Best on the Left or on the Right in a Likert Scale

Overview

In an informal poll of 150 educated research professionals attending the 2009 Sawtooth Conference, 100% of those who voted raised their hands to indicate they would put the best value in a scale at the right. Nevertheless, since then the author has seen a number of panel surveys and others with the best at the left.

Initially and at the time of the conference, we would have agreed with the 150 without reservation. Some of our results since are not as straightforward and we encourage researchers to continue examining this issue in different industries, with different respondent types, and with brand sets that are both similar and others that are quite diverse.

Prior Work

The first substantive work we found on this topic presented four to seven-scale point items on paper, in person, vertically (Belson, 1966). Belson reminds us that there is (as yet) no evidence which scale direction more accurately reflects the respondent's mindset … only that the scale presentation order influences the results. They tested "traditional" (high to low) [sic] order against the reversed, across respondents. Among the results they found (n=332):

- Items at the ends of the scale are particularly subject to order effects, in particular the first item presented experienced a bump
- The effect was consistent regardless of the length of the scale, or type of scale (approval, satisfaction, liking, agreement, interest)

Belson concludes by questioning if horizontal scales or products or issues where the respondent is less familiar or interested would see the same effect. This topic will be addressed later within this paper.

Holmes (1974) tested horizontal bipolar scales among 240 beer-drinking respondents. Two of his results relate:

- Respondents' responses were regressing toward the center from the beginning of the questionnaire to the end
- Respondents were more likely to choose the response at the left side of the page (that is, the first presented for an English reader)

Holmes notes that our assumption in sampling theory that measurement errors will be uncorrelated and cancel each other out may not be so.

Another variation tested was to include both favorable and unfavorable statements with a 5-point strongly agree to strongly disagree scale (or SD to SA) (Friedman, Herskovitz and Pollack, 1994). The researchers continued to find a bias towards the left side of the scale (n=208 undergraduates), but only for those statements where the attribute was worded positively. In general, the attitudes measured (towards their college) were quite positive and it appeared the students would disagree with the unfavorable statements no matter where 'disagree' was located.

Our Study

We intended to evaluate the pros and cons of each orientation so that we can could advise our clients on the design of future surveys. We wanted to be able to evaluate the respondent experience, as these are real people who we want to come back and take our surveys again. Indeed, our respondents may be future clients, so offering an intuitive and user-friendly survey is paramount to our company's current and future success. We also wanted to measure if either orientation increases discrimination between brands and the impact on rating differences.

Our study differs from the past studies we reviewed in that it was conducted online, with ratings of multiple targets or brands on multiple attributes.

Respondents were administered a roughly 13-minute questionnaire on some aspects of the healthcare industry. Greenfield Online provided the sample, with the following respondent qualification criteria:

- Age 18+
- Covered by health insurance
- Makes health insurance decisions for their household

Differences in the control and test groups were as follows:

	Control	Test
Location of best value	Left	Right
Interviews (n)	1,047	203
Field dates	June 11 – July 4, 2008	September 8 – 16, 2008

Respondents rated three brands with which they were familiar, one brand per screen, on a series of 14 attributes, on a grid with a bipolar seven-point scale (the one very top company, world class, stronger than most, average, weaker than most, much worse than other companies, the one worst company – and the reverse, followed by don't know) as seen in Figure 2.

Figure 2
Control and Test questions for Left-Right vs. Right-Left Study

Results

Our results were consistent with past studies conducted on paper: Given the negative

---

end of the scale first, respondents were significantly more likely to choose the negative attributes than when those scale points were placed to the far right of the screen as in Chart 1.
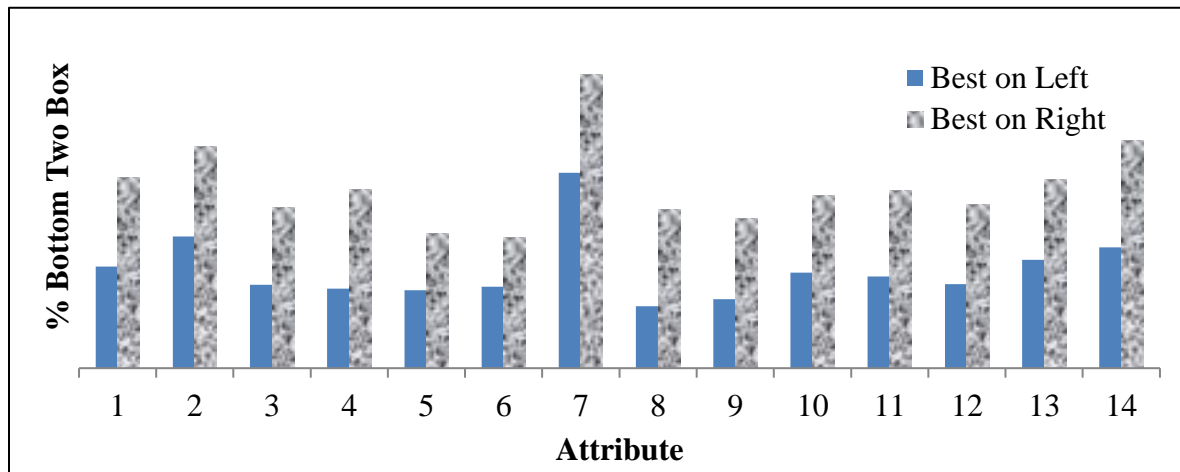

Chart 1

Twelve out of 14 of these attributes showed significant differences for bottom 2 box. The two attributes that did not show differences also had the lowest bottom 2 box ratings when best was on the right, but otherwise were not especially distinctive from the 12 that were significant. Top 2 box was unchanged, statistically, whether it was presented first or last.

Four statements had significantly higher means when best was on the left:

- helping people get the care they need
- helping people live a healthy lifestyle
- selling products and services in all segments of the market
- demonstrating ethical financial practices

These statements may differ from the others in being less operationally-defined, or at least less likely for the respondent to have had personal experience with the brand for that aspect. It is possible that the scale may have more of an impact on the respondents' choices when they have no capacity to measure the behavior of the brand themselves.

As for respondent experience, respondents who received the best on left format completed the questionnaire significantly faster (10%) than those who received the best on right format, trimming time to completion by 2%. It appears to be more of a cognitive burden to read across and choose when best is on the right. Regardless of the scale orientation, the majority of responses were on the positive end of the scale.

In addition, the standard deviations were consistently higher when the best was on the right, significantly so for five of the 14 attributes.

We then considered how well respondents were able to differentiate between brands.

Stacking the data by attribute shows that at least on a gross level (grand mean), respondents are separating the brands more distinctly when best is on the left, in particular between Brand 3 and 4 (Chart 2). Means could be masking differences occurring on a more granular level. Moreover, using discriminant analysis, measuring the Euclidean distance from the origin for centroids, shows 3 brands are undiscriminated when the best value is on the right, as in Chart 3.
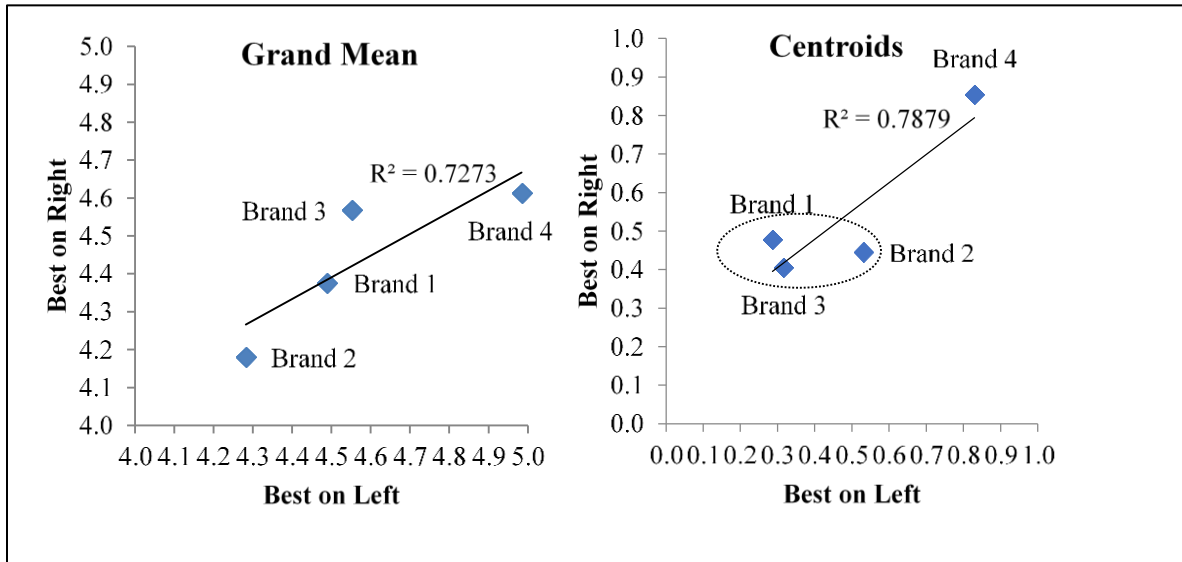
Charts 2 and 3

It would appear that while best on the right produces more variance within brands using this Euclidean distance algorithm, best on the left produces more variance between brands, for our limited dataset with four brands only, and one quite different from the rest. The means, on the other hand, demonstrate Brand 3 aligning with Brand 1 in one case and with Brand 4 in the other. The story told by this is therefore not clearly about differentiation but instead about a change in scores following a change in presentation style.

Comparing regression coefficients using "likelihood to recommend" as the dependent variable resulted in insignificant results using the Chow test. However, multicollinearity in all four studies is quite high.

Conclusion

Our results of a survey conducted online support past results from paper surveys: The orientation in which a scale is presented will influence the outcome, and the negative end of the scale is more likely to be selected when presented first. It might be further hypothesized that seeing the most negative end first gave respondents implicit permission to choose it.

Past results did not delve into the differentiation of the items being measured, only the

difference in means. Simply looking at means leaves the differentiation unclear. Using a Euclidean algorithm, with only four brands we see more differentiation between brands with the best on the left. This result deserves further exploration – would the results repeat with different questions, sample, brands, or context?

Furthermore, we hypothesize that repeating these tests with languages that are read in a different direction would show similar effects for primacy and not just for absolute orientation.