# The Effects of Grouping Response Options in Factual Questions with Many Options


## Cleo Redline, Roger Tourangeau, Mick Couper, Fred Conrad, and Cong Ye

**Cleo Redline, Joint Program in Survey Methodology, University of Maryland; Roger Tourangeau, Mick Couper, Fred Conrad, Joint Program in Survey Methodology, University of Maryland and Survey Research Center, University of Michigan; and Cong Ye, Joint Program in Survey Methodology, University of Maryland**

credline@survey.umd.edu; rtourangeau@survey.umd.edu; MCouper@umich.edu; fconrad@isr.umich.edu; cye@survey.umd.edu


### Introduction

Question comprehension depends upon social and communicative processes that go beyond the words of a question (e.g., Schwarz, 1996; Sudman, Bradburn, and Schwarz, 1996). Respondents are particularly likely to have difficulty answering questions if their situations map onto a question in a complicated or peripheral way (e.g., Conrad, Schober, and Coiner, 2007; Tourangeau, Conrad, Arens, Fricker, and Smith, 2006). In addition, studies in the last decade have shown that visual features of a survey question have an effect on comprehension (e.g., Christian and Dillman, 2004; Couper, Tourangeau and Kenyon, 2004; Tourangeau, Couper, and Conrad, 2004; Tourangeau, Couper, and Conrad, 2006; Schwarz, 1998). Still, there is much we do not know about the effects of formatting questions and response options on respondents' answers to survey questions. In this paper, we examine the impact of manipulating the presentation of response options in two factual questions. We examine whether presenting long lists of response options in conceptual groupings, either with or without headings and in a single column versus multiple columns facilitate respondents' answering of these questions.

### Background

The principles that guide peoples' interpretation of visual information, including those that are related to the grouping and spacing of information, were illustrated in Jenkins and Dillman (1997). For example, when similar figures are located in close proximity to one another, people tend to see them as belonging to the same group. Expanding upon these principles, Tourangeau and his colleagues (Tourangeau, Couper, and Conrad, 2004; Tourangeau, Couper, and Conrad, 2006) have suggested five heuristics that guide respondents in interpreting a number of visual cues in surveys. The heuristic most directly related to the topic of this paper is 'Near means related.' This heuristic states that respondents expect items that are physically near each other on the screen (or on the page) to be related conceptually. In line with this heuristic, the researchers found that correlations among a set of related questions were heightened when the item appeared together in a grid on a single Web page than when each question was presented in a single screen.

Recently, Smyth, Dillman, Christian and Stern (2006) found evidence in support of this heuristic as well. They found that respondents tended to respond differently when response options were offered in different formats. In one format, six to eight response options were separated into two groups; headings were given for each group and the options were arrayed down the screen in one column. In a second format, one continuous list of response options was offered (that is, there was no grouping of the response options and no headings). The researchers found that respondents were more likely to choose at least one answer from each of the subgroupings in the first format. The researchers also found that the percent of respondents selecting more than one response was more pronounced for an opinion question (which asked students to describe the benefit of a student recreational center) than a factual question (in which students were asked to describe the type of financial aid they received). The researchers did not find a difference in respondents' answers when the options were arrayed in two columns rather than one column (both of which had headings), nor did they find a difference within the two-column format between the versions with and without headings.

Do these findings extend to questions about education and transportation? These items figure prominently in national surveys (such as the American Community Survey). These questions ask about personal facts that respondents can generally be expected to know (that is, they involve little difficulty recalling). In addition, they contain twice as many response options from which to choose as those studied by Smyth and her colleagues. And they contain an instruction emphasizing that respondents should choose one option only (refer to Figure 1).

Headings have been shown to help readers. For example, in instructional text, headings help readers scan, select, and retrieve materials more easily (Hartley, 1985). More recently, Hyona and Nurminen (2006) found that topic structure processors (those that looked back in text to read topic sentences and headings) had a better understanding of the text they had read. Therefore, we thought that separating the response options into three conceptual groups and introducing headings (for example, grade school or less, high school, and college or above) would help respondents find their selection more easily in questions with many response options. If you know, for example, that your highest level of school is college, the headings should allow you to find that answer in the third grouping more quickly.

Also, it is still unclear whether it is better to divide long lists of response options into two or more columns so that all responses are visible on the screen simultaneously, or to present a single list that requires scrolling to see the later items (e.g., Couper, 2008). We thought that arraying the response options across the screen in columns would quicken respondents' selections in factual questions with many response options because this arrangement eliminates the need to scroll over two screens. It should be noted, however, that this is only so long as the list of response options are not so long or the columns are not spread out to the extent that they require horizontal scrolling to view.

We examine differences in response distributions across the designs. In addition, however, we examine three other sources of data to get a more detailed picture of the response process. We examined the eye-movement data from a relatively small scale laboratory study. Researchers have used eye-movement analysis for some time to study cognitive processing in reading and scene perception (e.g., Rayner, 1998). More recently, eye-movement data has been successfully used to study respondents' processing of survey questions, response options, and branching instructions (Galesic, Tourangeau, Couper, and Conrad, 2008; Redline, Hunfalvay, and Lankford,2006; Redline and Lankford, 2001). We also examined response latencies and click data, which were collected as part of a larger-scale Web survey (see Yan and Tourangeau, 2008, for a discussion of response latencies and Heerwegh, 2003, for a discussion of click data).

As mentioned earlier, we selected two factual questions for study, for which we expected respondents to know the answers: education attainment and mode of transportation used in commuting. We predicted that respondents' answers would not vary between various question designs, but that response times would. We thought that response times would be reduced when the groups were arrayed in columns across the screen and when headings were introduced to guide respondents' selections.

**Experiment 1: Eye Movement Study**

**Methods**

This experiment was conducted at the University of Maryland from August 14, 2007 through October 5, 2007. A total of 160 respondents were recruited to participate in this experiment through the use of fliers placed in and around the University of Maryland. Respondents were offered $20.00 to participate. Half of the respondents (80) were students, the other half were not. A little over half of the respondents were female (90); the remainder were male.

We used the Tobii 1750 eye-tracker with ClearView analysis software to collect eye-movements (www.tobii.com). Data is collected simultaneously for each eye through the use of reflected near infra-red light. The eye-tracker is fully automatic and non-intrusive. Average accuracy over a set of ten individuals has been tested to 0.5 degree. One degree of accuracy corresponds to an average error of about 1 cm between the measured and intended gaze point at 50 cm distance from the user to the screen.

After signing a consent form, respondents were seated at a desk with a computer and a monitor that housed the eye-tracking equipment (as noted above, however, this equipment was not obvious to the respondent). It is necessary to calibrate the respondents' eyes to the system. This was accomplished by asking the respondent to look at a dot that appeared at specific positions on the screen. After calibration, respondents were instructed to log into the survey. The eye-tracker failed to capture eye movements for four of the 160 respondents that completed the survey.

We included two questions in this experiment that were patterned after questions from the American Community Survey. The questions were part of a survey that contained 26 questions in total. Both questions offered respondents a long list of options from which to select a response. The first question concerned education: "What is the highest degree or level of school you have completed?". This question included 15 response options.

The second question concerned transportation to work. First respondents were asked if they worked outside the home, and if they did, they were asked our study question: "How did you usually get to work last week?" This question listed 13 possible answer options.

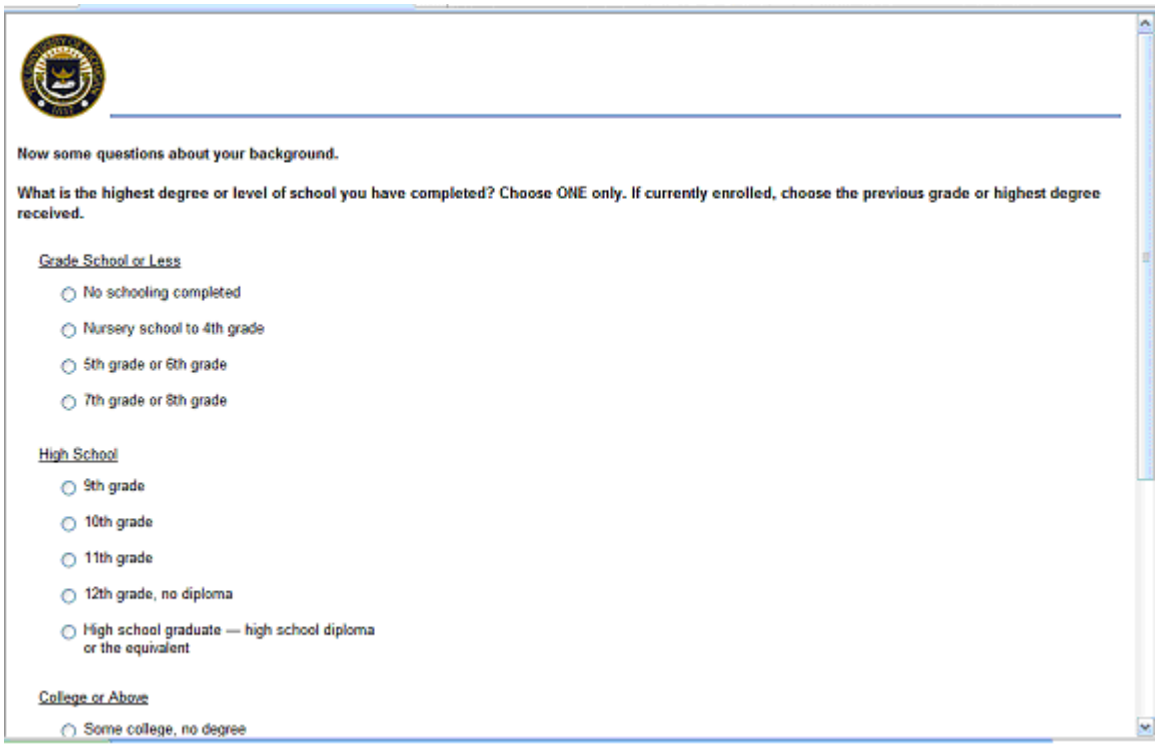
Respondents were randomly assigned to one of four conditions (displayed in Figure 1).

- Condition 1 presented the response options in one column in three conceptual groups that were separated by blank space. An underlined heading was placed before each group.
- Condition 2 presented the response options in one column, with the options separated into three conceptual groups by black space, but the groups did not have an underlined heading
- Condition 3 grouped the response options into three columns, each of which had an underlined heading.
- Condition 4 grouped the response options into three columns without any underlined headings.

Both the Education and Transportation Items are shown below for Condition 1. The Education Item is shown for the remaining conditions.

In the eye-movement study, respondents received both questions in the same condition.

**Condition 1, Education Item. One column separated into groups with underlined headings.**



**Condition 1, Transportation Item. One column separated into groups with underlined headings.**

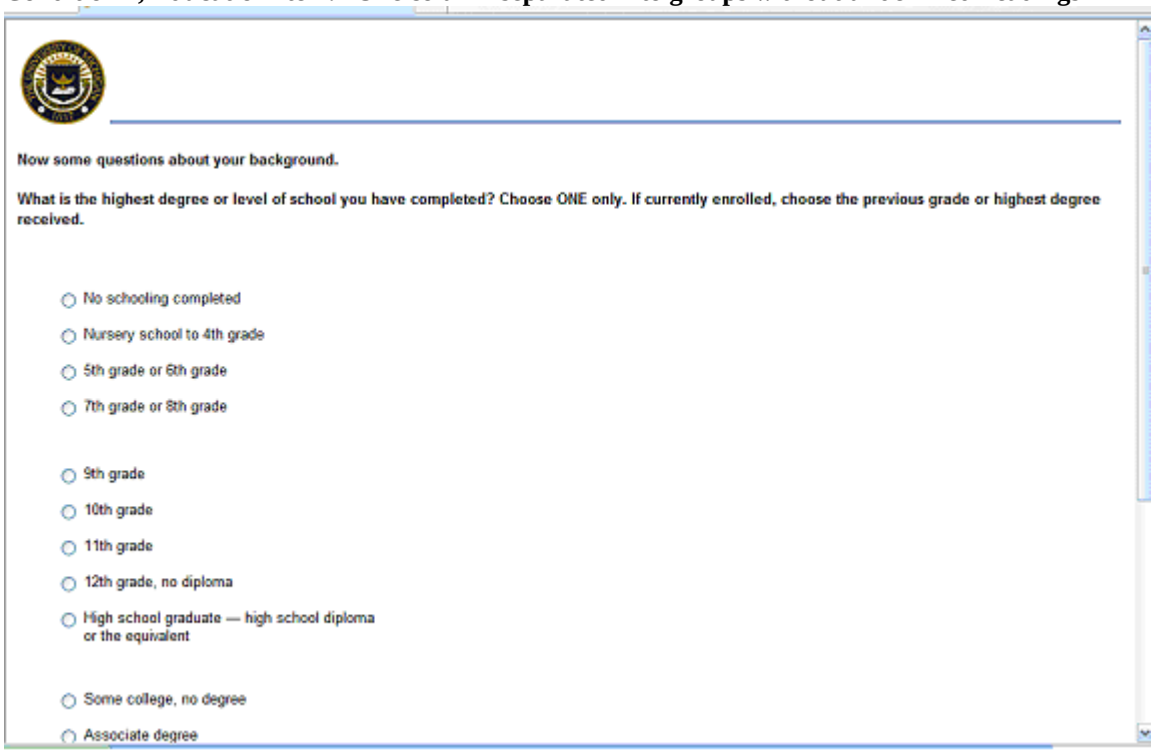
How did you usually get to work last week? If you usually used more than one method of transportation during the trip, choose the one used for most of the distance.

Own Vehicle
- Car
- Truck
- Van
- Motorcycle

Public Transportation
- Bus or trolley bus
- Streetcar or trolley
- Subway or elevated
- Railroad
- Ferryboat

Other
- Walked
- Bicycle
- Taxi or cab
- Other method

**Condition 2, Education Item. One column separated into groups without underlined headings**

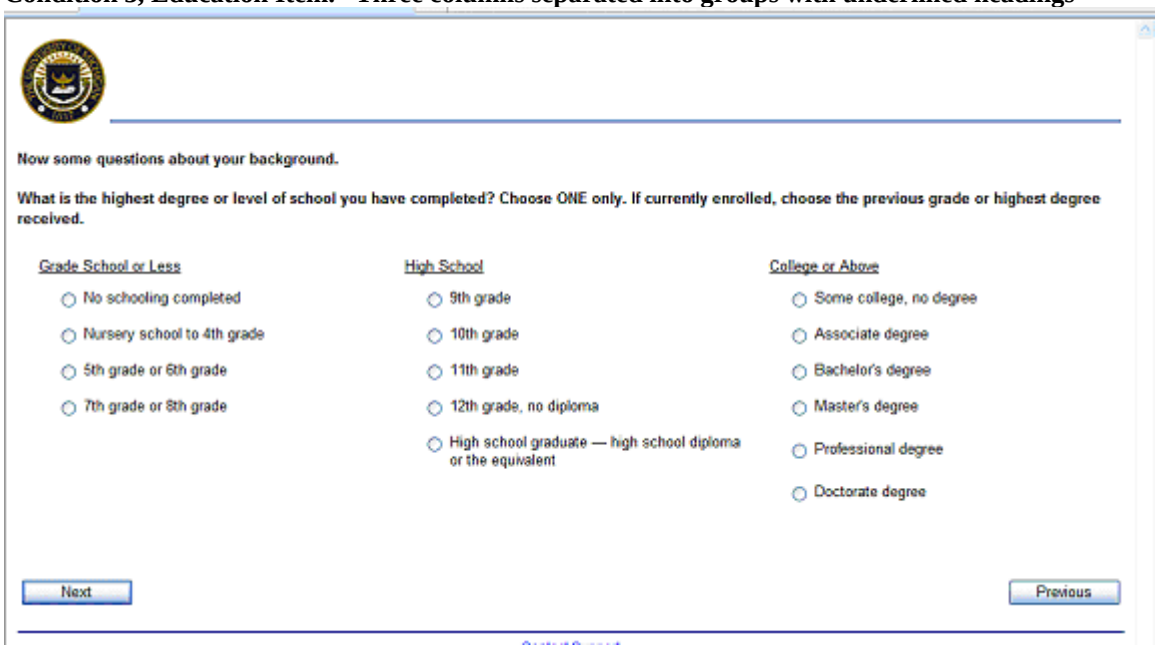


Now some questions about your background.

What is the highest degree or level of school you have completed? Choose ONE only. If currently enrolled, choose the previous grade or highest degree received.

- No schooling completed
- Nursery school to 4th grade
- 5th grade or 6th grade
- 7th grade or 8th grade
- 9th grade
- 10th grade
- 11th grade
- 12th grade, no diploma
- High school graduate — high school diploma or the equivalent
- Some college, no degree
- Associate degree

**Condition 3, Education Item. Three columns separated into groups with underlined headings**




**Condition 4, Education Item. Three columns separated into groups without underlined headings**

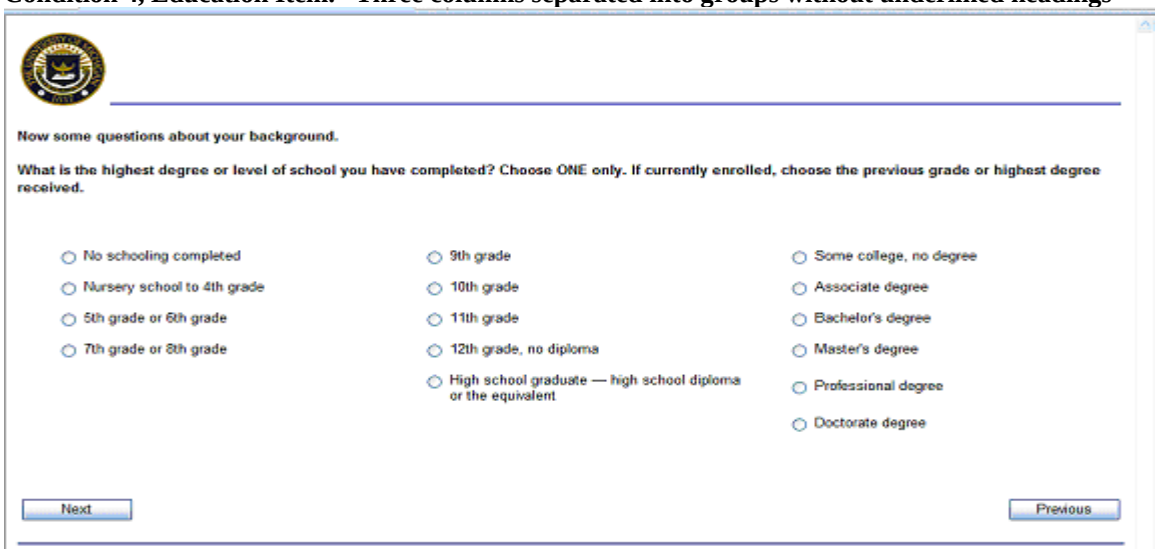



**Figure 1.** Screen shots of the four conditions for the education item, and one of the transportation item.

**Results**

As predicted, there were no statistically significant differences between the percentages of the responses in the sections of either item ($\chi^2$ (3, N = 155 ) = 2.34, p =.51 for education, $\chi^2$ (6, N = 97 ) = 1.15, p=.98 for transportation). A majority of respondents (83 percent) selected a response option from among the third group of response options in the education item, that is, from among the 'college or above' options. No one in the eye-movement study selected from among the first group of options "grade school or less" (hence, the degrees of freedom are lower for this item than for the transportation item). A smaller majority of respondents (59 percent) chose from among the first group of response options in the transportation item, that is, from among the group that included the respondent's 'own vehicle' options.

As Table 1 shows there was a significant difference in the mean total fixation duration times spent on each of the items by condition.

Table 1. Total fixation duration time (in seconds) spent on an item by condition.

| | Education | | Transportation | |
|---|---|---|---|---|
| Condition | Mean No. Seconds | (n) | Mean No. Seconds | (n) |
| 3 columns with headings | 6.3 | (39) | 3.5 | (20) |
| 3 columns without headings | 4.7 | (36) | 2.9 | (24) |
| 1 column with headings | 8.5 | (36) | 5.9 | (27) |
| 1 column without headings | 5.0 | (36) | 2.6 | (27) |
| F-value | F(3,146)=6.06*** | | F(3,97)=6.21*** | |

***<.001

Table 2 summarizes the overall trends in the data, after combining the fixation duration times in both items. Main effects for the headings were significant at the .001 level of the test. Respondents spent 4.7 seconds more on average fixating on both items when underlined headings were introduced. Main effects for the number of the columns and an interaction effect between the headings and number of columns were marginally significant at the .1 level of the test. Thus respondents spent marginally more time (2.5 seconds) fixating on both items when the response options were arrayed down a column. Furthermore, they spent marginally more time fixating on an item when headings were introduced in the one column condition than when headings were introduced into the three column condition (interaction effect).

Table 2. Total fixation duration time (in seconds) spent on both items by condition.

| | Both Items | |
|---|---|---|
| Condition | Mean No. Seconds | (n) |
| All Conditions | | |
| 3 columns with headings | 9.8 | (20) |
| 3 columns without headings | 7.8 | (24) |
| 1 column with headings | 14.5 | (27) |
| 1 column without headings | 7.8 | (26) |
| F-value | F(3,96)=6.12*** | |
| Headings | | |
| Headings | 12.5 | (47) |
| No Headings | 7.8 | (50) |
| F-value | F(1,96)=12.5*** | |
| Number of Columns | | |
| 1 column | 11.2 | (53) |
| 3 column | 8.7 | (22) |
| F-value | F(1,96)=2.81† | |
| Interaction Effect (Headings x Columns) | F(1,96) = 3.08† | |

***<.001, †<.1

One of the advantages of the conducting an eye-movement study is that we could remove the effects of looking at the headings from the effects of looking at the response options. After removing the time spent gazing at the headings, Table 3 shows that respondents still spent more time gazing at the 1 column with headings format that wasn't accounted for by looking at the headings.

Table 3. Total fixation duration time (in seconds) spent on an item by condition, less time spent on headings.

| | Education | | Transportation | |
|---|---|---|---|---|
| Condition | Mean No. Seconds | (n) | Mean No. Seconds | (n) |
| 3 columns with headings | 4.7 | (38) | 2.4 | (20) |
| 3 columns without headings | 4.7 | (36) | 2.9 | (24) |
| 1 column with headings | 7.2 | (36) | 5.0 | (27) |
| 1 column without headings | 5.0 | (36) | 2.6 | (27) |
| F-value | F(3, 145) = 4.12** | | F(3,97) = 4.62** | |

**<.01

A further advantage of conducting an eye-movement study is that we could also examine the scan paths of individuals to help us understand the summarized data better. We could view videos tracking respondents' eye movements, and get a sense of what respondents were looking at as well as the order in which respondents were looking at the information. Figure 2 presents a static snapshot of one respondent's scan path; it was selected to be representative of the problem we witnessed. As shown in Figure 2, respondents tended to spend more time when the response options were arrayed in one column with headings because they were confused. Some respondents had a tendency to circle between the question and the response options in this condition (especially between the first section and the question) in what appeared to be an effort to determine what the question was asking of them. They seemed to think they were supposed to select a response from each group of response options, and thus at least some of the additional amount of time spent when underlined headings were introduced or when the options were presented in one column appeared to be due to the additional time they spent clicking in these conditions. Clicking is denoted by x's in Figure 2

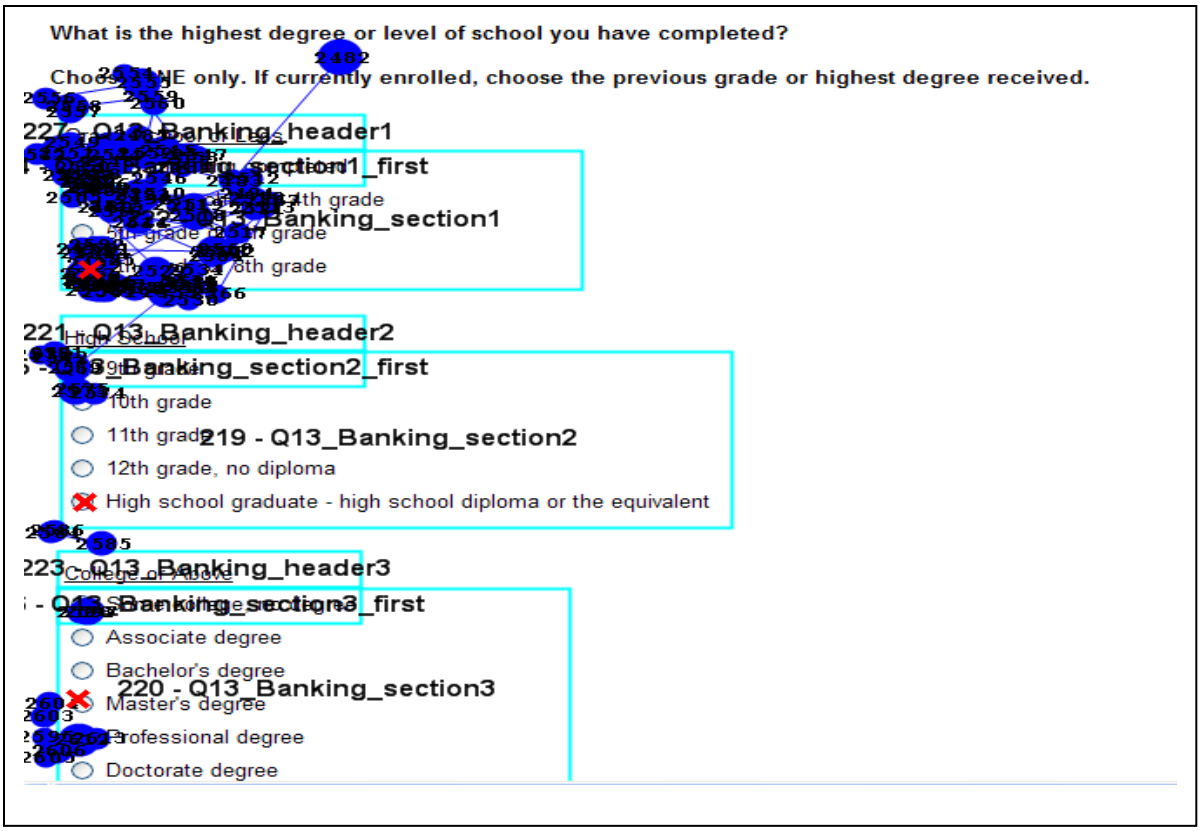



**Figure 2. Gaze plot illustrating a respondent's gaze points (as indicated by the filled circles) and clicking behavior (as indicated by the X's). It should be noted that the 'X' on the option '8th grade or above' is visible on screen, but may be obscured when printed out due to the large number of gaze points at this position.**

**Experiment 2: Opt-In Web Survey**

To confirm that respondents were attempting to click on more than one option and to ensure that the eye-movement findings generalized beyond the laboratory setting, we examined these issues in a larger field experiment.

**Methods**

This experiment was embedded in a Web survey that was conducted by Market Strategies, Inc. (MSI) from September 2, 2008 through September 23, 2008. One half of the sample for this survey was drawn from the Survey Sampling International (SSI) Survey Spot panel. This panel is an opt-in panel of over five million persons who have signed up on-line to receive survey invitations. The other half of the sample was drawn from a four-million member panel known as Authentic Response.

Respondents from both panels were invited using an email that asked them to participate in a study of attitudes and lifestyles that was being conducted by the University of Michigan. Each respondent was assigned a unique PIN to gain access to the survey. Once they completed the survey, Survey Spot panelists were given the opportunity to play an instant game, where they could win cash or gift cards from $50.00 to $250.00. They were also entered into a grand sweepstakes worth $25,000. Authentic Response panelists were paid $1.00 for completing the survey. Non-respondents received one follow-up email in both samples.

A total of 61,884 panel members were contacted by SSI and 1,200 completed the survey, for a 1.9 percent response rate. Authentic Response contacted 7,316 panel members, of whom 1210 completed the survey, for a 16.5 percent response rate. The overall response rate was 3.5 percent (AAPOR RR1).

This experiment included the same two questions in the eye-movement experiment and included the same four experimental conditions, except that, as shown in Figure 3, we added a control condition. The control condition presented the response options in one column vertically without any grouping or headings. This experiment differed from the eye-movement

experiment in that respondents answered 17 other questions between the education and transportation items. Respondents were assigned to conditions to the two items independently. This means that only a fraction of the respondents received both items in the same conditions.

**Control Condition. One column no separations—top of screen**

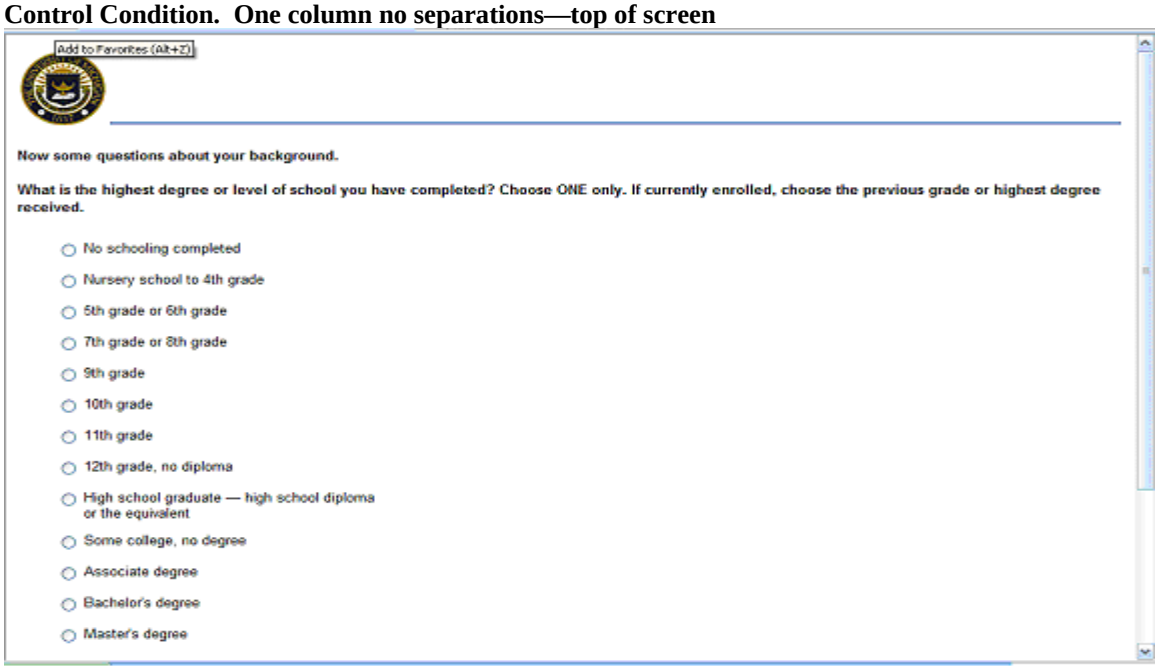



**Figure 3.** Screen shot of the control condition for the education item.

As before, we predicted that respondents' answers would not vary between the conditions, but that response times would. However, because of what we learned from the eye-movement study, we hypothesized that the number of clicks would differ between conditions. We thought that respondents would click more in the one-column condition and when headings were introduced.

**Results**

Again, the majority of respondents selected an answer from among the third group of response options in the education item (college or above) and from the first group of options in the transportation item (own vehicle). As with the eye-movement study, and as predicted, there were no statistically significant differences between the percentages of responses in the sections of either item ($\chi^2$ (8, N=2407)=5.94, p=.65 for the education item and $\chi^2$ (8, N=951) = 7.65, p=.47 for the education item). However, because the first group of response options in the education item had expected counts less than 5, rendering the $\chi^2$ potentially invalid, we collapsed the percentage of respondents who selected from the minority sections in both items and compared this percentage to the percentage choosing the majority section. Again, there were no statistical differences ($\chi^2$ (4, N = 2407) =3.06, p=0.55) in the education item and $\chi^2$ (4, N = 951 ) = 1.37, p=0.85 in the transportation item).

As can be seen in the 'All Conditions' panel of Table 4, the mean number of seconds a respondent spent on an item was statistically different across conditions. As suggested by Ratcliff (1993), we removed long response times from the analysis presented in Table 3,. As a result, we removed the slowest one percent of response times (24 observations with response times greater than or equal to 75 seconds in the education item and 9 observations with response times greater than 40 seconds in the transportation item).

The 'Headings' panel of the table confirms that respondents tended to spend almost 4 seconds longer on average on the education item when there were headings than when there were no headings, and almost 2 seconds longer on average on the transportation item when there were headings. As can be seen in the 3rd and 4th panels of Table 4, there was no difference in time spent between the one and three-column formats, and there was no interaction between the headings and number of columns (as there had been for the eye-movement study).

Respondents spent the least amount of time on the control condition. Respondents spent around 2.5 seconds less on the control condition in the education item ( $F(1,2384)=44.84$, $p<.0001$) and about a second less on the control condition in the education item ($F(1, 943)=5.14$, $p<.01$). This was the simplest of the formats: composed of one long column of response options with no conceptual groupings and no underlined headings.

Table 4. Mean number of seconds respondents spent on an item, by condition

| | Education | | Transportation | |
|---|---|---|---|---|
| Condition | Mean No. Seconds | (n) | Mean No. Seconds | (n) |
| All Conditions: | | | | |
| Control (1 column, no separations and no headings) | 12.4 | (481) | 9.6 | (188) |
| 3 columns with headings | 16.7 | (476) | 11.0 | (198) |
| 3 columns without headings | 13.1 | (480) | 9.5 | (175) |
| 1 column separated with headings | 17.1 | (471) | 12.0 | (199) |
| 1 column separated without headings | 14.2 | (477) | 10.1 | (184) |
| F-value | F(4,2384)=30.34**** | | F(4, 943)=6.21**** | |
| Headings: | | | | |
| Headings | 16.9 | (947) | 11.5 | (397) |
| No Headings | 13.2 | (1437) | 9.7 | (547) |
| F-value | F(1,2384)=108.2**** | | F(1,943)= 22.95**** | |
| Number of Columns: | | | | |
| 1 column | 14.5 | (1429) | 10.6 | (571) |
| 3 columns | 14.9 | (956) | 10.3 | (378) |
| F-value | F(1,2384)=1.43 | | F(1,943)=0.54 | |
| Interaction Effect (Headings*Columns) | F(1,2384) = 0.09 | | F(1,943) = 0.77 | |

*<.05, ****<.0001

Adjusted by removing the slowest one percent of response times (24 observations with response times greater than or equal to 75 seconds in the education item and 9 observations with response times greater than 40 seconds in the transportation item) see Ratcliff (1993).

Since the eye-movement analysis did not have the control condition, we removed the control condition from this analysis, and re-ran it. As shown in Table 5, the headings are still significant, but now the one-column conditions are marginally significantly different from the three- column conditions, and this difference is in the same direction as for the eye-movement study, that is, respondents take longer in the one-column conditions when the response options are grouped and separated compared to the three-column conditions. There is no interaction between the headings and the number of columns in the Web survey.

Table 5. Adjusted mean number of seconds spent on a condition, having removed the control condition from the analysis.

| | Education | | Transportation | |
|---|---|---|---|---|
| Condition | Mean No. Seconds | (n) | Mean No. Seconds | (n) |
| Headings: | | | | |
| Headings | 16.9 | (947) | 11.5 | (397) |
| No Headings | 13.7 | (957) | 9.8 | (359) |
| F-value | F(1,1903)= 66.85**** | | F(1,755)=15.38*** | |
| Number of Columns: | | | | |
| 1 column | 15.6 | (948) | 11.1 | (383) |
| 3 columns | 14.9 | (956) | 10.3 | (373) |
| F-value | F(1,1903)=2.88† | | F(1,755)=3.15† | |
| Interaction Effect | F(1,1903)=0.77 | | F(1,755)=0.28 | |

†<.1, ****<.0001

Adjusted by removing the slowest one percent of response times (24 observations with response times greater than or equal to 75 seconds in the education item and 9 observations with response times greater than 40 seconds in the transportation item) per Ratcliff (1993).

Table 6 confirms that the percent of respondents clicking more than once significantly differed by condition, and that this difference was clearly present when conditions with headings are compared to those without.

Table 6. Percent of respondents clicking more than once by each condition and item.

| | Education | | Transportation | |
|---|---|---|---|---|
| Condition | Percent Clicking >1 | (n) | Percent Clicking >1 | (n) |
| All conditions: | | | | |
| 1 column (no separations) | 3.3 % | (479) | 4.2 | (191) |
| 3 columns with headings | 19.6 | (479) | 15.7 | (198) |
| 3 columns without headings | 4.4 | (483) | 6.3 | (174) |
| 1 column separated with headings | 21.8 | (478) | 13.5 | (200) |
| 1 column separated without headings | 7.3 | (481) | 2.7 | (184) |
| $\chi^2$ (df=4) | 147.1**** | | 32.4**** | |
| With and without headings: | | | | |
| Headings | 20.7% | (957) | 14.6 | (398) |
| No Headings | 5.0 | (1440) | 4.4 | (564) |
| $\chi^2$ (df=4) | 142.1**** | | 30.4**** | |

****<.0001

**Discussion**

As predicted, respondents' final selections were not affected by the layout and labeling of the response options. This was in line with our original hypothesis, and lends evidence to the fact that these are factual questions that people tend to answer accurately, regardless of the kinds of design variations under examination here.

Originally, we were concerned that the eye-movement study might not have been powerful enough to detect differences in respondents answers to the questions because the sample size of this study was so small (n=155). In addition, we couldn't be sure that respondents who volunteered to take part in an eye-movement study and who were willing to come to the University were representative of general Web survey responders, which is one reason we followed the eye movement study with the larger scale Web survey. Since there were no differences in respondents' answers among the various conditions in either study—our concerns regarding the small sample size and the potential unrepresentativeness of the eye-movement sample were dispelled.

The fact that respondents were able to answer similarly across the various conditions in both studies, however, is not the same as saying that respondents were able to do so with equal efficiency. The paradata reveal underlying problems when the format included underlined headings. While either the eye-movement study or the Web survey by itself may have left some doubt as to why respondents took more time to answer when underlined headings were introduced, taken together, the two studies strongly suggest that respondents became confused when headings were introduced. The additional time it takes to answer a question when headings were introduced (Table 2) appeared to be due to spending time circling between the response options and the question (or at least parts of the question) (see Figure 2) and clicking more than once (refer to Table 5). The underlined headings seemed to persuade respondents to interpret the question as though it read: "What is the highest degree or level of school you have completed *in each of the following categories*: (1) *grade school or less, (2) high school, and (3) college or above?*" or "How did you usually get to work last week *in each of the following categories: (1) own vehicle, (2) public transportation, (3) and other?".*

This appears to be a false implicature (Grice, 1975). An implicature is something meant, implied, or suggested distinct from what is said. False implicatures occur when respondents extract unintended meanings from questions and or response categories because they presume the survey questions follow the Gricean maxims of being truthful, relevant, informative, and clear. It would seem that respondents falsely drew the conclusion that the headings were essential parts of the question, not just the response guideposts they were meant to be.

Some respondents were induced to select an answer from each of the three groups in the transportation item (as well as the education item), even though it is harder to see how respondents might re-interpret the transportation item as asking for a response in each category. Respondents could more easily select the highest level of schooling they completed in all three education groups —for example, a respondent could have completed "7$^{th}$ or 8$^{th}$ grade" in grade school and completed "High school graduate" in high school and completed "Some college, no degree" in college. Attempting to answer more than one section in the transportation item seems an even stronger indicator of confusion than attempting to do so in the education item. However, because the radio button design excluded their being able to select more than one option, respondents seemed to figure out within seconds that answering more than one section was not really the intention of either question, and they settled on an answer—the same answer they would have selected had the headings not misled them to select more than one.

The results of our study are consistent with Smyth, Dillman, Christian and Stern (2006), who found that when respondents were offered two groups of response options (with three response options per group), both of which had headings, respondents were more likely to choose at least one response option from each group of response options, and this was true whether the two groups of response options were arrayed vertically or horizontally. A difference between their study and ours is that they found actual differences in responses distributions between conditions, whereas we did not. This is likely due to their using check boxes and our using radio buttons. The radio buttons prevented respondents from selecting more than one response.

There were two outcomes between our first experiment (the eye-movement study) and our second (the Web survey) that did not replicate if we compare all conditions in both studies. For one, respondents in the eye-movement study were marginally slower when the response options were arrayed down one column as opposed to when they were presented in three columns (Table 2). We wondered if the control condition was responsible for this effect, as it wasn't one of the conditions in the eye movement study and it differed from the other 1-column conditions in that the response options were not grouped

conceptually and separated from one another. When we removed the control condition from the analysis, essentially limiting the analysis to only those conditions in which the response options were separated from one another and grouped, we found that respondents were marginally slower in the 1-column conditions in the Web survey too compared to the 3-column conditions (Table 5). The interaction between the headings and the number of columns, however, never did transpire in the Web survey, with or without the control as part of the analysis

**Conclusions and Future Research**

We thought that grouping the options conceptually and giving respondents headings would guide them to their preferred answer so they would answer the items more quickly. The headings, at least as they were designed here, led to confusion and appeared to induce a false implicature. We thought that grouping the options and arraying them in three columns would make the respondent's task easier, by eliminating the need to scroll. This did appear to help in the eye-movement study and was marginally successful in the Web survey too, if we remove the effects of the control condition (the one column condition that had no separations and no headings). In other words, respondents may be a little better off if the options are presented in three columns on one screen, if the plan is to separate a long list of response options into conceptual groups that would otherwise span two screens. This conclusion is based on the fact that respondents answered the 3-column format marginally faster than the one column format in both studies (Tables 2 and 5).

In the end, however, it seems respondents were able to answer the control format fastest, and with the least amount of double or triple clicking. Probably because this was the simplest format, it required the least amount of eye-movement and the least amount of effort. This condition was composed of one long column of response options that were not separated from each other in any way (no spacing between options and no underlined headings). This condition spanned two screens and necessitated scrolling. Nevertheless it outperformed the other conditions because it did not mislead respondents into thinking that they were supposed to provide more than one answer. Until it can be improved upon, this is the format that should be used.

This research is one more in a body of research to suggest that what appear to be helpful design features on the surface may actually harm respondent performance. This research also demonstrates that although our understanding has progressed, still there is much we do not know about how to design questions and responses. For example, can the underlined headings used here be redesigned in ways to aid respondents' selections without misleading them? These questions and more are left for future research.

**References**


Christian, L. and Dillman, D. 2004. The Influence of Graphical and Symbolic Language Manipulations on Response to Self-Administered Questions. *Public Opinion Quarterly*. 68(1): 57-80.

Couper, M. 2008. *Designing Effective Web Surveys*. Cambridge, NY: Cambridge University.

Conrad, F.G., Schober, M.F., and Coiner, T. 2007. Bringing Features of Human Dialogue to Web Surveys. *Applied Cognitive Psychology*, 21: 165-188.

Couper, M.P., Conrad, F.G., Tourangeau, R. Visual Context Effects in Web Surveys. *Public Opinion Quarterly*. 71(4): 623-634.

Couper, M.P., Tourangeau, R., and Kenyon, K. 2004. Picture this! Exploring Visual Effects in Web Surveys. *Public Opinion Quarterly*. 68(x): 255-266.

Couper, M.P. and Lyberg, L.E. 2005. The Use of Paradata in Survey Research. *Proceedings of the 54th Session of the International Statistical Institute*. Sydney, Australia.

Couper, M.P. 1998. Measuring Survey Quality in a CASIC Environment. *Proceeding of Research Methods Section American Statistical Association.* Pp. 41-49.

Galesic, M., Tourangeau, R., Couper, M. Conrad, F. 2008. Eye-Tracking Data: New Insights on Response Order Effects and Other Signs Of Cognitive Shortcuts In Survey Responding. *Public Opinion Quarterly*. 72(5): 892-913.

Grice, H.P. 1975. Logic and Conversation. In P.Cole and J.L. Morgan (Eds.*). Syntax and Semantics: 3. Speech Acts* (pp.41-58). New York: Academic Press.

Hartley, 1985. *Designing Instructional Text*. New York: Nichols.

Heerwegh, D. 2003. Explaining Response Latencies and Changing Answers Using Client-Side Paradata from a Web Survey. *Social Science Computer Review*. 21: 360-373

Hyona, J. and Nurminen, A. 2006. Do Adult Readers Know How They Read? Evidence from Eye Movement Patterns and Verbal Reports. *British Journal of Psychology*. 97: 31-50.

Jenkins, C. and Dillman, D. 1997. Towards a Theory of Self-Administered Questionnaire Design. In Lyberg, L., Biemer, P., Collins, M., DeLeeuw, E., Dippo, C., Schwarz, N., and Trewin, D. (Eds.) *Survey Measurement and Process Quality.* New York: John Wiley and Sons, Inc.

Mockovich, B. 2005. Comparing the Effectiveness of Alternative Approaches for Displaying Edit Error Messages in Web Forms. Paper Presented at FedCASIC, Bureau of Labor Statistics: Washington, DC.

Ratcliff, Roger (1993). Methods for Dealing with Reaction Time Outliers. *Psychological Bulletin*, 114 (3): 510-532.

Rayner, K. 1998. Eye-movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*. 124: 372-422.

Redline, C., Hunfalvay, M., and Lankford, C. 2006. Eye-movement Analysis of a Web Survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Redline, C. and Lankford, C. 2001. Eye-movement Analysis: a New Tool for Evaluating the Design of Visually Administered Instruments (paper and Web), *Proceedings of the Section on Survey Research* Methods, American Statistical Association.

Schwarz, N., Grayson, C.E., and Knauper, B. 1998. Formal Features of Rating Scales and the Intepretation of Question Meaning. *International Journal of Public Opinion Research*. 10:177-183.

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

Smyth, J.D, Dillman, D., Christian, L.M. and Stern, M.J. 2006. Effects of Using Visual Design Principle to Group Response Options in Web Surveys. *International Journal of Internet Science* 1(1): 6-16.

Sudman, S., Bradburn, N. and Schwarz, N. (1996) *Thinking About Answers: The Application Of Cognitive Processes To Survey Methodology*. San Francisco: Jossey-Bass Publishers.

Tourangeau, R. Conrad, F.G., Arens, Z., Fricker, S., Lee, S., and Smith, E. 2006. Everyday Concepts and Classification Errors: Judgements of Disability and Residence. *Journal of Official Statistics*, 22 (2): 385-418.

Tourangeau, R., Couper, M.P., Conrad, F.G., 2004. Spacing, Position, and Order: Interpretive Heuristics for Visual Features of Survey Questions. *Public Opinion Quarterly.* 68(3): 368-393.

Tourangeau, R., Couper, M.P., Conrad, F.G. 2006. Color, Labels, and Interpretive Heuristics for Response Scales. *Public Opinion Quarterly.* 71(1): 91-112

Yan, T., and Tourangeau, R. 2008. "Fast Times and Easy Questions: The Effects of Age, Experience and Question Complexity on Web Survey Response Times." *Applied Cognitive Psychology*, 22 (1): 51-68.

**Appendix A**

**Survey Invitation Email**

Subject Line: University of Michigan Lifestyles and Health Survey

Dear Survey Respondent,

We have a new opportunity for you to participate in a study of attitudes and lifestyles that is being conducted by the University of Michigan. There are questions about a variety of health, lifestyle and social topics. This is an important study being funded by the National Institutes of Health.

Reward: An Instant Win game play and an entry in the $25,000 sweepstakes.

Click this link to start: <field>LINK</field>

[If you are not able to click this address, copy the entire line and paste it into your browser's address line. If you encounter problems while taking the survey, please contact https://www.msisurvey.com/ContactSupport/default.aspx?script=G08003_1">.]

Thanks for your participation.