo. Box (V3 vs. V4)	4.68	.031*
FOUR-DIGIT YEAR		
Symbols Right vs. Symbols Below (V1 vs. V2)	5.93	.015*
Symbols Below vs. Words Below (V2 vs. V3)	99.07	.000*
Words Below vs. Words Below Sm. Mo. Box (V3 vs. V4)	9.95	.002*

Notes: Desired format consists of a two-digit month and a four-digit year. Mean times are not available for this item. * $p \leq .05$

Table 6: Results for Month and Year of an Activity from the Second Survey

Q1: When did you first begin your studies at WSU?				
	VERSION 1	VERSION 2	VERSION 3	VERSION 4
	Symbols Below, Stem Change	Symbols Below	Words Below	Drop Down
MONTH	%	%	%	%
1 digit	1.02	2.69	17.49	0.00
2 digits	90.33	88.12	50.59	0.00
Alpha	1.53	1.79	25.06	95.49
Other	7.12	7.40	6.86	4.51
YEAR				
2 digit	1.02	0.90	7.80	0.00
4 digits	92.11	91.70	85.11	94.81
Other	6.87	7.40	7.09	5.19
Desired Format	89.31	87.22	45.39	0.00
Mean Time (sec)	22.0	22.3	23.1	19.6
n	393	446	423	443

TESTS OF SIGNIFICANCE		
	Chi Squared	p
DESIRED FORMAT		
Symbols Below, Stem Change vs. Symbols Below (V1 vs. V2)	0.88	.348
Symbols Below vs. Words Below (V2 vs. V3)	171.43	.000*
MEAN TIME		
Symbols Below, Stem Change vs. Symbols Below (V1 vs. V2)	-0.31	.754
Symbols Below, Stem Change vs. Words Below (V1 vs. V3)	-1.00	.317
Symbols Below, Stem Change vs. Drop Down (V1 vs. V4)	1.90	.058
Symbols Below vs. Words Below (V2 vs. V3)	-0.71	.477
Symbols Below vs. Drop Down (V2 vs. V4)	2.27	.024*
Words Below vs. Drop Down (V3 vs. V4)	2.91	.003*

Notes: Desired format consists of a two-digit month and a four-digit year. Respondents taking over two minutes were removed from calculations of mean times. * $p \leq .05$

Table 7: Results for Month and Year of an Activity from the Second Survey

Q22: When do you expect to complete your studies at WSU?				
	VERSION 1	VERSION2	VERSION 3	VERSION 4
	Symbols Below, Stem Change	Symbols Below	Words Below	Drop Down
MONTH	%	%	%	%
1 digit	0.76	1.12	15.84	0.00
2 digits	69.72	71.97	40.43	0.00
Alpha	1.78	1.79	18.44	79.23
Other	27.74	25.11	25.30	20.77
YEAR				
2 digit	0.00	0.22	6.15	0.00
4 digits	73.54	73.99	67.38	76.75
Other	26.46	25.78	26.48	23.25
Desired Format	69.47	70.85	34.99	0.00
Mean Time (sec)	19.6	19.2	20.6	21.3
n	393	446	423	443

TESTS OF SIGNIFICANCE			
		Chi Squared	p
DESIRED FORMAT			
Symbols Below, Stem Change vs. Symbols Below (V1 vs. V2)		0.19	.661
Symbols Below vs. Words Below (V2 vs. V3)		112.21	.000*
MEAN TIME			
Symbols Below, Stem Change vs. Symbols Below (V1 vs. V2)	t	0.39	.698
Symbols Below, Stem Change vs. Words Below (V1 vs. V3)		-0.92	.358
Symbols Below, Stem Change vs. Drop Down (V1 vs. V4)		-1.48	.138
Symbols Below vs. Words Below (V2 vs. V3)		-1.38	.169
Symbols Below vs. Drop Down (V2 vs. V4)		-1.94	.053
Words Below vs. Drop Down (V3 vs. V4)		-0.68	.494

Notes: Desired format consists of a two-digit month and a four-digit year. Respondents taking over 2 minutes were removed from calculations of mean times. * $p \leq .05$

DISCUSSION AND CONCLUSIONS

In this paper we have investigated, experimentally, how certain aspects of the visual design and layout of questions influence how respondents process survey questions and provide answers in web surveys. Our focus is on understanding multiple aspects of the answering process, from whether visual layout changes people's answers to how best to communicate exactly what the designer wants so response errors, and corrections, are minimized. We are also interested in the efficiency by which people arrive at their answers. Thus, our interest is not only in obtaining accurate answers, but also in minimizing respondent burden and frustration.

The central issue explored in this paper is how different visual groupings of information affect how respondents arrive at answers—which designs promote accuracy, which are most efficient, and which will minimize the need for corrections. Because grouping is only one of several considerations that may influence respondent answers, the experimentation reported here has in some cases gone beyond grouping to investigate related aspects, such as word instructions, an understanding of which may help us better understand the relative importance of visual grouping.

Each of the questions investigated here represented a format used in a web survey prototype developed by the National Science Foundation in pursuit of changing a mail questionnaire to an electronic data collection instrument. And, the three questions chosen for experimentation had each presented problems to respondents in cognitive interviews of the prototype instrument.

The first issue investigated was whether the use of response category subheadings influenced respondent answers. Use of two headings which not only subcategorized categories topically, but resulted in graphical separation that made them appear as two distinct groups had a dramatic effect on people's answers. Many respondents felt compelled to provide answers in each sub-group when the headings and spacing were used (e.g. 70% marked an answer in both sub-groups, but only 41% did so when no headings or separation were used). We also found that the word instructions, "Please select the best answer," only partially mitigated the effect. Respondents checked significantly fewer items when the instructions were present (2.4 vs. 3.0), but were still just as likely to mark an answer in each of the sub-groups. It is obvious that visual grouping and separation have a dramatic effect on answers with or without the explicit instructions. At the same time, when the response choices were double banked on a different topic, very few differences were observed. These results suggest that separating a vertical list of categories into sub-groups is not a desirable construction practice. Doing so appears to set an expectation that the respondent provide an answer in each group. Thus, visual layout has an obvious and significant impact on responses.

The second issue focused on how the grouping of a large number of response options influences answering behavior. Of particular importance is the fact that the question asks for information that the respondent already knows, but must find among a long list of options. Asking for field of study or student major seems like a straightforward question, but the results indicate that how the options are grouped greatly influences the answering process. First, the results show that respondents can process some formats more efficiently than others. For example, a two-step grouping (organizing majors under general areas) is less efficient for respondents than a single grouping (alphabetical list) of the choices, but simply asking for an open-ended response and allowing respondents to report is the most efficient format. Second, differences in response patterns occur because of how response information is grouped. In the first survey, nearly 10 percent of respondents to the open ended question provided double or triple majors. When this was allowed in all versions of the second survey, slightly higher numbers of respondents indicated they had multiple majors (15% averaged across versions). The drop-down menu used in the second survey provides visually separate places for listing first and

second majors, and as such encouraged a greater number of people to do that (16.4%). Thus, we see time and response differences depending upon how the information is presented to respondents. It seems important to recognize that for what appears to be a straightforward question, response burden and response patterns differ greatly depending on how the question is visually presented to the respondent.

The third issue involved communicating to respondents the format they were expected to use when reporting a date. The experiments included a number of treatments such as grouping of information, the use of symbols versus words to communicate expectations, and graphical manipulations. Results indicate that grouping the symbols with the answer spaces in order to communicate expectations is effective in conveying to respondents the desired format. Regardless of the position of the symbols, most respondents tended to use a two-digit month; however, when the symbols “MM YYYY” were placed directly below their respective answer boxes instead of to the right of the answer boxes respondents were significantly more likely to use a four-digit year (94% vs. 89.9%). This led to a larger proportion of respondents reporting the date using the desired format when the symbols were placed below the answer boxes (90.6 vs. 88.5) and overall, indicated that symbols are most efficient at communicating expectations when they are grouped, through proximity, with answer boxes. While grouping the symbols with the answer boxes appeared to increase the use of the desired format, other manipulations also encouraged its use. For example, one experiment on the first survey and two on the second showed indisputably that the use of the symbols “MM YYYY” were more effective in inspiring use of the desired format than was the use of the words “Month” and “Year” in their place. In addition, graphical manipulation of the size of the answer box significantly affected the proportion of respondents using the desired format. These findings suggest that grouping in combination with the use of other visual design aspects such as symbols and graphical manipulation of the month box would result in greater use of the desired format. In addition, they indicate that response patterns can be greatly affected by the design of such questions.

In all three question types the grouping of information had significant effects on both the efficiency with which respondents were able to answer and the way they answered the questions. In addition, the evidence suggests that grouping interacts with other visual design elements as well as question formats and wording and as such its effect may vary depending on these other elements. Although much remains to be done to determine the effects of asking questions using different grouping schemes, the main conclusion to be reached at this point is that different approaches lead to different results and therefore need to be evaluated carefully before using them.

APPENDIX A
Experimental Treatments

Figure 1: Experimental Treatments for the Effects of Category Headings in the First Survey

VERSION 1

Q18. What best describes the benefit of the Student Recreation Center?

Health Benefits

- The variety of physical fitness offerings
- The health and wellness offerings
- Helps reduce stress

Academic Benefits

- Improves academic productivity
- Enhances learning experience
- Provides information for students to learn about their health

Don't Know

Next Question

VERSION 2

Q18. What best describes the benefit of the Student Recreation Center? Please select the best answer.

Health Benefits

- The variety of physical fitness offerings
- The health and wellness offerings
- Helps reduce stress

Academic Benefits

- Improves academic productivity
- Enhances learning experience
- Provides information for students to learn about their health

Don't Know

Next Question

VERSION 3

Q18. What best describes the benefit of the Student Recreation Center?

- The variety of physical fitness offerings
- The health and wellness offerings
- Helps reduce stress
- Improves academic productivity
- Enhances learning experience
- Provides information for students to learn about their health
- Don't Know

Next Question

VERSION 4

Q18. What best describes the benefit of the Student Recreation Center?

Health Benefits

- The variety of physical fitness offerings
- The health and wellness offerings
- Helps reduce stress

Academic Benefits

- Improves academic productivity
- Enhances learning experience
- Provides information for students to learn about their health

Don't Know

Next Question

Figure 2: Experimental Treatments for the Effects of Category Headings in the Second Survey

VERSION 1

Question 21 of 25

Have you received financial support from each of the following sources while attending WSU?

<u>Financial Aid</u>	<u>Other Sources</u>
<input type="checkbox"/> Stafford Loans	<input type="checkbox"/> Parents
<input type="checkbox"/> Other Loans	<input type="checkbox"/> Other family
<input type="checkbox"/> Scholarships	<input type="checkbox"/> Full-time employment
<input type="checkbox"/> Work-study	<input type="checkbox"/> Part-time employment

VERSION 2

Question 21 of 25

Have you received financial support from each of the following sources while attending WSU?

<input type="checkbox"/> Stafford Loans	<input type="checkbox"/> Parents
<input type="checkbox"/> Other Loans	<input type="checkbox"/> Other family
<input type="checkbox"/> Scholarships	<input type="checkbox"/> Full-time employment
<input type="checkbox"/> Work-study	<input type="checkbox"/> Part-time employment

VERSION 3

Question 21 of 25

Have you received financial support from each of the following sources while attending WSU?

<u>Financial Aid</u>	Yes	No	<u>Other Sources</u>	Yes	No
Stafford Loans	<input type="radio"/>	<input type="radio"/>	Parents	<input type="radio"/>	<input type="radio"/>
Other Loans	<input type="radio"/>	<input type="radio"/>	Other Family	<input type="radio"/>	<input type="radio"/>
Scholarships	<input type="radio"/>	<input type="radio"/>	Full-time employment	<input type="radio"/>	<input type="radio"/>
Work-study	<input type="radio"/>	<input type="radio"/>	Part-time employment	<input type="radio"/>	<input type="radio"/>

VERSION 4

Question 21 of 25

Have you received financial support from each of the following sources while attending WSU?

	Yes	No		Yes	No
Stafford Loans	<input type="radio"/>	<input type="radio"/>	Parents	<input type="radio"/>	<input type="radio"/>
Other Loans	<input type="radio"/>	<input type="radio"/>	Other Family	<input type="radio"/>	<input type="radio"/>
Scholarships	<input type="radio"/>	<input type="radio"/>	Full-time employment	<input type="radio"/>	<input type="radio"/>
Work-study	<input type="radio"/>	<input type="radio"/>	Part-time employment	<input type="radio"/>	<input type="radio"/>

Figure 3: Experimental Treatments for the Effects of Grouping on Field of Study Question in the First Survey

VERSION 1

Q20. What is your major?

Next Question

VERSION 2

Q20. What is your major?

Click for area of study Click for major

Next Question

VERSION 3

Q20. What is your major?

- Accounting
- Accounting & Information Systems
- Agribusiness
- Agricultural Communications
- Agricultural Economics
- Agricultural Education
- Agricultural Technology & Management
- Agriculture (General)
- American Studies
- Animal Sciences
- Anthropology
- Apparel, Merchandising, & Textiles
- Architecture
- Asian Studies
- Athletic Training
- French
- Genetics and Cell Biology
- Geology
- German
- Health and Fitness
- History
- Horticulture
- Hotel and Restaurant Administration
- Human Development
- Human Nutrition and Foods
- Human Resources and Personnel
- Humanities (General)
- Integrated Pest Management
- Interdisciplinary
- Interior Design

VERSION 4

Q20. What is your major?

Agricultural Sciences

- Agricultural Technology & Management
- Agriculture (General)
- Animal Sciences
- Crop Science
- Food Science
- Horticulture
- Integrated Pest Management
- Soil Science

Architecture, Visual Arts, and Design

- Apparel, Merchandising, & Textiles
- Architecture
- Fine Arts
- Interior Design
- Landscape Architecture

Figure 4: Experimental Treatments for the Effects of Grouping on Field of Study Question in the Second Survey

VERSION 1

Question 23 of 25

What is your major or majors?

VERSION 2

Question 23 of 25

What is your major or majors?

Major 1

Major 2
(If applicable)

VERSION 3

Question 23 of 25

What is your major?

- Accounting
- Accounting Information Systems
- Agribusiness
- Agricultural Communications
- Agricultural Economics
- Agricultural Education
- Agricultural Technology Management
- Agriculture (General)
- American Studies
- Animal Sciences
- Anthropology
- Apparel, Merchandising, Textiles
- Architecture
- Asian Studies
- Athletic Training
- Biochemistry

VERSION 4

Question 23 of 25

What is your major or majors?

- Agricultural Sciences
- Agricultural Technology & Management
- Agriculture (General)
- Animal Science
- Crop Science
- Food Science
- Horticulture
- Integrated Pest Management
- Soil Science
- Architecture, Visual Arts, and Design
- Apparel, Merchandising, & Textiles
- Architecture
- Fine Arts
- Interior Design
- Landscape Architecture

Figure 5: Experimental Treatments for Asking for the Month and Year of an Activity in the First Survey

VERSION 1

Q2. When did you first begin your studies at WSU?

MM YYYY

VERSION 2

Q2. When did you first begin your studies at WSU?

MM YYYY

VERSION 3

Q2. When did you first begin your studies at WSU?

Month Year

VERSION 4

Q2. When did you first begin your studies at WSU?

Month Year

Figure 6: Experimental Treatments for Asking for the Month and Year of an Activity in the Second Survey (First Question)

VERSION 1

Question 1 of 25
When did you begin your studies at Washington State University?

MM YYYY

VERSION 2

Question 1 of 25
What month and year did you begin your studies at Washington State University?

MM YYYY

VERSION 3

Question 1 of 25
What month and year did you begin your studies at Washington State University?

Month Year

VERSION 4

Question 1 of 25
What month and year did you begin your studies at Washington State University?

-Month- -Year-

Figure 7: Experimental Treatments for Asking for the Month and Year of an Activity in the Second Survey (Second Question)

VERSION 1

Question 22 of 25

When do you expect to complete your studies at WSU?

MM YYYY

Don't know

VERSION 2

Question 22 of 25

What month and year do you expect to complete your studies at WSU?

MM YYYY

Don't know

VERSION 3

Question 22 of 25

What month and year do you expect to complete your studies at WSU?

Month Year

Don't know

VERSION 4

Question 22 of 25

What month and year do you expect to complete your studies at WSU?

-Month- -Year-

Don't know

REFERENCES

- Altheimer, Irshad and Don A. Dillman. 2002. "Results from Cognitive Interviews of NSF Earned Doctorate Web Survey." Social & Economic Sciences Research Center Tech Report #02-30.
- Best, Samuel J. and Brian Krueger. 2003. *Internet Data Collection*. Unpublished Manuscript.
- Christian, Leah Melani. 2003. "The Influence of Visual Layout on Scalar Questions in Web Surveys." Unpublished Master's Thesis, Washington State University.
- Christian, Leah Melani and Don A. Dillman. 2004. "The Influence of Graphical and Symbolic Language Manipulations on Responses to Self-Administered Questions." *Public Opinion Quarterly*. 68:58-81.
- Couper, Mick P. Roger Tourangeau, and Frederick G. Conrad. 2004. "What They See is What We Get: Response Options for Web Surveys." *Social Science Computer Review*. 22(1):111-127.
- Dillman, Don A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. 2nd ed. New York, NY: John Wiley and Sons, Inc.
- Dillman, Don A., Jolene D. Smyth, Leah M. Christian, and Michael J. Stern. 2003. "Multiple-Answer Questions in Self-Administered Surveys: The Use of Check-All-That-Apply and Forced-Choice Question Formats." Paper presented at the meetings of the American Statistical Association, San Francisco, CA, August 2003.
- Heerwegh, D. & G. Loosveldt. 2002. "Describing Response Behavior in Websurveys Using Client Side Paradata." Paper Presented at the International Workshop on Web Surveys, Mannheim, Germany, October 2002.
- Jenkins, Cleo R. and Don A. Dillman. 1997. "Towards a Theory of Self-Administered Questionnaire Design." in *Survey Measurement and Process Quality* edited by Lyberg, Biemer, Collins, de Leeuw, Dippo, Schwarz, and Trewin. John Wiley & Sons, Inc.
- Neisser, Ulric. 1967. *Cognitive Psychology*. New York, NY: Meredith Publishing Company.
- Palmer, Stephen E. 1999. *Vision Science Photons to Phenomenology*. Cambridge, MA: MIT Press.
- Rasinski, Kenneth A., David Mingay, and Norman M. Bradburn. 1994. "Do Respondents Really 'Mark All That Apply' on Self-Administered Questions?" *Public Opinion Quarterly*. 58: 400-408.

- Redline, Cleo D and Don A. Dillman. 2002. "The Influence of Alternative Visual Designs on Respondents' Performance with Branching Instructions in Self-Administered Questionnaires." in *Survey Nonresponse* edited by Robert M. Groves, Don A. Dillman, John L Eltinge, and Roderick J. A. Little. New York, NY: John Wiley & Sons.
- Redline, Cleo, Don A Dillman, Aref N. Dajani, and Marry Ann Scaggs. 2003. "Improving Navigational Performance in U.S. Census 2000 by Altering the Visually Administered Languages of Branching Instructions." *Journal of Official Statistics*. 19(4):403-419.
- Schwarz, Norbert. 1996. *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Schwarz, Norbert and Hans J. Hippler 1992. "Context Effects and Questionnaire Layout." Unpublished Research, ZUMA, Mannheim, Germany. Cited in *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. by Norbert Schwarz (1996).
- Stern, Michael J, Don A. Dillman, Jolene D. Smyth, and Leah Melani Christian. 2004. "The Uses of Paradata for Evaluating Alternative Versions of Web Survey Questions." Paper presented at the 57th annual conference of the World Association for Public Opinion Research. Phoenix, Arizona. May 11-13, 2004