Impact of Visual Layout in Vertically-Oriented Ordinal Scale Questions

Overview

In the new research described in this paper, a set of scalar questions were tested with different design characteristics: Displaying response options in two forms - either incremental (low value to high value) or decremental (high value to low value) order. Similar to Toepoel's, our research findings indicate that the incremental group yielded higher mean scores then the decremental group. Also similar was the finding that the amount of time respondents took to answer the questions in both orders was not statistically different.

Introduction

Researchers have found that the visual layout of a scale question can be an important source of information for respondents in selecting a response (Christian, Parsons, and Dillman 2009). Survey respondents take a heuristic approach to interpreting the visual design elements of a survey; among these related to scalar items include the middle response as the most typical, the presumption of lists of items being organized in some logical fashion, and in a vertically arranged list, the top option perceived as the optimal choice. Cues, including verbal (i.e. words) and nonverbal (i.e. design features) both independently and jointly may play a role in impacting the survey answers. This article seeks to replicate the approach in assessing whether the use of incremental or decremental designs has an effect on survey outcomes and what lessons can be drawn related to optimal survey administration.

A fair amount of research has focused on the impact that verbal cues and nonverbal cues have on respondent choices to scalar items in questionnaires. One specific area being considered here is the effect of the *verbal orientation* of a scale, or whether respondents will answer differently if response options either start or end with a positive or negative extreme. These are also known as *response-order effects.*

Toepoel (2009) contributes to this research by presenting results from a study which consisted of two separate experiments[1]; the first investigated the impact of response option presentation between linear (columns) and non-linear (three columns, two rows) and the second investigated the effects of verbal, graphical, and numerical language. Response-order effects are considered under the effects of verbal language. Toepoel summarizes these response-order effects as follows:

Primacy effects lead to options at the beginning of a response list being selected more often, while recency effects lead to options near the end of a response list being chosen more often (Krosnick and Alwin 1987). Satisficing occurs when respondents are more likely to choose items earlier in a list because they settle for the first response option they consider satisfactory, rather than going through all of them.

In this second experiment, responses were compared between one set running from positive and negative and another running from negative to positive; these were organized according to vertical layouts, horizontal layouts, and layouts with various numerical labels (1 to 5, 5 to 1, and 2 to -2). Study findings indicated statistically significant different response distributions and mean scores between a decremental (positive to negative) and [1]incremental (negative to positive) scale. The mean score in the positive to negative scale (2.91) is lower than the mean of the negative to positive scale (3.28). This appears to support the hypothesis that recency effects play a more pronounced role in impacting responses.

In addition to hypotheses related to the effect of scale orientation on responses chosen, researchers have also investigated the mean response times associated with either the decremental or incremental scale. Related to the use of visual heuristics that respondents rely upon to assist them during survey administration, Tourangeau et al. (2004) present the

---

[1] Toepoel, Vera, Marcel Das, and Arthur van Soest. "Design of Web Questionnaires: The Effect of Layout in Rating Scales" Journal of Official Statistics 25.4 (2009): 509-528

results of six experiments which attempt to measure the impacts of visual changes on questionnaire design and survey results. The heuristic tested in this case is "top means first", in which the hypothesis is described as follows:

When the list is a series of ordered response categories or scale values, respondents will expect the topmost or leftmost option to represent one of the two endpoints ("Agree strongly"); in addition, they will expect each of the successive options to follow in some logical order ("Agree," "Neither agree nor disagree") and they will expect the final option in the list – the one in the bottom or rightmost position – to represent the opposite endpoint ("Disagree strongly"). ***If the list does not conform to these expectations, respondents may become confused, make mistakes, and take longer to respond***.

The two studies cited thus far based results on a fairly homogenous population; in the case of (Christian, Parsons, & Dillman, 2009), respondents were drawn from students attending Washington State University (WSU), and in (Toepoel, Das, & van Soest, 2009), respondents were drawn from a web-based household panel in the Netherlands. Cambia previously investigated horizontal grid questions where response options are oriented lowest to highest vs. highest to lowest (Ploskonka & Srinivasan, 2009). In that study, among a number of experiments, Cambia researchers found results that were consistent with previous studies investigating response order orientation. Respondents were randomly assigned to a control or test group in which three brands were rated across 14 attributes. One group was presented with an order orientation placing the "best" rating on the left and the other group presented with an order orientation placing the "best" on the right. The results supported the hypothesis of a strong primacy effect as respondents were significantly more likely to select ratings when the scale was oriented with the "best" on the right. These respondents were also more likely to expend significantly more time in completing the questionnaire. Study findings also stipulated that some results indicating significantly higher scalar means when "best" was on the left may be due to the subject matter of the attribute content, a discussion item generally not considered in the literature.

In this study, we attempt to replicate the research design and analyze respondent data from a random sample of the general US Population. Our hypothesis focuses on two main components treated in previous studies: that the visual organization of scales has a significant impact on the response distribution and means scores between those respondents presented with an incremental or decremental orientation. Second, we hypothesize that the mean response time per question will be higher for respondents presented with an incremental scale.

Methods

The data used in this study is based on an online survey conducted January 25-29, 2012 in which 2,022 participants responded. The respondents were split among a nationwide strata (n=1012) as well as additional sample drawn from four major metropolitan areas: New York (n=302), Chicago (n=304), Los Angeles (n=202), and San Francisco (n=202). The universe was defined as the US voting population. The study was conducted by Hill+Knowlton ("H&K"). The subject matter of the online survey measured registered voters' opinion in politics and business issues. H&K wants to capture the public's view of current economic problems, the level of public trust, democratization of society, corporate reputation, corporate-public interactions and information flow.

Rather than isolate the treatment of scale orientation to one question, scale orientations were presented consistently based on the initial random assignment. Respondents were screened to ensure they met the criteria of selection, which included being registered to vote. Upon the first scale question, a random assignment was made which presented response options in either the decremental format (excellent to poor) or incremental format (poor to excellent). Once assigned, the response options were presented in an identical orientation throughout the rest of the survey. Not all questions consisted of options like poor and excellent, but the direction of responses maintained either an incremental or decremental nature[2].

Once data collection took place, t-tests were performed to assess whether there were any significant differences in mean scores between the two groups under study, with Chi Square tests performed on the response distributions. Tests were performed at the $\alpha=0.05$ level in both cases. In the case of response time, t-tests were conducted on mean response times between the two groups; mean response times are based on the paradata provided during the questionnaire administration and represent the number of seconds between initial presentment of each question and response.

Results

Table 1 (page 5) provides results for those questions where there were significant differences found between the two groups of respondents receiving either the decremental or incremental orientation. Before embarking on a discussion of those items, we focus on those areas where we found no significant differences.

As noted, the questionnaire used in this research solicited respondent opinions on a fairly wide range of current affairs and policy issues. Issues covered in these questions are included in the lists below, beginning with those items where significant differences were found2:

Significant Differences Found

- Current business conditions in the US and your local area
- How local areas have been impacted by the recent economic situation
- Level of access to information about the business practices of corporations
- Level of difficulty for public to hold companies accountable
- Emphasis corporations currently place on integrity and should place on integrity

---

[2] An illustrative example would be a question that asks "How often do you think corporate interests align with the public's best interest?" Response options to this question would include most of the time, some of the time, rarely, and never. In this instance, the incremental would present "never" first and the decremental would present "most of the time" first.

No Significant Differences Found

- Respondent perception of the relationship between trust and integrity
- The level of trust respondents place in policy actors and a range of entities
- The media's treatment of news dealing with business issues
- Whether corporate and public interest is aligned
- Meaningful actions of corporate interaction with the public
- Frequency of public use of social media to communicate opinions about specific companies

One area that has not been covered to a large extent in previous research literature is the extent to which the subject matter of a survey might impact the strength or direction of response-order effects. For example, one might hypothesize that a survey dealing with issues to which respondents might attach a great deal of affect makes it less likely that response-order orientation will have less of an impact. This outcome is due to the respondent being emotionally connected to the questions or more engaged in the survey leading to a more careful consideration of responses. The questions under study here are similar in nature to those used by (Toepoel, Das, & van Soest, 2009), which surveyed Dutch households on their opinion of quality of life and education.

Those items that were found to have significant differences between the two groups based on response order included opinions on current economic conditions and in general, public perception of corporations in terms of integrity and trust and access to information. Specific results are in contrast to results specified in (Toepoel, Das, & van Soest, 2009) which found the following:

Our two questions show statistically different answer distributions and mean scores between a decremental and an incremental scale, indicating that respondents are affected by verbal language…The mean score in the positive to negative scale is lower than the mean of the negative to positive scale, providing evidence for a primacy effect. Our results thus provide empirical support, in a different country and culture than the literature, for the theory of satisficing and primacy effects.

Table 1 indicates similar outcomes in assessing response-order effects, and provides for each question in which significant differences were found, the sample size for each group, the scale type (High to Low; Low to High), the range (whether the question consisted of four or five options), the proportion of respondents answering low to high, mean scores, results of statistical tests (t-test and Chi-square) on proportions and means, and on average time to complete questions.

				Percentage of Respondents Selecting Category						Difference of Means			Difference of Means	
				Low ←———————→ High								Mean Time		
Q	N	Scale Type†	Range	1 Poor	2 Not Very Good	3 Good	4 Excellent	5	Mean	t test Chi Sq	p χ		t test	P
2	1012	Dec	4 point	20.2	63.6	15.5	0.7		2.0	4.47	0.00	11.0	-.517	.605
	1010	Inc		13.8	64.7	20.7	0.9		2.1			13.8		
3	1012	Dec	4 point	18.0	50.9	29.9	1.2		2.1	3.27	0.00	11.8	.532	.595
	1010	Inc		12.5	51.8	34.6	1.2		2.2			9.7		
4	1012	Dec	4 point	23.0	51.1	19.4	1.8			43.58	0.00	55.9	-.537	.591
	1010	Inc		21.3	43.9	22.6	7.3					60.3		
9	1012	Dec	5 point	7.2	13.7	30.8	31.9	16.3	3.4	5.14	0.00	15.6	-1.177	.239
	1010	Inc		5.0	9.3	25.2	40.0	20.4	3.6			24.6		
10	1012	Dec	5 point	13.4	27.0	30.3	24.3	4.9	2.8	4.60	0.00	99.6	.263	.792
	1010	Inc		10.3	21.3	31.9	28.3	8.2	3.0			94.1		
15	1012	Dec	4 point	24.9	59.2	14.2	1.7		1.9	3.82	0.00	50.7	1.142	.254
	1010	Inc		19.1	61.1	16.7	3.1		2.0			11.8		
16	1012	Dec	4 point	3.8	5.1	52.5	38.6		3.3	4.46	0.00	11.9	1.563	.118
	1010	Inc		0.4	3.8	52.4	43.5		3.4			9.4		

Table 1: Differences Between Decremental and Incremental Scale Direction

Note: Shaded cells indicate significant differences between proportions at α=0.05 level; †Dec=Decremental Scale, Inc=Incremental Scale. Difference of means t test assume equal variance.

The two questions rated on a five-point scale indicates that there were significant differences nearly across all responses[3], and those presented with the incremental scale orientation tended to respond more positively. For instance, in Q9[4], which measured public perception of access to information about business practices of corporations, higher proportions of respondents presented with the incremental orientation answered "much more access" (20.4% cf. 16.3%) and lower proportions of respondents answered "much less access" (5.0% cf. 7.2%). Results of Q10[5], which measured public perception of the

---

[3] Only the middle response in Q10 ("About the same") did not indicate significant differences.

[4] Compared to ten years ago, do you feel like you have more access, less access, or about the same access to information about the business practices of corporations?

[5] Compared to ten years ago, do you feel like it is easier or more difficult today for the public to hold companies accountable for their actions?

level of difficulty in holding companies accountable indicated similar results,

whereby respondents presented with the incremental orientation answered "much easier" (8.2% cf. 4.9%) and lower proportions of respondents answered "much more difficult" (10.3% cf. 13.4%).

Of the other items which indicated significant differences, all were based on a four-point scale and all exhibited somewhat similar outcomes. In general, significant differences in response option distributions were found among the low and middle categories (in Q2, Q3, and Q15). Similar to the five point scales described above, the responses from those presented with the incremental scale orientation tended to answer more positively to all of these questions. This positive result was indicated in all of the questions in which significant differences were indicated.

For instance, when asked about the current business condition in the US, the poor to excellent group registered a "good" rating of 21%, significantly higher than the excellent to poor group's "good" rating of 15%. When asked about local area business conditions (Q3), the poor to excellent group generated a 35% "good" rating, significantly higher than 30% "good" rating of the high to low group. In both Q2 and Q3, low to high groups both had significantly higher scale means than high to low groups. Very few respondents answered that business conditions in the US or their local area could be rated "excellent" and the proportions that did answer in this manner were identical between the two comparison groups.

The low to high group had a significant rating on "it is a top priority" (3% vs. 2%) and "it is one of the most important priorities" (17% vs. 14%) compared to the comparison group when asked "how much emphasis do you feel US Corporations currently place on integrity?" (Q15). Mean of Q15 low to high group was 2.0, significantly higher than mean of the high to low group of 1.9 on the 5-point scale. When asked "how much do you feel US corporations should place on integrity", significantly more low to high group respondents selected "it should be a top priority" (43% vs. 39%) compared to high to low comparison group (Q16). The low to high mean of 3.4 on the 5-point scale was significantly higher than the high to low mean of 3.3.

Q15 and Q16 tended to indicate differences between the two groups at either end of the response distribution (with Q15 indicating significant differences among response options "It is a top priority", "It is lower on the list of priorities", and "It is not a priority" and Q16 indicating significant differences among the two polar opposite responses in that question, "It is a top priority" and "It is not a priority").

All of the tested questions had associated timestamps which tracked how much time it took respondents to get through the tested questions. Means of those times across all respondents between the two groups were tested with independent mean T-tests. None of the T-test results yielded any significant differences. In addition, total time across all questions did not indicate any difference. In conclusion, respondents did not spend a significantly different amount of time between the two question orders (low to high and high to low).

Discussion and Conclusions

Similar to Toepoel's research, our research findings indicated that the incremental (low to high) group yielded higher means and more positive ratings compared to the decremental (high to low) group. Results generally support the conclusions found in (Toepoel, Das, & van Soest, 2009) that both primacy and satisficing effects play a role in responses based on the verbal orientation of scalar items. This research also contributes to the current literature by establishing similar results using randomly selected respondents from the general population.

Additionally, total time to take tested questions had no significant difference at both the individual question level and at the aggregate level. Further research including the same type of experiment and numerical language attached to scale should be conducted to assess differences between scale orders. In conclusion, we recommend randomizing both decremental and incremental orders on scalar questions in future research. This solution will likely neutralize any bias created by scale orders.