October, 1998

An alternative method of reporting customer satisfaction scores

By: Rajan Sambandam, George Hausser

Editor's note: Rajan Sambandam, Ph.D., is director of research, and George Hausser is director of project management, at The Response Center, a Philadelphia, Pa., research firm. They can be reached at 610-352-2800.

Though customer satisfaction evaluations are widely used, reporting of these scores has varied from one study to another. This is likely the result of each method's advantages and disadvantages, as well as the personal preferences and habits of the researcher. We recently had the opportunity to report customer satisfaction scores in a unique format that assimilates the advantages of various methods and provides the manager with a clearer picture of where to take action. In this article we review various reporting methods and outline our method with an example. Further, we also discuss a type of reporting that is becoming increasingly common especially in the health care arena, i.e., the issue of comparing the performance of various facilities or centers that belong to a single network or organization. We show how our method can be applied for this purpose and why it is advantageous.

Current reporting formats

First, consider the prevailing methods for reporting attribute satisfaction scores.

Mean scores

Top two box scores

Top box scores

Top two box and bottom two box scores

(Whether it is the top two box or top three box score that is reported is irrelevant because the principle remains the same.)

The advantage of reporting the mean is that it is a summary score that takes into account the frequency of answers for each scale point. As often found in practice, the disadvantage is that the manager doesn't see much difference between the mean satisfaction scores on various attributes.

For example, on a seven-point satisfaction scale, the mean satisfaction scores on many attributes tend to be clustered in the five to six region. Thus, even though some of them may be statistically different from others, for a manager, it is not easy to understand where efforts need to be focused to improve overall satisfaction. Further, means of attributes which have bimodal distributions (high top two box and bottom two box scores) could have the same mean as those with normal distributions.

Reporting top two (or top three) box scores has the advantage of providing more variation in the data. Thus, it is easier to identify the attributes where the company is performing poorly. The disadvantage with a top two box score is that it is often quite high. This sometimes leads to complacency because its high value seems to indicate that respondents are very satisfied. However, reality could be quite different because the measure doesn't take the full distribution into account. Consequently, even though the bottom box scores may be different, two attributes with equal top two box scores could be considered to be equal.

Top box score has been suggested more recently as a better way of reporting satisfaction scores. It includes only "totally satisfied" customers and hence the lower value associated with a top box score discounts any notion of managerial complacency. However, as in the previous case, top box scores ignore the rest of the scale, potentially masking trends in bottom box scores.

To overcome this problem, in some cases researchers report both top two box and bottom two box scores. While this provides a more complete picture, it forces the manager to integrate two pieces of information for each attribute. This is especially hard when there are many attributes included in the study and, more important, when scores are compared across many attributes and locations.

Alternative format

The solution we developed for this situation was to create a single statistic which utilized both top two box and the bottom three box data from a seven-point scale. More specifically, we subtracted the bottom three box score from the top two box score to provide a single rescaled score that not only varied between attributes but also took both ends of the distribution into account.

Having constructed this new score, we took it a step further. Rather than just subtract the bottom three box score, we subtracted twice the bottom three box score from the top two box score. (Refer to the example below.)

This formulation has several implications. Marketing studies conducted by us and other researchers have shown that there exists an asymmetric effect on satisfaction. That is, the impact of negative attribute performance on overall dissatisfaction is higher than the impact of positive attribute performance on overall satisfaction. By using twice the bottom three box score, the formulation proposed here takes asymmetry into account by saying that a customer in the bottom box hurts a company more than the gain provided by a customer in the top box.

It also means that moving a customer out of the bottom two box is harder than moving a customer into the top two box. Therefore, the former achievement should be rewarded more than the latter. Because the bottom two box score is weighted to twice its value in this formulation, moving customers in and out of the bottom boxes has a greater impact on the scaled attribute score than moving customers in or out of the top boxes.

The number of scale points to be included in the top and bottom boxes is dependent on the scale used in the study and the distribution. Similarly, the extent to which the bottom boxes should be weighted (twice, thrice, etc.) is dependent on the particular study. But the principle remains the same. A single score is obtained that has more variation than the mean score, includes both ends of the scale and is weighted to include the asymmetric effect.

An example

Consider a company that has completed an attribute satisfaction survey with 100 respondents, where scores are measured on a seven-point scale.

Scale	Attribute 1	Attribute 2
7	41%	45%
6	30%	26%
5	10%	15%
4	5	9%
3	7%	3%
2	4%	2%
1	3%	0%
Mean	5.69	5.95
Scaled Score	43	59

While there is a statistical difference between the means, for a manager the difference may not appear substantial. The top two box scores are the same for both attributes, and the top box scores are very close. If we use the formulation described here, the scaled score for Attribute 1 would be (71 - 28 =) 43 and that for Attribute 2 would be (71 - 10 =) 61. This is a substantial difference and it occurs because of the higher proportion of bottom box scores on Attribute 1 (even though the top two box score is the same).

Now consider what happens if for Attribute 1, 3 percent of the customers are moved out of the bottom three box while for Attribute 2, 3 percent of the customers are moved into the top two box. The score for Attribute 1 is now (71 - 22 =) 49 and that for Attribute 2 is (74 - 10 =) 64. While there is still a substantial difference, clearly Attribute 1 is seeing more movement in the score than is Attribute 2. Depending on the scale used in the study and the nature of the industry, the formulation can be varied to provide an appropriate score.

Comparative reporting

While the new formulation described here seems simple and intuitive, its usefulness does not become fully apparent until we consider a specific situation. Consider the situation that companies are increasingly facing where they need to not only measure the performance of different organizational entities, but also to compare the performance of the entities for evaluation purposes. An example of this is health plans that need to evaluate physicians or health facilities that are affiliated with them. Another example is of financial services companies that have branches in different parts of the country. We were faced with the latter situation where 20 centers belonging to a financial services company needed to be evaluated on a variety of attributes to determine the overall level of customer satisfaction with each center.

Each center was evaluated on a list of 10 attributes using seven-point satisfaction scales. For each attribute a rescaled score was calculated using the convention described previously (top two box score less twice the bottom three box score). Thus, when the centers were compared on a single attribute, those with high top two box scores and low bottom three box scores had the highest scores.

These attribute scores had to be made meaningful to the managers of the centers. To do this, they had to be expressed in a form that managers could easily understand. As a first step, the 20 re-scaled scores for Attribute 1 were averaged. Next, the average score for Attribute 1 was standardized by applying a multiplier to convert it to 75. Thus, centers scoring above 75 were above average on Attribute 1 and centers scoring below 75 were

below average. The same process was applied to each of the other nine attributes. Thus, each center had a standardized score for each attribute and the mean for each attribute across centers was 75. We used 75 to denote the average score because it was felt that it would be in line with the perceptions of the managers who would ultimately look at the data. This average can be any number depending on the situation.

The attribute scores for each center had to be converted into a single score that was representative of the performance of the center, so that the centers could be easily compared. To do this, the scores on the 10 attributes for each center were averaged, to create an overall satisfaction score for each center. Since this makes the average of the overall scores across centers also 75, determining above and below average performance for each center is not hard.

This process provides many advantages. First, as explained in the previous section, centers with high top two box and low bottom three box scores are rewarded. Centers are rewarded more for moving customers out of the bottom boxes than for moving them into the top boxes. Second, the standardization allows each center manager to compare his score against all of the other centers without having to look at the individual scores for each of the other centers. Finally, the standardization to 75 allows each manager to clearly understand performance, since excellent scores ranged close to 100 and low scores were in the 60s or 50s. Using any of the existing methods would require the manager to process either more information or incomplete information.

Summary

In this article we have provided an alternative method of reporting customer satisfaction scores that incorporates the advantages of existing methods while avoiding some of the disadvantages. We have also demonstrated an application of this method by way of providing standardized comparative evaluation scores. This method will allow researchers to help managers understand the results of customer satisfaction studies better, by providing more information in an easier format.