



Designing an Attribute Battery to Measure Reputation

Overview

Although randomizing brands within attributes is often accepted as the best approach for gathering comparative data in image ratings, not all agree. In fact, in our literature review we found at least one strong argument for also randomizing attributes within brands. Nevertheless, our attributes are strongly intercorrelated and it was our hypothesis that they might be less so if the question were formatted to ask the respondent to rate a group of brands for one attribute before proceeding to the next attribute. You will see as you read ahead that we did succeed, but only minutely.

Prior Work

Ninety-four years ago, Edward Thorndike observed that respondents had difficulty separating their ratings of individual attributes for a person from that of their overall perception of the person. As a result, correlations between attributes were higher than reality. The term “halo effect” comes from this study, published in 1920.

As researchers, we generally find halo effect to be a bad thing. We want respondents to be able to distinguish brands one from another, to distinguish attributes one from another, and in the end, provide data that allows us to identify which attributes are distinctive drivers of success for our clients.

It is possible to take another viewpoint, though. The ‘halo’ itself can be extracted as a single dimension and treated as brand reputation (Laroche, 1978) or brand equity (Leuthesser, et al., 1995, Dillon, et al., 2001) – with the remainder examined for differences. Leuthesser suggests double-centering the data to make it ipsative and then running analyses on it. Other authors (McClellan and Chissom, 1986) suggest this is unwise. Rossi, Gilula, and Allenby (2001) have followed up with a Bayesian alternative to ipsatization, Dillon et al with a decompositional model – I recommend these resources to you for a more detailed explanation and examination.

However, for the purposes of this research we sought to diminish the halo effect, and we believe that randomizing brands within attributes is a better way to achieve this than randomizing attributes within brands. But before we jump to conclusions, Torres and Bijmolt (2009) found

... when the association between brands and attributes is measured asking brand-to-attribute associations, which is a non-comparative format, the stronger links from the brands to the attributes dominate the associations. On the other hand, if a researcher measures brand image asking attribute-to-brand associations (a comparative format), stronger links from the attributes to the brands will determine the perceptions of the consumers...we suggest that both directions of associations should be considered when brand image is assessed to make managerial recommendations.

To translate this to plain English, think of a brand in a particular category, for example cars. What attributes come to mind? Now, think about just one of those attributes. Which cars come to mind when you think of this attribute? The order in which the question is asked may result in an asymmetrical correspondence between brands and attributes, depending on the strength of the brand’s personality and the impact of the particular attribute.

Our Study

In our study, we hoped to evaluate the level of halo effect in our data, and decrease it by randomizing brands within attributes in comparison to the control group where attributes would be randomized within brands.

In addition, pre-tests showed a shorter study time when one brand was asked at a time for all attributes. We would monitor and report on this as well.

Respondents were administered a roughly 13-minute questionnaire on some aspects of the healthcare industry. To participate, respondents had to fulfill the following 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
Randomization	Attributes within Brands	Brands within Attributes
Interviews (n)	1,047	266

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, don't know) as in Figure 1.

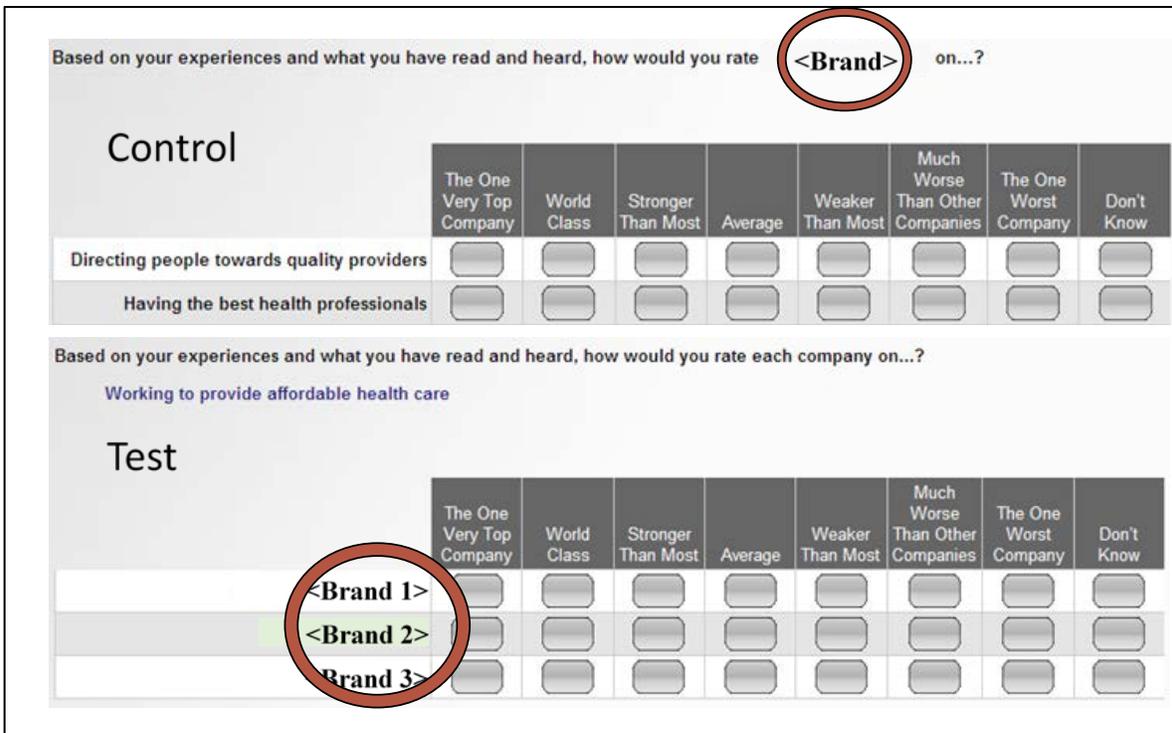


Figure 1
Screen shots of Brand within Attribute vs. Attribute within Brand questions

Results

In brief, we found that the mean is higher in the “across brand” (Test) condition, as is top 2 box (by 9% on average) and standard deviation. These results strongly indicate that it would be unwise to switch formats from one wave of research to the next. It remarkably changes the order of the means ... the ranking of the attributes only correlates .44 from one group to another (compared to .92 for the left-right, right-left study).

Again, Chow tests of differences in regression coefficients across the two samples using the rating variables are not significant, using likelihood to recommend as a dependent variable. We advise caution due to the high multicollinearity. The average inter-correlation

for attribute within brand was .81. For brand within attribute the average inter-correlation was .76. We cannot say that we succeeded in reducing multicollinearity. A rule of thumb suggesting the presence of a halo effect is an intercorrelation between .60 and .70. We beat that handily.

Below in Chart 1 is differentiation on the rotation order data:

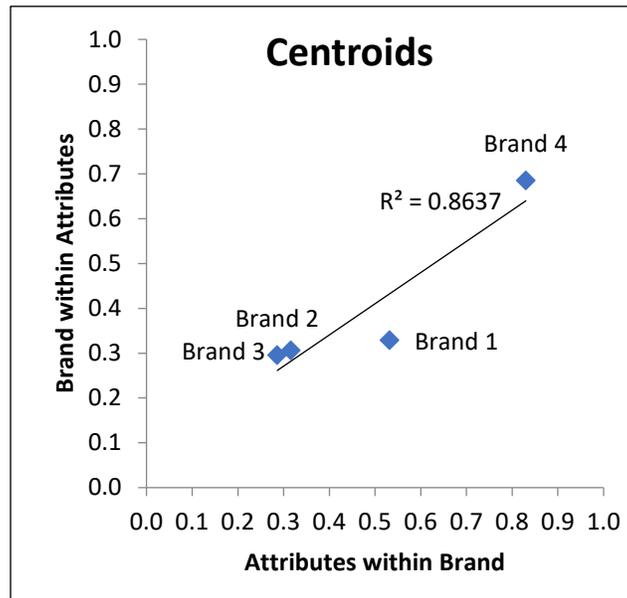


Chart 1

While at first there appears to be better differentiation asking attributes within brand, it makes sense that one could be reducing multicollinearity at the expense of discrimination. This leads us back to Torres and Bijmolt (2009)'s advice ... we may be provoking a different response and gathering a different (and yet still valid) story in changing how the question is asked.

One interesting finding was in applying the ipsative approach suggested by Leuthesser, et al. With the normalized, standardized healthcare data we derived two very interesting varimax factors, each easily labeled on a bipolar scale and very intuitive. The financial data was not as amenable to our attempts to reduce the multicollinearity this way, however.

Finally, there was no statistical difference in the time to complete the survey.

Conclusion

The halo effect is very high, and only slightly less so when brands are asked within attribute. The means changed and the mean order changed. Brands were rated more highly when they were compared with other brands.