Words, Numbers and Visual Heuristics in Web Surveys

Toepoel, V.; Dillman, D.A.

Publication date:
2008

No. 2008–92

WORDS, NUMBERS AND VISUAL HEURISTICS IN WEB SURVEYS: IS THERE A HIERARCHY OF IMPORTANCE?

By Vera Toepoel, Don A. Dillmann

October 2008

ISSN 0924-7815

Words, Numbers and Visual heuristics in Web Surveys: Is there a Hierarchy of Importance?

Vera Toepoel[*] and Don A. Dillman[1]

Abstract In interpreting questions, respondents extract meaning from how the information in a questionnaire is shaped, spaced, and shaded. This makes it important to pay close attention to the arrangement of visual information on a questionnaire. Respondents follow simple heuristics in interpreting the visual features of questions. We carried out five experiments to investigate how the effect of visual heuristics affected the answers to survey questions. We varied verbal, numerical, and other visual cues such as color. In some instances the use of words helps overcome visual layout effects. In at least one instance, a fundamental difference in visual layout (violating the 'left and top means first' heuristic) influenced answers on top of word labels. This suggests that both visual and verbal languages are important. Yet sometimes one can override the other. To reduce the effect of visual cues, it is better to use fully labeled scales in survey questions.

Keywords: questionnaire design, layout, visual language, response effects, visual cues

JEL codes: C42, C81, C93

[*] CentERdata, Tilburg University, postal address: CentERdata, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Corresponding author: Vera Toepoel; e-mail: V.Toepoel@uvt.nl

[♦] Social and Economic Sciences Research Center, Washington State University, postal address: 133 Wilson Hall, PO Box 644014, Washington State University, US.

1. Introduction

Traditionally, it has been assumed that what survey questions mean to respondents depends almost entirely on words, a view fostered by the dominance of telephone interview surveys that depend on aural communication. However, considerable research has cast doubt on that view, suggesting instead that the visual layout of questionnaires also has a significant influence on how respondents answer questions, in interview surveys (Smith, 1995) as well as paper questionnaires (Jenkins and Dillman, 1997).

Considerable experimental research now shows that respondent behaviors in web as well as paper surveys are significantly influenced by a variety of visual aspects, including overall graphical composition, symbols and the use of numbers (e.g. Christian and Dillman, 2004; Redline and Dillman, 2002, Toepoel, Das, and Van Soest, 2006). In addition, Tourangeau, Couper, and Conrad (2004) argue that respondents follow simple heuristics in interpreting the visual features of questions:

- Middle means typical: respondents will see the middle option as the most typical
- Left and top means first: the leftmost or top option will be seen as the 'first' in conceptual sense
- Near means related: options that are physically near each other are expected to be related conceptually
- Up means good: the top option will be seen as the most desirable
- Like means close: visually similar options will be seen as closer conceptually.

These authors propose that each heuristic assigns a meaning to a visual cue, and they provide evidence for three of the heuristics that respondents follow when selecting answers for scalar questions.

These authors (2007) also suggest that there may be a hierarchy of features that respondents attend to, with verbal labels taking precedence over numerical labels and numerical labels taking precedence over purely visual cues. The impact of the nonverbal cues is conjectured to be greater when the verbal labels leave more uncertainty in the question or answer categories. This suggests that respondents search for additional cues when the verbal language is not clear. By using verbal language effectively, the effect of visual heuristics may be diminished.

The purpose of this paper is to investigate whether the effects of visual layout in scalar questions can be diminished through more extensive use of verbal language and numbers assigned to words. More specifically, we address the question whether there is a hierarchy of features that respondents attend to, with verbal labels taking precedence over numerical labels and numerical labels taking precedence over visual cues. Each of the five visual heuristics is tested on verbal, numerical, and visual language, to determine whether effective question writing reduces the effect of visual language.

2. Background

Respondents are cooperative communicators and will process survey questions by drawing on all information provided by the survey researcher (Schwarz, 1996). They find cues in the verbal language (words) in the questionnaire, but in the nonverbal language as well. Nonverbal cues include graphical, numerical and symbolic languages that convey meaning in addition to the verbal language (Christian and Dillman, 2004). The visual layout of a question is an important source of information that respondents use when deciding which answer to select (Christian, 2003). A conceptual framework to explain how visual languages may influence respondent

behavior has been provided by Jenkins and Dillman (1997). Considerable research shows that independent and combined effects of visual manipulations affect respondents' answer choices (see e.g. Christian and Dillman, 2004; Christian, Dillman, and Smyth, 2005; Jenkins and Dillman, 1997; Redline and Dillman, 2002; Redline et al., 2003; Toepoel, Das, and Van Soest, 2006; 2009, Forthcoming; Tourangeau et al., 2004; 2007).

The act of responding to a question (the question-answering process) contains basically four steps: comprehending the question, retrieving information, generating an opinion or a representation of the relevant behavior, and reporting it (Tourangeau, 1987). Mistakes can be made because of problems at any one of these steps, resulting in measurement error. During the response process, respondents draw information from the visual layout of questions. Tourangeau et al. (2004) argue that respondents follow simple heuristics in interpreting the visual features of questions.

The 'middle means typical' heuristic states that respondents see the visual midpoint of a scale as representing the typical or middle response. Tourangeau et al. (2004) show that when the answer categories for an item are unevenly spaced, so that the conceptual midpoint for an item does not coincide with the visual midpoint, the response option that was displaced toward the visual midpoint was selected more often. They also demonstrate that when the non-substantive options ("Don't know" and "No opinion") are not visually distinct from the scale points, respondents use the visual midpoint of the scale as a reference point and are more likely to select answers from that conceptual end of the scale. Christian and Dillman (2004) find an effect of equal versus unequal spacing between response categories for nominal questions as well. When one category was more widely separated from the remaining three categories, respondents were more likely to choose that response category. The

researchers concluded that unequal spacing can increase the visual prominence of the isolated category. Research by Christian, Parsons and Dillman (forthcoming) shows that varying the distance between categories in a symmetrical fashion, i.e. providing varied distances between adjacent categories on both sides of the midpoint, does not change respondent answers.

The 'left and top means first' heuristic states that respondents see response options in a logical order, with the first item also being first in a conceptual sense. Tourangeau et al. (2004) find that respondents answer questions most quickly when items follow a logical progression in line with the heuristic. A rearrangement of response options also affects the distribution of responses. Respondents overlook the unconventional order and select an option based on the order they expected.

The heuristic 'near means related' suggests that placing items near each other affects the answering process. Items are more likely to be viewed as related if grouped on one screen, reflecting a natural assumption that blocks of questions bear on related issues, much as they do during ordinary conversation (Schwarz and Sudman, 1996; Sudman, Bradburn, and Schwarz, 1996). Couper, Traugott, and Lamias (2001) and Toepoel et al. (2009) conclude that correlations are consistently higher among items appearing together on a screen than among items distributed across several screens, but the effect they find is small and differences between pairs of correlations are insignificant. Tourangeau et al. (2004) find significant differences between correlations, however. Peytchev et al. (2006) find little evidence that grouping affects respondents' answers.

The 'up means good' heuristic states that people expect categories to appear from positive to negative, i.e., that the first response option in a list is the most desirable. Both the meaning of the verbal label and its position on the scale can

influence the appraisal made by respondents (Hofmans et al., 2007). Research on orientation effects in rating scales yield inconsistent results. While in some studies respondents altered their responses when the orientation of a scale changed, in other studies responses remained unaffected (Weng and Cheng, 2000). For example, Christian et al. (forthcoming) show that if one consistently arranges all categories from negative to positive, this does not produce different answers than when arranged positive to negative. However, Toepoel et al. (2006; forthcoming) find differences in a decremental scale compared to an incremental scale.

People view options as conceptually closer when they are similar in appearance, compared to when they are dissimilar in appearance ('like means close heuristic'). Schwarz et al. (1991) demonstrate this by adding numbers ranging from -5 to +5 or 0 to 11 to verbal labels. This results in lower scores for the 0 to 11 version, compared to the –5 to 5 format; respondents apparently hesitate to assign a negative score to themselves (scale label effect, see Tourangeau, Rips, and Ransinski, 2000, p.248). To the respondents, the ends of the scale seemed conceptually further apart when the numerical labels differed both in sign and value than when they differed only in value. The principle that dimensions are seen as differing more sharply when they vary along two dimensions than when they differ only in a single dimension, also applies to the shading of response options. Tourangeau et al. (2007) demonstrate that when the verbal and numerical labeling of scale points provide respondents with minimal interpretive help, different colors for the endpoints of the scale have a noticeable effect on answers to the questions.

The results of Tourangeau et al. (2007) suggest that there is a hierarchy of features that respondents attend to. They demonstrate that only when the verbal and numerical labeling of scale points provide minimal help to respondents, the shading of

the response options has a noticeable effect on answers to the questions. In addition, the effects of the numerical labels are consistently larger than the effects of the shading of response options.

Visual design theory hardly makes any reference to respondent characteristics (Dillman, 2007), and few empirical studies have analyzed how the effects of questionnaire format vary with respondent characteristics. Tourangeau et al. (2007) find no systematic variation in the impact of the layout of a response scale in relation to gender, age or education group. Stern, Dillman, and Smyth (2007) also show that the layout of survey questions affects different demographic groups in similar ways. Knauper, Schwarz and Park (2004) and Borgers, Hox and Sikkel (2004), on the other hand, find that the generally poorer memory of older adults results in increased design effects. Deutskens et al. (2004), Dillman, Caldwell and Gansemer (2000), and Stern et al. (2007) all conclude that further research is needed on the effects of questionnaire format for different populations.

3. Design and implementation

To determine whether there is a hierarchy of features that respondents attend to when answering survey questions, five experiments were conducted on visual heuristics in a five-point rating scale (horizontally orientated). The experimental treatments varied verbal, numerical, and visual language. All experiments were fielded in the CentERpanel (see also http://www.centerdata.nl/en/CentERpanel), that has existed for 18 years. Panel members fill out questionnaires every week; our experiments were conducted in week 27 to 31, 2008. Respondents' panel duration ranges from 18 years to a few months. Although the CentERpanel is an Internet-based panel, there is no need to have a personal computer with an Internet connection. If necessary,

equipment is provided by CentERdata. The recruitment of new panel members is done through a random sample of landline numbers of candidates. If a household drops out of the panel, a new household is selected from a database of potential panel members on the basis of demographic characteristics. The panel is designed to be representative of the Dutch population.

The first experiment tested the 'middle means typical' heuristic; 2167 respondents were selected and 1541 responded (71%). Five questions were put to respondents about their satisfaction in terms of living in Europe, in the Netherlands, in their town, their neighborhood, and their street (with answer options ranging from "totally disagree" to "totally agree"), testing the effect of uneven spacing of response options. Respondents were randomly assigned to eight different formats in a 2*2*2 design, varying the response options with a polar point or fully labeled scale, the adding of numbers (1 to 5[2]) or not, and uneven spacing of response options or not. Uneven spacing was accomplished by inserting greater distance between the two right-most categories as shown in Figure 1. The effect of this change was to shift the midpoint somewhat to the left of a mid-distance location. We hypothesized that the uneven spacing of response options would result in different answers in the polar point format, but that the effect would decrease when numbers are added to the polar point format, and would disappear in the fully labeled format. We added numbers to the fully labeled format to complete the design and to determine whether numbers help to interpret response options in addition to the verbal labels.

[insert figure 1]

[2] In all experiments we used numbers 1 to 5.

The second experiment tested the 'left and top means first' heuristic; 2038 respondents were selected and 1375 responded (67%). Five questions were asked about verbalism[3], testing the effect of the ordering of response options. Respondents were again randomly assigned to eight different formats: (1) fully labeled, (2) fully labeled with numbers, (3) fully labeled with color[4], (4) fully labeled with numbers and color, (5) fully labeled with inconsistent ordering, (6) fully labeled with numbers and inconsistent ordering, (7) fully labeled with color and inconsistent ordering, (8) fully labeled with numbers and color and inconsistent ordering. Figure 2 shows the inconsistent ordering of response options (format 7, fully labeled with color). We hypothesized that the inconsistent order of response options (not in a logical progression) would result in different answers, but that the effect would decrease if colors are added to the verbal labels, and would decrease even further when numbers are added to the verbal labels, with the lowest effect of an inconsistent scale when both numbers and color are added to the verbal labels. Note that all formats are fully labeled and that, in this experiment, numbers and colors are expected to help respondents interpret the response options, in other words that visual language will help deal with ambiguous verbal cues.

[insert figure 2]

The third experiment tested the 'near means related' heuristic; 2054 respondents were selected and 1375 responded (67%). Five questions were asked about visualization[5]. Respondents were randomly assigned to eight different formats in a 2*2*2 design,

[3] See Appendix A for the questions used.
[4] In all experimental treatments with color, shadings of green were used for positive answers and shadings of red for negative answers.
[5] See Appendix A for the questions used.

varying the response options with a polar point or fully labeled scale, the adding of numbers or not, and the placement of the five items on separate screens or a single screen. Figure 3 shows the fully labeled with numbers single screen format (format 8). We hypothesized that the placement of all five questions on a single screen would affect respondents' answers most in the polar point format, followed by the polar point with numbers format. We expected no effect in the fully labeled formats (with and without numbers).

[insert figure 3]

The fourth experiment tested the 'up means good' heuristic; 2145 respondents were selected and 1688 responded (79%). Five questions were asked about their satisfaction in terms of the functioning of the United Nations, European Parliament, Dutch Parliament, the Dutch education system and the Dutch health care system (answer options ranging from "totally dissatisfied" to "totally satisfied"). Respondents were randomly assigned to eight different formats in a 2*2*2 design, varying the response options with a polar point or fully labeled scale, the adding of numbers or not, and the placement of the five items in an incremental or decremental order. Figure 4 shows the polar point with numbers decremental format (format 6). We hypothesized that orientation effects would be largest in the polar point format, followed by the polar point with numbers format. We expected no effect in the fully labeled formats (with and without numbers). Note that we used a horizontally orientated scale instead of a vertically orientated scale. People read top to bottom and left to right, therefore we conjectured that the effect of 'up means good' is the same as 'left means good'.

[insert figure 4]

The fifth experiment tested the 'like means close' heuristic; 2004 respondents were selected and 1340 responded (67%). Five questions were put to respondents about their satisfaction in terms of themselves, their family, their friends, their neighbors, and people they meet in ordinary life (answer options ranging from “totally dissatisfied” to “totally satisfied”), testing the effect of different numerical and visual labels. Respondents were again randomly assigned to eight different formats: (1) polar point, (2) polar point with numbers (1 to 5), (3) fully labeled, (4) fully labeled with numbers (1 to 5), (5) polar point with different numbers (-2 to 2), (6) fully labeled with different numbers (-2 to 2), (7) polar point with color, (8) fully labeled with color. Figure 5 shows the polar point with different numbers format (format 5). We hypothesized that the adding of negative signs and color would affect the polar point format but not the fully labeled format.

[insert figure 5]

We are aware that the layout manipulations might interact with the formats that respondents have dealt with in previous questionnaires. CentERdata tries to keep the format they use in line with previous research on the topic of the questionnaire. Since panel respondents complete questionnaires every week, they are confronted with all kinds of formats. Items are grouped together based on construct, on several screens to reduce the use of scrolling, but items are also presented one-item-per-screen if this is convenient in a particular questionnaire. Various numbers are added (CentERdata

generally uses the same numbers as in previous research related to the topic of the questionnaire, for example if a specific personality scale commonly attaches numbers 2 to -2 to the verbal labels, these numbers are used, whereas if in another scale no numbers are added in literature, CentERdata also does not add numbers). The target items were a self-contained questionnaire, so no items preceded them. An incremental scale is used more commonly in Dutch-speaking countries (Hofmans et al., 2007), so we used an incremental scale as reference level.

4. Results

4.1 Middle means typical heuristic

To analyze the effects of the experimental treatments on the response distributions, Chi Square Tests were conducted for each of the five questions. Table 1 shows that there were significant differences for the polar point format without numbers for all five questions. When numbers were added to the polar point format and with fully labeled scales, the uneven spacing of response options did not have a significant effect on responses.

The difference in the polar point format between the even and uneven spacing resulted in selecting the fourth option less and the fifth response option more often in the uneven format (for example, in Q1 40.6% selected the fourth option and 47.2% selected the fifth option in the even format, compared to 30.6% and 58.1% in the uneven format, respectively). Our results do not show that the visual midpoint was selected more often, but rather that respondents were more likely to choose the visually most distinct category (option 5), as demonstrated by Christian and Dillman (2004) in nominal questions. Our insignificant results for the formats with verbal and

numerical cues suggest that the effect of visual variations may be eliminated when verbal and numerical languages are added. ANOVA showed that the response times for the fully labeled scales were significantly longer compared to the polar point scales[6] ($F(df,7)=7.02$, $p<.01$), which indicates that respondents were attending more carefully to details of the layout.

[insert table 1 around here]

Table 2 presents ordinal regression analyses in which six dummy-coded variables were included. The first three variables represent the main effects of the experimental treatments: whether a fully labeled scale was used (as opposed to a polar point scale), whether numerical language was used, and whether the 'middle means typical' heuristic was violated (uneven spacing of response options). The last three variables represent the interaction effects (multiplication of two main effect variables): the interaction between verbal and numerical language, verbal and visual language, and numerical and visual language. The results confirm that verbal language is more powerful than numerical and visual language: whether or not a fully labeled scale is used is significant in all five questions. The adding of numbers is only significant in Q2, while the violation of the visual heuristic 'up means good' by an uneven spacing of response options is significant in Q1, Q2, and Q3. The adding of verbal labels, numerical language, and the visual separation of response options all lead to significantly more positive answers. A closer look at the interaction effects shows that a combination of numerical and visual cues has a converse effect on the responses. Both in question 2 and 4 the combination of numerical language and

[6] Four cases were deleted with response times longer than 2 times the standard deviation.

uneven spacing makes the responses more negative. It can be noted that all interaction effects between numerical language and uneven spacing of response options are negative, but only two out of five reach statistical significance.

[insert table 2 around here]

Since we use a panel designed to be representative of the Dutch population, we can also look at effects of personal characteristics. We computed 3 dummy variables for gender (1=man), age (1=50 years and older), and education (1=high). In addition, interaction terms of these three dummies with three dummies for verbal, numerical, and visual language (resulting in 9 interaction effects) were computed. In four out of five questions we found a significant interaction between age and verbal language (Q1 B=.50, p<.01, Q2 B=.39, p<.05, Q3 B=.33, p<.05, Q5 B=.43, p<.01): older respondents selected more positive answers when a scale was fully labeled. We also found a significant interaction effect for verbal language and gender (Q3 B=.59, p<.01, Q4 B=.43, p<.05): women also selected more positive answers in a fully labeled scale.

In short, this experiment shows that verbal language takes precedence over non-verbal cues such as numbers and spacing. Only when verbal and numerical language is absent or ambiguous does an uneven spacing of response options appear to have an effect on respondents' answers.

4.2 Left and top means first heuristic

According to this heuristic, the leftmost or top item in a list should be the 'first' in some conceptual sense, and the remaining options should follow in some logical

progression. To test whether respondents expect response options to follow in a logical progression, the placement of response options was changed in the experimental treatment. Instead of the response order “totally disagree”, “disagree”, “neither disagree nor agree”, “agree”, “totally agree”, we used the inconsistent order “totally disagree”, “totally agree”, “disagree”, “agree”, “neither disagree nor agree”. Note that in this experiment all response options were verbally labeled. We used numbers, colors, and a combination of numbers and colors to determine whether nonverbal cues help respondents order the response options. Table 3 shows that the comparison of the consistent and inconsistent scale resulted in significant differences in 18 out of 20 comparisons. Differences between the consistent and inconsistent scale basically result from respondents selecting the response option “disagree” more often in the inconsistent format. In the inconsistent format, this option is presented as the third alternative as opposed to the second alternative in the reference format. This indicates that respondents use the 'middle means typical' heuristic and view the third option as being the middle response. The response option “neither disagree nor agree” was selected less often in the inconsistent format. In this format the option was presented last (fifth), suggesting that respondents may shortcut the response process to select the first satisficing answer instead of processing each option individually (cf. Krosnick, Narayan and Smith, 1996). This may be due to the fact that we used a panel with trained respondents: Toepoel et al. (forthcoming) show that trained respondents are more sensitive to satisficing compared to untrained respondents. Only in Q5, the formats with color did not show significant differences between response distributions; color helped respondents order the response scale in the last question. This hints at a learning effect. Response times show that the inconsistent formats took

significantly longer compared to the consistent (reference) formats[7] (ANOVA, $F(df,7)=10.43$, $p<.01$). This is in line with the results of Tourangeau et al. (2004).

[insert table 3 around here]

Table 4 presents the ordinal regression analyses which again included six dummy-coded variables. The first three variables represent the main effects of the experimental treatments: whether color (visual language) was used, whether numerical language was used, and whether the 'top means first' heuristic was violated (inconsistent ordering of response categories). The last three variables represent the interaction effects (multiplication of two main effect variables): the interaction between colors and numbers, colors and inconsistent ordering of response categories, and numbers and inconsistent ordering. The results show that respondents more often select positive answers when these answers are colored green compared to no adding of colors. Also, the inconsistent arranging of response options causes more negative answers in two out of five questions. We find a significant interaction between color and the 'left and top means first' heuristic. Numbers did not significantly help the respondents interpret the verbal labels.

[insert table 4 around here]

Again we looked at effects of personal characteristics. We computed 3 dummy variables[8] for gender, age, and education, and interaction terms of these three dummies with three dummies for color, numbers, and inconsistency (resulting in 9

[7] Six cases were deleted with response times longer than 2 times the standard deviation.
[8] See experiment 1 for the cut-off levels that were used.

interaction effects). A significant interaction effect was found for age and color in Q1 and Q2 (B=.47, p<.01, B=.32, p<.05, respectively): older respondents react more strongly to the adding of color. Older respondents were also more sensitive to the inconsistent ordering of response options (Q2 B=-.43, p<.01, Q3 B=-.32 p<.01, Q4 B=-1.76, p<.01), and women were more sensitive to the inconsistent ordering compared to men (Q2 B=-.148, p<.01, Q4 B=-1.84, p<.01). Respondents with high education showed stronger ordering effects compared to respondents with middle or low education (Q2 B=-2.22, p<.01, Q3 B=-.46, p<.01, Q4 B=-3.41, p<.01).

In short, this experiment shows that respondents expect response categories to follow in a logical progression ('left and top means first'). Additional visual cues such as numbers and color do not overcome the fundamental problem of the original non-logical visual layout. Words (fully labeled scale) do not overcome it either, in contrast to what was found in the first experiment. Our results suggest that some aspects of visual layout are more powerful than words and also more powerful than words plus numbers and color. In this case a visual layout appears to be higher than words in the hierarchy of influence on respondent behavior.

4.3 Near means related heuristic

In this experiment we tested whether the placement of five questions on a single screen instead of on separate screens causes differences in responses and whether this is influenced by verbal and numerical language. Chi Square Tests (Table 5) showed significant differences for only one question (Q5): the polar point format without numbers showed significant differences between a one versus multiple items per screen format (χ^2 (df,4)=14.83, p<.01, N=315).

Reliability analysis shows that inter-item correlations (Cronbach's alpha) on the single screen format are higher in the polar point format compared to the one-item-per-screen format (alpha=.483, N=153 in the one-item-per-screen format and alpha=.631, N=162 in the single screen format). The difference between one and multiple items per screen format becomes smaller if numbers are added to the polar point format (alpha=.557, N=175 in the one-item-per-screen format and alpha=.640, N=175 in the single screen format). In the fully labeled formats, we did not find higher inter-item correlations when the items are placed on a single screen (alpha=.567, N=174 in the one-item-per-screen format and alpha=.567, N=191 in the single screen format for the fully labeled scale, and alpha=.634, N=174 in the one-item-per-screen format and alpha=.605, N=171 in the single screen format for the fully labeled scale with numbers). This finding offers additional evidence that respondents are more likely to use visual language when verbal and numerical labels provide minimal support. Our hypothesis is supported: the placement of items near each other produces the largest inter-item correlations in the polar point format. The difference in inter-item correlations becomes smaller when numbers are added to the polar point format. In the fully labeled formats, the effect of the 'near means related' heuristic is diminished. ANOVA shows that response times for the single screen formats are consistently lower compared to the one-item-per-screen formats[9] ($F(df,7)=5.09$, $p<.01$). In addition, the polar point format took less time to process compared to the other three formats.

[insert table 5 around here]

[9] Seven cases were deleted with response times longer than 2 times the standard deviation.

Table 6 shows ordinal regression analyses (with the same dummies as used in experiment 1). This revealed again that the fully labeled scales produced more positive answers; differences were significant in Q1 (B=.60, p<.01) and Q5 (B=.35, p<.01). The adding of numbers did not reveal significant differences. A closer look at the interaction effect between design and personal characteristics shows a significant interaction effect between age and verbal labels: in Q1 (B=.501, p<.01) and Q4 (B=.460, p<.01), older respondents selected positive answers in a fully labeled format more often than their younger counterparts.

[insert table 6 around here]

In sum, this experiment provides evidence that the effect of visual heuristics becomes smaller when numbers are added to a polar point scale and the effect disappears in a fully labeled scale. This finding may provide an explanation for the differences in results between Tourangeau et al. (2004) and Couper et al. (2001) and Toepoel et al. (2009): Tourangeau et al. used a polar point format and found significant differences between one and multiple items per screen formats, while Couper et al. and Toepoel et al. used a fully labeled scale and found no significant differences.

4.4 Up means good heuristic

In this experiment we compared an incremental scale to a decremental scale, again in 4 different versions: polar point, polar point with numbers, fully labeled, and fully labeled with numbers. Note that the response options were horizontally aligned. We found only one significant difference, strikingly, in the fully labeled with numbers

version (see Table 7; Q3 χ^2 (df=4)=10.03, p<.05, N=411). Response times did not vary between formats[10] (F(df,7)=.67, p=.70).

[insert table 7 around here]

Ordinal regression computed in the same way as in the previous experiments showed a significant positive effect of verbal labels (see Table 8; Q1 B=.543, p<.01, Q2 B=.415, p<.01). In Q3 we found a significant effect of violating the 'up means good' heuristic (B=.295, p<.01): the use of a decremental scale produced more positive answers. Respondents more often selected one of the first options, which were the positive verbal labels. This indicates that respondents select the first satisficing response option instead of processing each option individually and also that a positive tone of the first option changes reports in a positive manner (anchoring effect, as suggested by Schwarz, 1996). There was also a significant interaction effect in Q3 between the violation of the heuristic and the adding of numbers (B=-.462, p<.01): the numbers 5 to 1 produced more negative answers in a decremental scale. This indicates that numbers helped respondents interpret the question.

[insert table 8 around here]

Ordinal regression on personal characteristics shows that women more often selected negative answers when numbers were added to the verbal labels (Q2 B=-.350, p<.05, Q3 B=-.374, p<.05). They also reported more negative answers in a decremental scale (Q3 B=-.389, p<.05). Respondents in the age of fifty and older

[10] Five cases were deleted with response times longer than 2 times the standard deviation.

selected more positive answers in a decremental scale (Q2 B=.334, p<.05). They furthermore selected more negative answers in a fully labeled scale (Q3 B=-.324, p<.05). Respondents with high education more often selected positive answers in a fully labeled scale (Q2 B=.315, p<.05).

Overall, we found little evidence for the 'up means good' heuristic. This supports the hypothesis that if one consistently places all categories negative to positive, this does not produce different answers than running from positive to negative. In addition, Dutch respondents are generally more used to incremental scales. Although their interpretation of heuristics might suggest that the first option is the 'best' in conceptual sense, their experience may make them less sensitive to violations of the 'up means good' heuristic. Our use of a horizontal format could also be the reason for the lack of significant differences when the heuristic 'up means good' is violated.

4.5 Like means close heuristic

In this experiment we again used a polar point scale and a fully labeled scale as reference levels. We added different numbers (-2 to 2 compared to 1 to 5) and color (shadings of red and green) to test whether respondents see the ends of the scale as further apart conceptually when the numerical labels differed both in sign and value and the shading of response options differed in color.

Table 9 shows that the adding of numbers -2 to 2 compared to numbers 1 to5 resulted in different response distributions for both the polar point and fully labeled scale. Negative numbers might be interpreted as implying more extreme judgments than low positive numbers (scale label effect, see Tourangeau et al., 2000, p.248; see also Tourangeau et al., 2007, who make a similar argument and provide additional

evidence for the added attention that negative signs receive). This suggests that respondents interpret the verbal labels differently when negative signs are added. This could explain the fact that the adding of numbers -2 to 2 also affected the fully labeled scale. The adding of color only affected the polar point scale (all five questions showed significant differences; two at $p<.05$ and the other at $p<.10$). The answers to the fully labeled scale were not affected by the adding of color; respondents did not search for additional visual cues in interpreting the answer options. Response times did not vary between formats[11] (ANOVA, $F(df,7)=1.15$, $p=33$).

[insert table 9 around here]

Table 10 shows the results of the ordinal regression. We again found a strong positive effect of verbal labels. Although color and numbers did not show significant results individually, the dummy indicating that different endpoints were used (color or numbers -2 to 2) showed a significant positive effect in two out of five questions: respondents more often selected a response option with a green color or a positive number (when numbers -2 to 2 were added). Remarkably, the interaction effect between verbal labels and different endpoints shows a negative effect, indicating that the positive effect of different endpoints is overruled by a fully labeled scale (this applies to all five questions, although the difference only reaches statistical significance in two questions).

[insert table 10 around here]

[11] Three cases were deleted with response times longer than 2 times the standard deviation.

Ordinal regression on personal characteristics shows that older respondents (fifty years and older) selected positive answers in a fully labeled scale more often than their younger counterparts (Q1 B=.682, p<.01, Q2 B=.477, p<.01, Q3 B=.691, p<.01, Q4 B=..536, p<.01). Older respondents also reacted more strongly to different endpoints (Q1 B=.384, p<.05, Q2 B=.360, p<.05, Q3 B=.413, p<.05, Q4 B=.458, p<.01). In Q1 we found a significant interaction between education and verbal labels: respondents with high education more often selected positive answers in a fully labeled scale (B=.512, p<.01). In Q4 women reacted more strongly to different endpoints than men (B=.350, p<.05).

In short, our results support the heuristic. The effect of different numerical endpoints (both in sign as in value) and endpoints with different colors show that respondents use the 'like means close' heuristic. Color was only used by respondents when no numbers or verbal labels were presented, while different numbers also affected the fully labeled scale. This suggests that numerical language takes precedence over visual language when negative numbers are added. This also explains the hierarchy found by Tourangeau et al. (2007); they also used negative numbers in addition to verbal labels and color.

5. Discussion and Conclusions

In this paper we have evaluated the relative importance of verbal and visual language in communicating survey questions to respondents.. The effect of visual heuristics was depended on verbal labels, numerical labels, and (other) visual cues such as color. Respondents used the 'middle means typical' heuristic only in a polar point scale. When numbers were added to the polar point format and with fully labeled scales, the effect of spacing between response options disappeared. Respondents were confused

when options did not follow the 'left and top means first' heuristic. This visual effect (ordered or not ordered in a logical progression) could not be overcome with verbal and numerical labels, while color only helped respondents to order the answer options in some cases. Presenting multiple items per screen instead of using separate screens influences answers to questions in a polar point format, but not in a fully labeled format. The effect of the heuristic 'near means related' became smaller when numbers were added to the polar point format. We found no evidence that respondents use the 'up means good' heuristic. This may be due to the fact that we used a between-subjects design and not a within-subjects design. Also, in Dutch-speaking countries an incremental scale is more commonly used; this might interact with our results. Toepoel (2008) suggests that placing response options horizontally reduces the effect of visual language. Our use of a horizontal format could also explain the lack of significant differences when the heuristic 'up means good' is violated. Future research could make the effect of horizontally or vertically aligning response options clearer. We also found evidence that respondents use the 'like means close' heuristic: the adding of numbers that differed both in sign and value and the adding of different shadings of red and green color affected respondents' answers. The effect of different numbers was apparent in both the polar point format and the fully labeled format, suggesting that respondents change their interpretation of verbal labels when negative signs are used ('scale label effect'). The effect of color was only apparent in the polar point format.

In some instances it seems that the use of words really helps overcome visual layout effects. In at least one instance, a fundamental difference in visual layout (violating the 'left and top means first' heuristic) influenced answers on top of verbal labels. From a hierarchical standpoint, the data suggest that both visual and verbal

language is important. However, sometimes one can override the other and we have at least one instance of each.

Answers to a fully labeled scale were significantly more positive compared to a polar point scale. Also, the adding of numbers -2 to 2 resulted in more positive answers. Our results suggest that we may have found larger effects of numerical language when using numbers -2 to 2 instead of numbers 1 to 5. Toepoel et al. (2006; forthcoming) also found no effect of adding numbers 1 to 5 in a fully labeled scale, while they did find a significant result when adding numbers -2 to 2 to the verbal labels. The hierarchy of features suggested by Tourangeau et al. (2007) was also based on adding negative numbers.

Our analyses on respondent personal characteristics showed that older respondents are more sensitive to cues in a questionnaire than younger respondents. Also, women showed stronger design effects compared to men. Education showed little effect on respondents' reaction to (visual) layout manipulations.

A limitation of our study is the fact that we used trained respondents (by means of a panel) and this might interact with our results. It could be true that prior experience makes respondents more or less aware of layout manipulations. They may be used to a particular format, so that a different format makes them more alert. On the other hand, untrained respondents might rely more on visual cues in interpreting the survey questions. Future research using untrained respondents should show whether our results are indifferent to survey experience. Also, our experiments were fielded in five subsequent weeks. This could influence respondents' attention. It would have been better to perform all experiments simultaneously, but the number of panel members was too small to allow this.

Our results suggest that it is better to use fully labeled scales to reduce the effect of visual language. This might be related to the number of scale points, however. A scale with 7 or more response options is difficult to fully label, in terms of constructing the verbal labels and of interpreting the verbal labels. In this case, we would suggest using a polar point format with numbers starting at 1 to help respondents interpret the scale points.

Finally, for survey practitioners our results suggest that one should not assume that all visual effects can be overcome by better question wording, or that wording problems can be overcome by better visual design. It is important that we learn to work with both words and visual effects in order to improve answers to ordinal survey questions.

6. References

Borgers, Natascha, Joop Hox, and Dirk Sikkel. 2004. "Response Effects in Surveys on Children and Adolecents: The Effect of Number of Response Options, Negative Wording, and Neutral Mid-Point." *Quality & Quantity* 38: 17-33.

Christian, Leah, M. 2003. "The Influence of Visual Layout on Scalar Questions in Web Surveys. Unpublished Master's Thesis." Retrieved on http://survey.sesrc.wsu.edu/dillman/papers.htm.

Christian, Leah M. and Don A. Dillman, 2004. "The Influence of Symbolic and Graphical Language Manipulations on Answers to Paper Self-Administered Questionnaires." *Public Opinion Quarterly* 68: 57-80.

Christian, Leah M., Don A. Dillman, and Jolene D. Smyth. 2005. "Instructing Web and Telephone Respondents to Report Date Answers in Format Desired by the Surveyor." *Social and Economic Sciences Research Center* Technical Report 05-067.

Christian, Leah Melani, Nicholas L. Parsons and Don A. Dillman. Forthcoming. "Measurement in Web Surveys: the Importance of Visual Layout and Design." *Sociological Methods and Research.*

Couper, Mick P. , Michael W. Traugott, and Mark J. Lamias. 2001. "Web Survey Design and Administration." *Public Opinion Quarterly* 65: 230-253.

Deutskens, Elisabeth, Ko De Ruyter, Martin Wetzels, and Paul Oosterveld. 2004. "Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study." *Marketing Letter*, 15: 21-36.

Dillman, Don. A. 2007. *Mail and Internet Surveys. The Tailored Design Method*. Hoboken NJ: Wiley.

Dillman, Don A., Scot Caldwell, and Mary Gansemer. 2000. *Visual Design Effects on Item Nonresponse to a Question About Work Satisfaction That Precedes the Q-12 Agree-Disagree Items.* Paper supported by the Gallup Organization and Washington State University http://survey.sesrc.wsu.edu/dillman/papers.htm.

Hofmans, Joeri, Peter Theuns, Sven Baekelandt, Olivier Mairesse, Niels Schillewaert, and Walentina Cools. 2007. "Bias and Changes in Perceived Intensity of Verbal Qualifiers Effected by Scale Orientation." *Survey Research Methods* 1: 97-108.

Jenkins, Cleo R. and Don A. Dillman. 1997. " Towards a theory of Self-administered Questionnaire Design." P. 165-196, in *Survey Nonresponse*, edited by R. Groves, D. Dillman, J. Eltinge and R. Little. New York: Wiley.

Knauper, Barbel, Norbert Schwarz, and Denise Park. 2004. "Frequency Reports Across Age Groups." *Journal of Official Statistics* 20: 91-96.

Krosnick, Jon A., Sowmya Narayan, and Wendy R. Smith (1996). Satisficing in Surveys: Initial Evidence. *New Directions for Program Evaluation* 70: 29-44.

Peytchev, Andy, Mick P. Couper, Sean Esteban McCabe, and Scott D. Crawford. 2006. "Web Survey design. Paging Versus scrolling." *Public Opinion Quarterly* 70: 596-607.

Ramsey, Rosemary P. and Dawn R. Deeter-Schmelz. 2008. "An Assessment of the Psychometric Properties of the Style-of-Processing (SOP) Scale: How Do We Measure Individuals' Verbal/Visual Information-Processing Preferences." *Journal of Marketing Theory and Practice* 16: 41-55.

Redline, Cleo R. and Don A. Dillman. 2002. " The Influence of Alternative Visual Designs on Respondents'Performances with Branching Instructions with Self-Administered Questionnaires." In *Survey Measurement and Process Quality*, ed. L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin. New York: Wiley.

Redline, Cleo D., Don A. Dillman, Lisa Carley-Baxter, and Robert Creecy. 2003. "Factors that Influence Reading and Comprehension in Self-Administered Questionnaires." Paper presented at the Workshop on Item-Nonresponse and Data Quality, Basel Switzerland, October 10, 2003. Retrieved on http://survey.sesrc.wsu.edu/dillman/papers.htm.

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

Schwarz, Norbert, Barbel Knauper, Hans-J. Hippler, Elisabeth Noelle-Neumann, and Leslie Clark. 1991. “Rating Scales: Numeric Values May Change the Meaning of Scale Labels.” *Public Opinion Quarterly* 55: 570-582.

Schwarz, Norbert and Seymour Sudman. 1996. *Answering Questions*. San Francisco: Jossey-Bass Publishers.

Smith, Tom W. 1995. “Little Things Matter: A Sampler of How Differences in Questionnaire Format Can Affect Survey Responses.” *Proceedings of the American Statistical Association, Survey Research Methods Section*: 1046-1051.

Stern, Michael J., Don A. Dillman, Jolene D. Smyth. 2007. “Visual Design, Order Effects, and Respondent Characteristics in a Self-Administered Survey.” *Survey Research Methods* 1: 121-138.

Sudman, Seymour, Norman M. Bradburn, and Norbert Schwarz. 1996. *Thinking About Answers*. San Francisco: Jossey-Bass Publishers.

Toepoel, Vera. 2008. *A Closer Look at Web Questionnaire Design.* CentER Dissertation Series, 220, CentER: Tilburg.

Toepoel, Vera, Marcel Das, and Arthur van Soest. 2006. “Design of Web Questionnaires: the Effect of Layout in Rating Scales.” CentER Discussion Paper 2006-30, Tilbueg: CentER.

Toepoel, Vera, Marcel Das, and Arthur van Soest. (Forthcoming). “Design Effects in Web Surveys: Comparing Trained and Fresh Respondents.” *Public Opinion Quarterly*.

Toepoel, Vera, Marcel Das, and Arthur van Soest. 2009. “Design of Web Questionnaires: the Effect of Number of Items per Screen.” *Fieldmethods* 21(2).

Tourangeau, Roger. 1987**.** "Attitude Measurement: A Cognitive Perspective." In: *Social Information Processing and Survey Methodology*. Eds. H. J. Hippler, Norbert Schwarz, and Seymour Sudman, Pp. 149-162. New York: Springer Verlag.

Tourangeau, Roger, Mick P. Couper, and Frederick Conrad. 2007. “Color, Labels, and Interpretive heuristics for Response Scales.” *Public Opinion Quarterly* 71: 91-112.

Tourangeau, Roger, Mick P. Couper, and Frederick Conrad. 2004. “Spacing, Position, and Order. Interpretive Heuristics for Visual Features of Survey Questions.” *Public Opinion Quarterly* 68: 368-393.

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

Weng, Li-Jen and Chung-Ping Cheng. 2000. “Effects of Response Order on Likert-type Scales.” *Educational and psychological measurement* 60: 908-924.

Table 1. Chi Square tests for the middle means typical heuristic: even versus uneven spacing of response categories in five questions.

χ^2	Q1	Q2	Q3	Q4	Q5
Polar point: even (N=180) versus uneven (N=186)	8.01*	15.04**	11.92 *	18.98 **	12.92 **
Polar point with #: even (N=209) versus uneven (N=198)	5.11	6.43	7.37	2.25	3.07
Fully labeled: even (N=191) versus uneven (N=171)	3.69	5.12	6.51	7.46	2.21
Fully labeled with #: even (N=205) versus uneven (N=201)	1.52	5.41	5.35	5.46	6.90

*p<.05. **p<.01

Table 2. Ordinal regression parameters on five items with dummies for labels, numbers, uneven spacing, and interaction terms for the middle means typical heuristic.

	Q1	Q2	Q3	Q4	Q5
Fully labeled (1=yes)	.532**	.788**	.985**	.663**	.873**
Numbers (1=yes)	.026	.356*	.302	.276	.142
Uneven spacing (1=yes)	.412*	.541**	.444*	.256	.253
Interaction label*number	.102	-.138	-.111	-.036	-.168
Interaction label*uneven spacing	-.067	-.306	-.382	-.063	-.312
Interaction number*uneven spacing	-.277	-.391*	-.381	-.436*	-.283

*p<.05. **p<.01

Table 3. Chi Square tests for the left and top means first heuristic: consistent versus inconsistent scale in five questions.

χ^2	Q1	Q2	Q3	Q4	Q5
Fully labeled: consistent (N=146) versus inconsistent (N=186)	17.55**	108.97**	28.65**	174.73**	10.38*
Fully labeled with #: consistent (N=174) versus inconsistent(N=170)	16.53**	109.76**	34.33**	210.68**	14.61**
Fully labeled with color: consistent(N=199) versus inconsistent(N=181)	9.74*	134.56**	33.30**	242.01**	6.87
Fully labeled with # and color: consistent (N=143) versus inconsistent (N=180)	26.18**	121.72**	30.23**	197.21**	1.92

*p<.05. **p<.01

Table 4. Estimates of ordinal regression on five items with dummies for labels, numbers, inconsistent scale, and interaction terms for the left and top means first heuristic.

	Q1	Q2	Q3	Q4	Q5
Color (1=yes)	.325	.456*	.288	.345*	.480**
Numbers (1=yes)	-.100	-.018	-.101	-.106	.051
Inconsistent (1=yes)	.135	-2.029**	-.292	-3.466**	-.012
Interaction color*number	.079	-.110	.090	.081	-.224
Interaction color*inconsistent	-.143	-.394*	-.238	-.508*	-.286
Interaction number*inconsistent	.094	.090	-.122	.004	.315

*p<.05. **p<.01

Table 5. Chi Square tests for the near means related heuristic: one versus multiple item per screen in five questions.

χ^2	Q1	Q2	Q3	Q4	Q5
Polar point: one (N=153) versus multiple items per screen (N=162)	.68	2.63	5.49	6.57	14.83**
Polar point with #: one (N=175) versus multiple items per screen (N=175)	8.29	4.22	3.47	2.48	4.43
Fully labeled: one (N=174) versus multiple items per screen (N=191)	1.23	3.53	.59	1.87	4.43
Fully labeled with #: one (N=174) versus multiple items per screen (N=171)	3.55	5.04	3.46	.10	3.21

**p<.01

Table 6. Ordinal regression estimates on five items with labels, numbers, multiple items per screen, and interaction terms for the near means related heuristic.

	Q1	Q2	Q3	Q4	Q5
Fully labeled (1=yes)	.604**	.061	.285	.233	.346*
Numbers (1=yes)	-.015	-.205	.140	-.215	.107
Multiple items (1=yes)	.035	.041	.113	-.005	.554**
Interaction label*number	-.119	.232	-.142	.222	.044
Interaction label*multiple items	-.180	.037	-.155	-.143	-.384*
Interaction number*multiple items	.257	.006	-.110	.117	-.426*

*p<.05, **p<.01.

Table 7. Chi Square tests for the up means good heuristic: incremental versus decremental scale in five questions.

χ^2	Q1	Q2	Q3	Q4	Q5
Polar point: incremental (N=224) versus decremental (N=221)	5.60	4.71	7.98	3.54	7.22
Polar point with #: incremental (N=205) versus decremental (N=199)	4.36	2.55	5.08	4.78	.28
Fully labeled: incremental (N=203) versus decremental (N=228)	.53	2.33	3.10	.46	1.60
Fully labeled with #: incremental (N=174) versus decremental (N=237)	1.71	2.78	10.03*	2.17	4.36

*p<.05

Table 8. Estimates of ordinal regression on five items with labels, numbers, decremental scale, and interaction terms for the up means good heuristic.

	Q1	Q2	Q3	Q4	Q5
Fully labeled (1=yes)	.532**	.415**	.058	-.029	-.133
Numbers (1=yes)	-.079	-.107	.221	-.020	.016
Decremental (1=yes)	.179	.086	.295*	.207	.217
Interaction label*number	-.126	-.008	-.272	.079	.275
Interaction label*decremental	-.192	-.185	-.215	-.128	-.194
Interaction number*decremental	-.021	.144	-.462**	-.049	-.015

*p<.05, **p<.01

Table 9. Chi Square tests for the like means close heuristic: different endpoint labels in five questions.

χ^2	Q1	Q2	Q3	Q4	Q5
pp with # (N=149) versus pp with different # (N=182)	3.17	9.31*	1.02	3.05	24.72**
fl with # (N=167) versus fl with different numbers (N=163)	4.51	29.26**	19.92**	30.42**	8.35
pp (N=177) versus pp with color (N=168)	17.70**	8.15	18.62**	8.70	8.09
fl (N=162) versus fl with color (N=172)	2.66	2.04	.73	1.94	5.47

*p<.05. **p<.01

Note pp=polar point, fl=fully labeled

Table 10. Estimates of ordinal regression on five items with dummies for labels, numbers, colors, and interaction terms for the like means close heuristic.

	Q1	Q2	Q3	Q4	Q5
Label (1=yes)	1.358**	.854**	1.082**	.706**	.280
Color (1=yes)	-.324	-.411	-.145	-.097	-.039
Numbers (1=yes)	.095	-.089	.152	.149	.237
Different endpoints (1=yes)	.591**	.486*	.266	.252	.252
Interaction label*color	.566	.649	.047	.519	.661
Interaction label*number	-.342	.062	-.271	-.167	.115
Interaction label*different	-.745*	-.650*	-.188	-.120	-.383

*p<.05. **p<.01

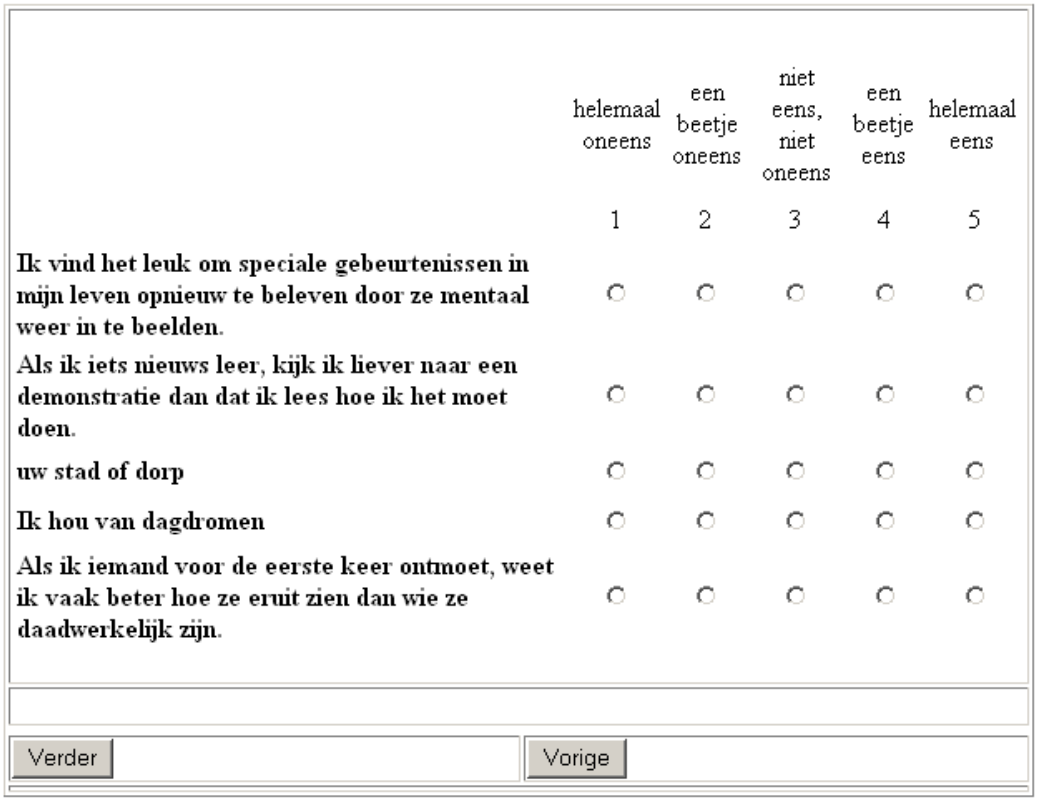
Figure 1

Experiment 1, format 5: polar point unevenly spaced

Figure 2

Experiment 2, format 7, fully labeled with color, inconsistent ordering

Figure 3

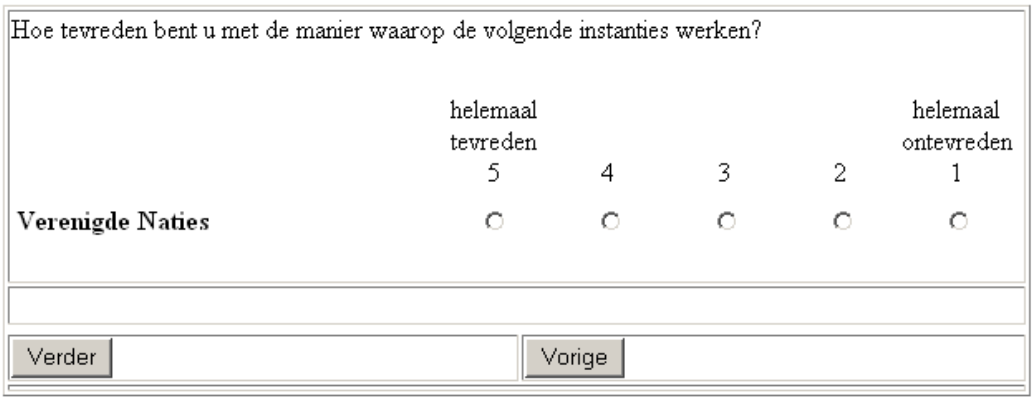


	helemaal oneens	een beetje oneens	niet eens, niet oneens	een beetje eens	helemaal eens
	1	2	3	4	5
Ik vind het leuk om speciale gebeurtenissen in mijn leven opnieuw te beleven door ze mentaal weer in te beelden.	○	○	○	○	○
Als ik iets nieuws leer, kijk ik liever naar een demonstratie dan dat ik lees hoe ik het moet doen.	○	○	○	○	○
uw stad of dorp	○	○	○	○	○
Ik hou van dagdromen	○	○	○	○	○
Als ik iemand voor de eerste keer ontmoet, weet ik vaak beter hoe ze eruit zien dan wie ze daadwerkelijk zijn.	○	○	○	○	○

Verder Vorige

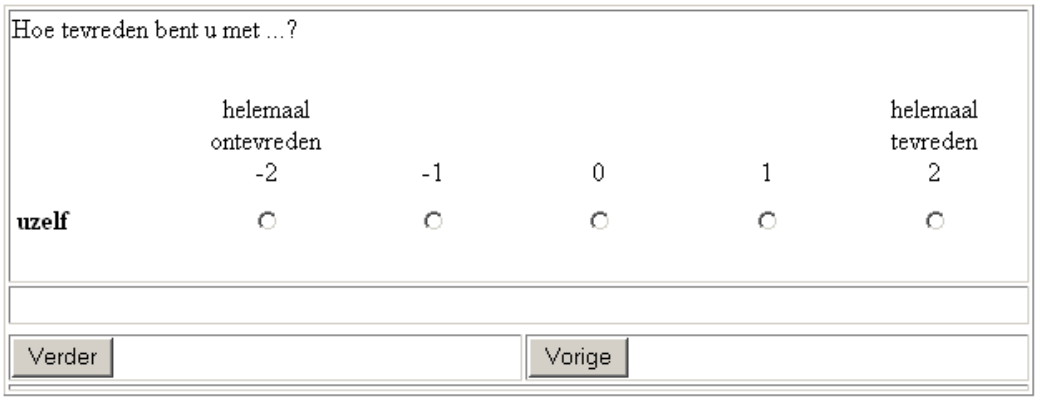
Experiment 3, format 8: fully labeled with numbers, single screen

Figure 4



Experiment 4, format 6: polar point with numbers, decremental

Figure 5



Experiment 5, format 5: polar point with different numbers

Appendix A: Items used in experiment 2 and 3

Experiment 2:
Five questions on verbalism

1. I enjoy doing work that requires the use of words.
2. I can never seem to find the right word when I need it.
3. I do a lot of reading.
4. I think I often use words in the wrong way.
5. I enjoy learning new words.

Experiment 3:
Five questions on visualisation

6. There are some special times in my life that I like to relive by mentally "picturing" just how everything looked.
7. When I am trying to learn something new, I would rather watch a demonstration than read how to do it.
8. I like to picture how I could fix up my apartment or room if I could buy anything I wanted.
9. I like to daydream.
10. After I meet someone for the first time, I can usually remember what they look like, but not much about them.

Answer options ranging from "totally disagree" to "totally agree".

All questions are based on a Style-of-Processing (SOP) Scale (Ramsey and Deeter-Schmelz, 2008).