



**Research on Research White Paper Series:**

## **Card Sort: Validation Study**

## Product Description

Toluna has developed a "Solitaire" like card sorting technology for use in online interviews. The technique provides alternatives to typical online questions and helps maximize respondent cooperation, and thoughtfulness of response.

### Card Sort is ideal for:

- Projects that contain lengthy attribute lists.
- Studies conducted with demographic groups that tire easily, or are difficult to engage (ie. teens, ethnic groups.)
- Multiple images to be evaluated and compared.

Using Dynamic HTML (DHTML) technology, Toluna can ask respondents to complete a visual exercise of sorting statements as they would a deck of cards.

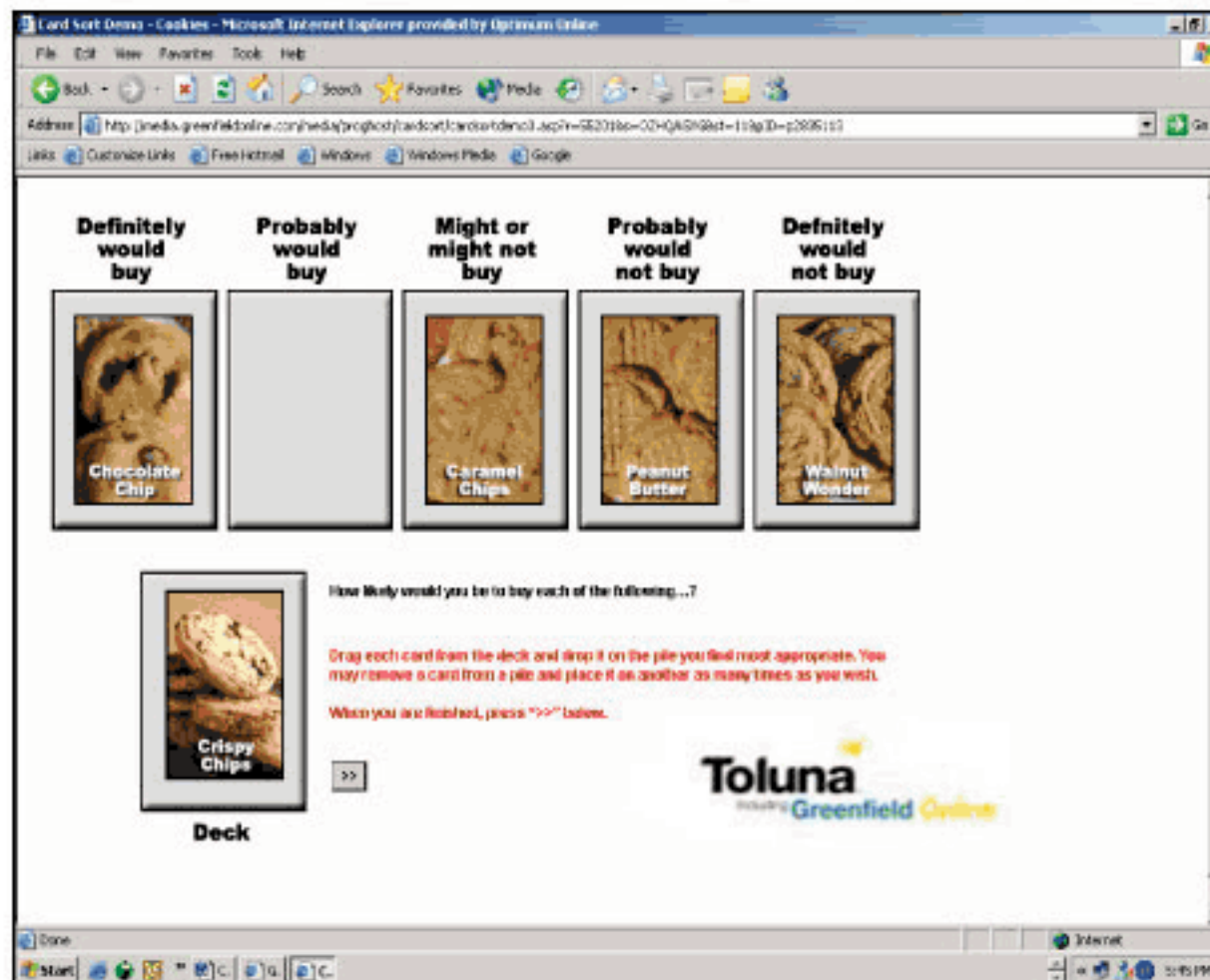
The deck of statements, names or product variations appears on the screen with

a rating scale, similar to the way a sort board would be presented to a live respondent. The respondent is asked to sort the cards by clicking on the statement and then drag and drop each to appropriate rating.

## Background

The Card Sort technology was evaluated in a side-by-side comparison within two separate Toluna Omnibus studies in July. Sample size of each Omnibus is N=1000. The test was set-up to randomly select respondents for one of two cells (of equal probability) within each Omnibus. The question consisted of 32 ratings of the likelihood of using the Internet to perform a specific task.

The survey consists of 32 ratings of the likelihood of using the Internet to perform a specific task. The rating scale used was "Definitely would use", "Probably would use", "Might or might not use", "Probably would not use", "Definitely would not use". Five hundred six respondents completed the Card-Sort task while 475 used the traditional rating scale grid methodology.



## Card Sort: Validation Study

To enhance the user experience, the card sort was developed with a "snap" functionality that allowed respondents to drop each card near a pile, whereby the browser would instantly affix the card to the nearest pile. Respondents also had the ability to remove a card from a pile and place it on another pile. Validation code ensured that every card was placed prior to proceeding to the next question.

The survey was programmed to capture the length of time each respondent spent completing the question.

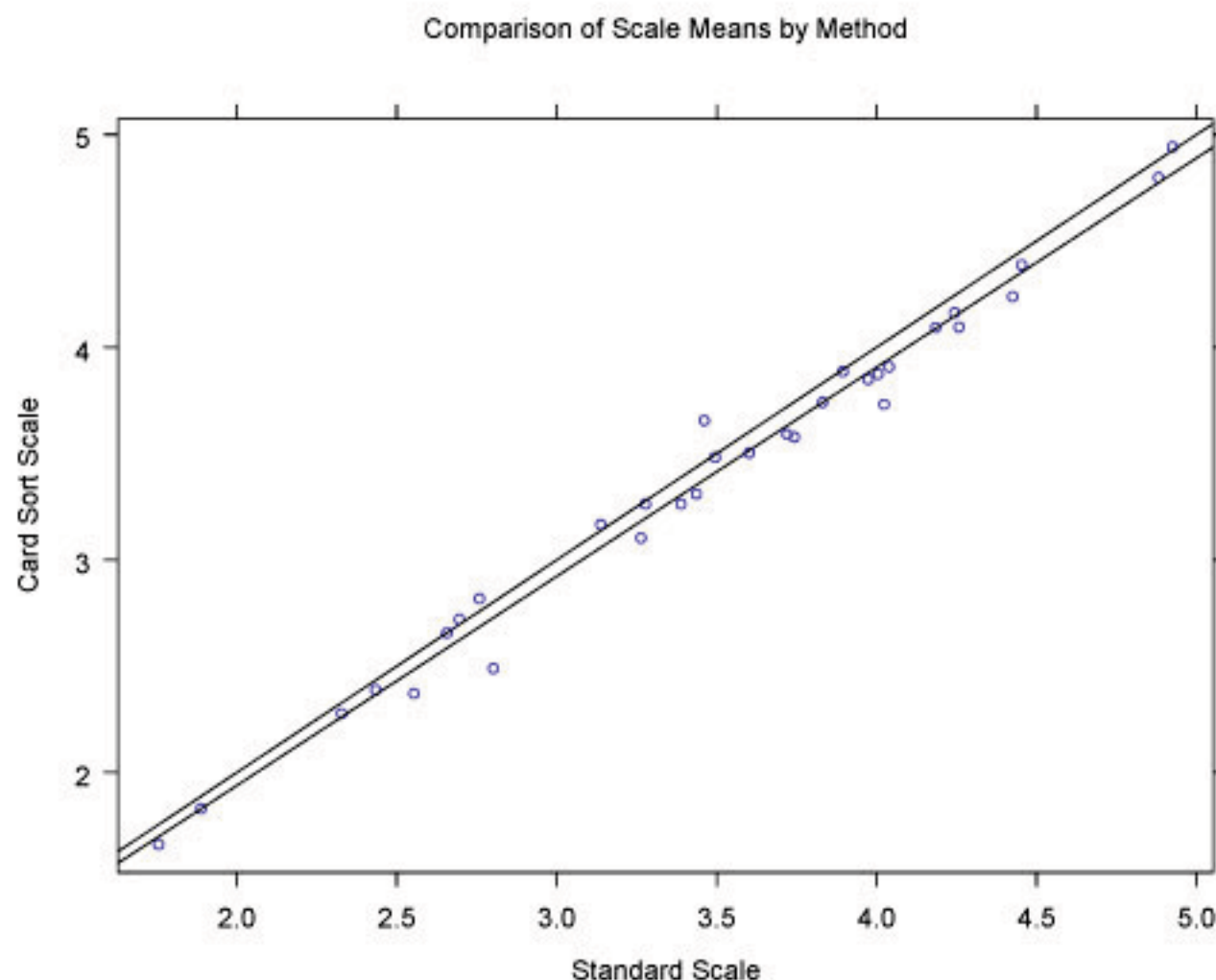
### Analysis

The first set of analyses deal with the consistency of the data between the two approaches. Clearly, since we do not know the "true" answers of the respondents we would like to reach the same conclusions using the two methods.

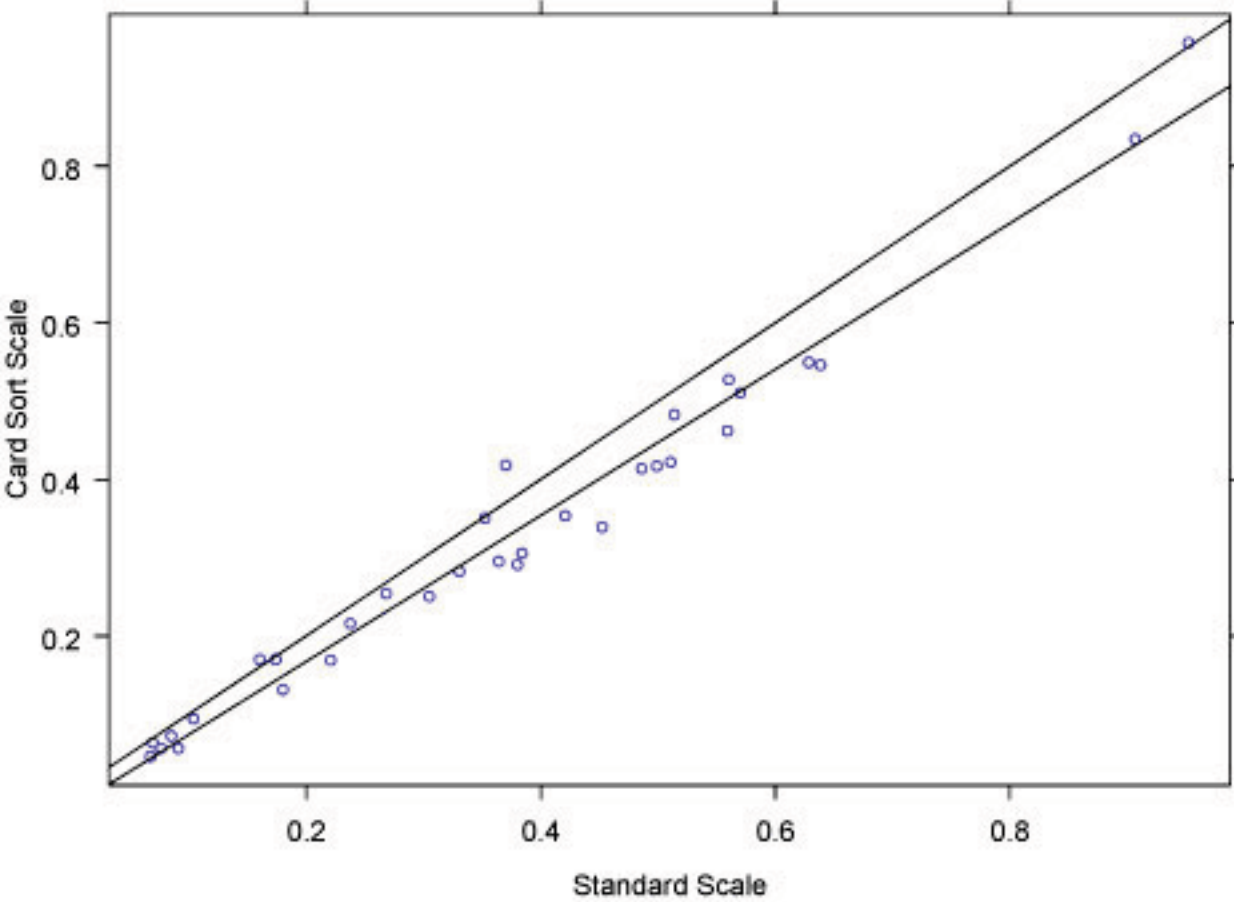
The chart below compares the mean scale scores of the 32 items across the two methods. The upper diagonal line shows the point of exact equality between the

methods. The line that is slightly lower represents the least squares regression line of this data. This indicates that the mean scores in the Card Sort procedure are generally slightly lower than the scores in the traditional procedure. However, the correlation between the mean scores is .997 and the rank correlation is .987. These scores (as well as the visual inspection of the chart) show that conclusions based on mean scores would be virtually identical in either method. The overall difference in scores does not appear to be related to the actual score (i.e. the difference is about the same whether the actual score is high or low).

Similar analyses on top box and top 2 box scores show that the biggest discrepancy is in top box scores. The average top box score is about 4 points higher in the standard methodology. However, the correlation between scores is still excellent (.979 correlation and .987 rank correlation). This indicates that decisions would rarely be different between the two methods.



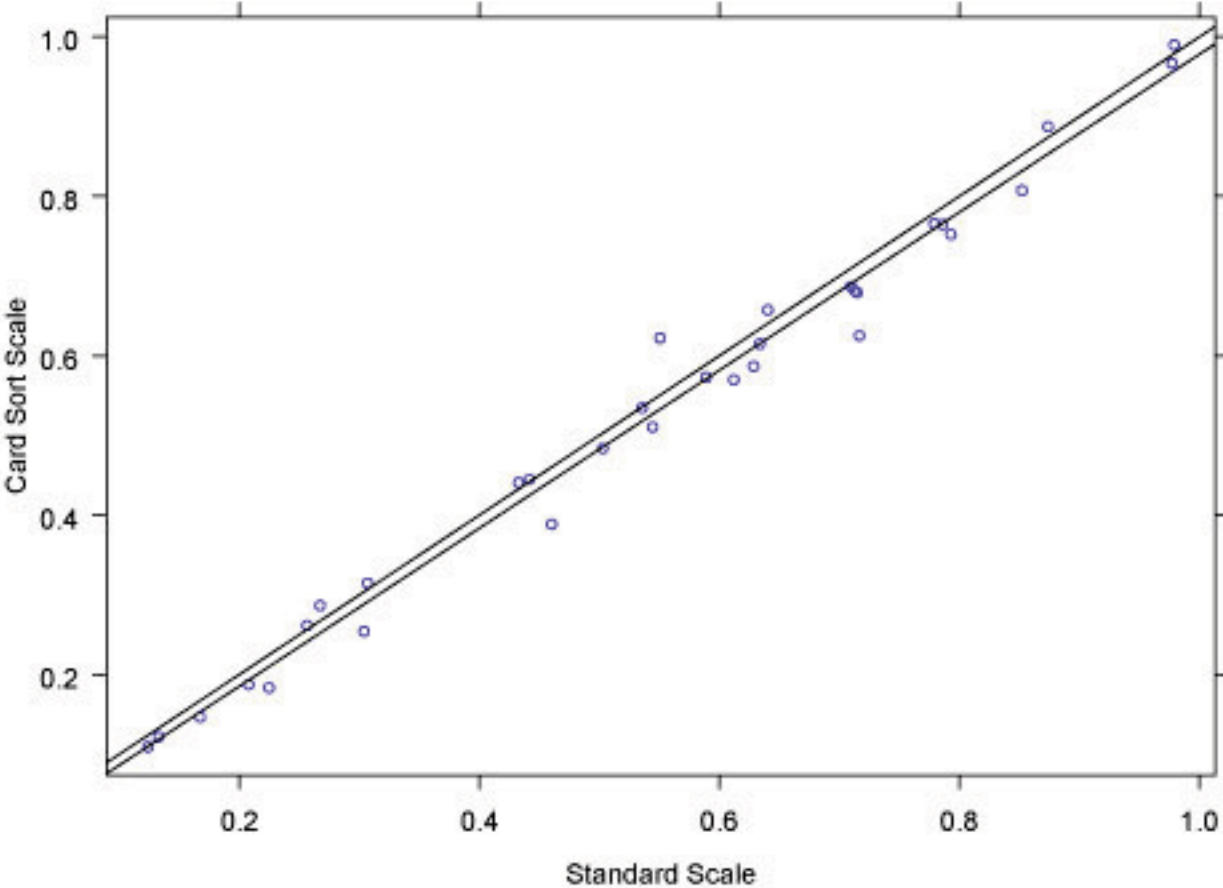
Comparison of Scale Top Box by Method



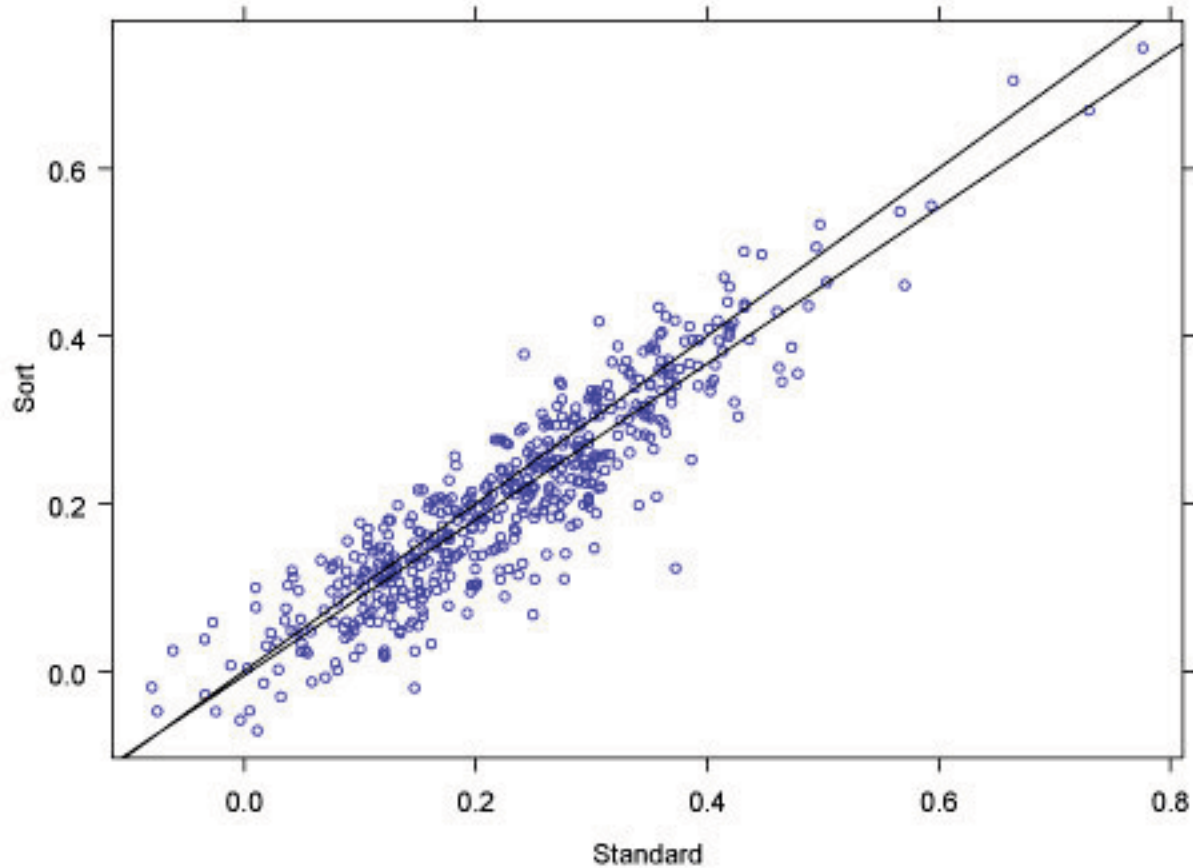
With a top 2-box analysis, the results are more similar. The mean difference shrinks to 1.7 points and the correlations

remain quite high (.986 correlation, .993 rank correlation).

Comparison of Scale Top 2 Box by Method



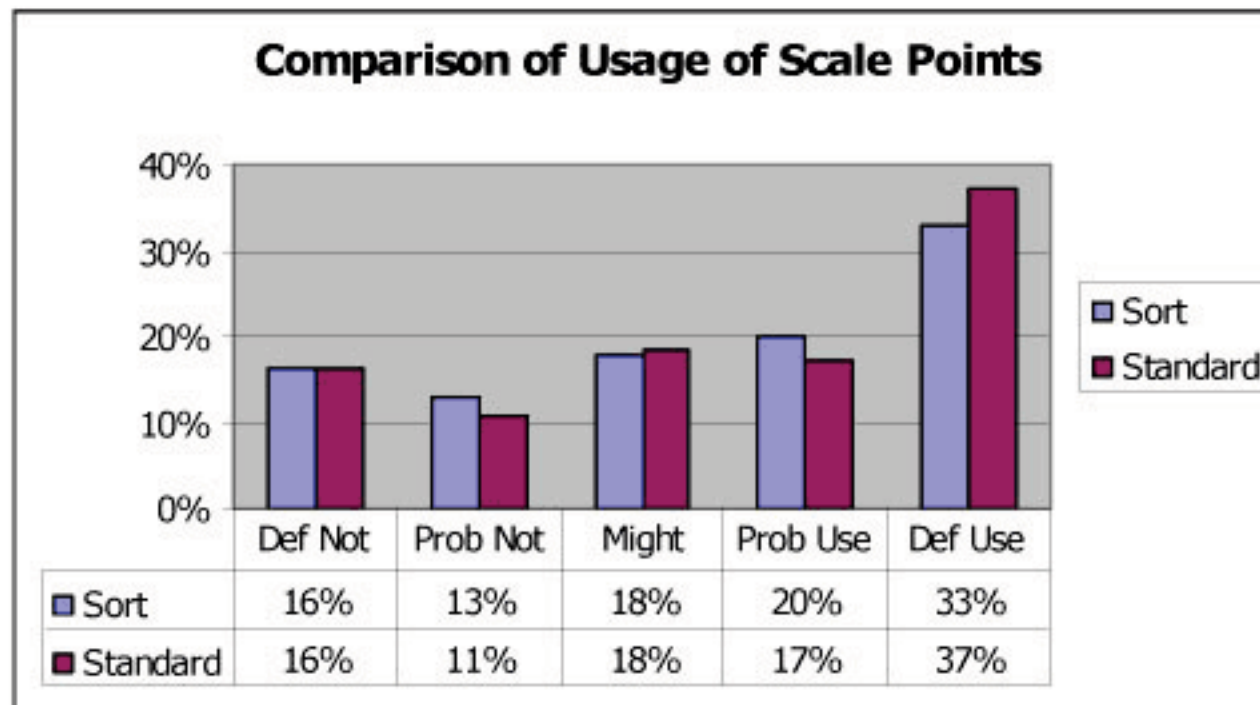
Comparison of Pairwise Correlations by Method

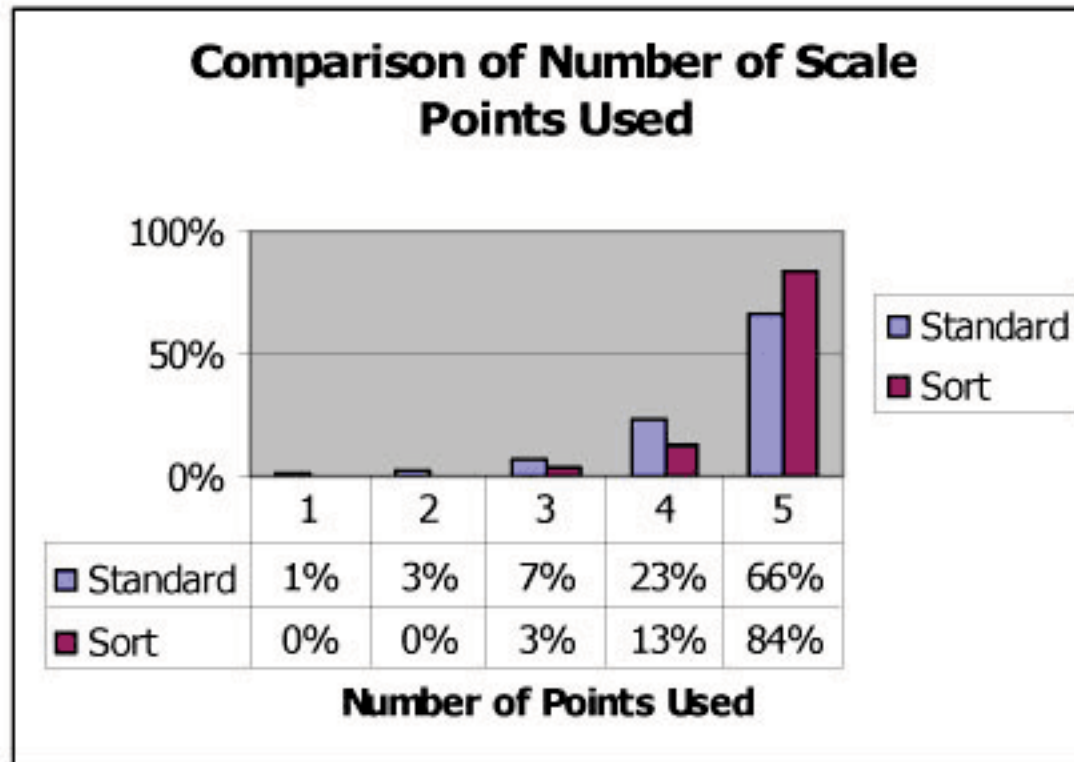


A second important aspect of scale data is the amount of correlation between variables that is present in the data. We can see in the next chart that the correlations generally follow the same pattern, i.e. highly correlated items will be highly correlated in both methods. The average correlation between items in the data is about .02 higher in the standard scale data than in the Card Sort data. You can

see in the chart, that the difference is more pronounced among the more highly correlated items. The question is whether the higher correlation in the standard scale data is meaningful.

Looking further at the differences in scale usage between the methods we can compare the proportion of times that each scale point was used.





Here we can see a slight tendency to use the top box more in the standard scale and to use the second and fourth boxes more in the sort method.

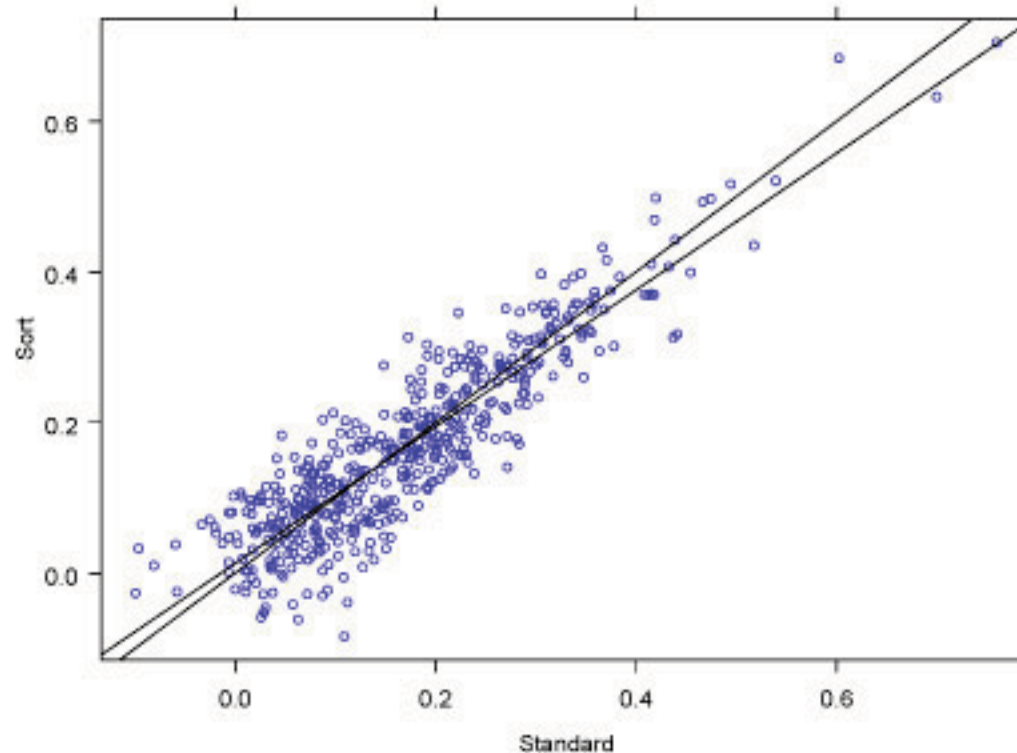
We can also look at the distribution of the number of scale points used by each respondent. Here we see a nice feature of the card sort technique – no one in the sample used only a single scale response for all 32 items. Also, a much larger percentage of respondents used the entire

scale (all 5 points).

Going back to our earlier discussion of the correlation between items we can wonder if the higher correlations in the standard rating data are due to a tendency for respondents to confine their answers to one portion of the scale.

If we limit the correlations analysis to those who used all five-scale points in either method we find that the average correlations are virtually identical.

Comparison of Pairwise Correlations by Method  
Respondents Using All 5 Scale Points



## Conclusions

Overall, we find no evidence that any different conclusions would be drawn from data collected under the sort methodology. Given the expectation that respondents will find the task more interesting and engaging, Card Sort is a viable method for collecting data in online surveys.

The average length of time to complete each question type was:

- Card sort – 159.3 seconds
- Standard grid – 144.1 seconds

This indicates respondents spent an average 10.5% more time to complete the card sort.