

# The “Perfect” Scale

The number of points on a scale is less important than its application.

By Diane H. Schmalensee

**W**ith so many businesses recognizing the importance of customer satisfaction, many researchers are taking a serious look at the viability of their own measurement methods. The articles under discussion in this issue illustrate the industry's struggle toward refinement. Given the variety of scales tested and the criteria used by the authors, it's not surprising that they all reach different conclusions.

After reading these articles, I have drawn several general conclusions:

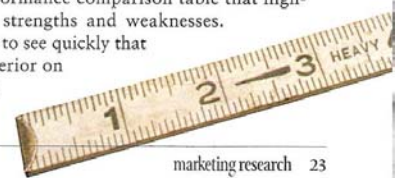
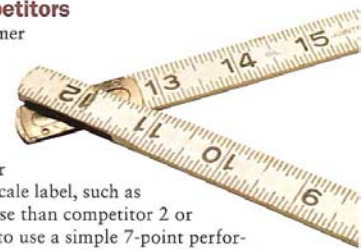
- It's virtually impossible for a single article to consider all the possible combinations of labels and number of scale points. There are simply too many possible combinations to be analyzed at one time.
- The choice of scales is highly dependent on the criteria used. Although all the authors' criteria are important to them, the criteria themselves are highly subjective.
- The choice of scales may be (and should be) strongly influenced by the research objectives when measuring customer satisfaction.

Some case examples may help illustrate why the choice of scales is so dependent on research objectives and market conditions. Rather than using the 4-, 5-, and 10-point scales under contention, I'll use a neutral example of a 7-point scale in most cases.

## Fewer Strong Competitors

Company 1's customer satisfaction goal was to outperform its two best and largest competitors to build customer retention and accelerate customer acquisition. Rather than use a comparative scale label, such as “performs better or worse than competitor 2 or 3,” this company chose to use a simple 7-point performance label (from “poor” to “excellent”), with respondents rating each company separately on each attribute.

This led to a performance comparison table that highlighted competitive strengths and weaknesses. Executives were able to see quickly that Company 1 was superior on variables A and B, in a dead heat on vari-



marketing research 23

## Executive Summary

After analyzing three articles dealing with the issue of the best scaling choices for customer satisfaction research, the author concludes that the selection of labels and number of points on the scale should be driven largely by company objectives and market situations.

able C, and behind on variable D. This scale tested well for reliability and permitted Company 1 to monitor how well it was meeting its strategic objectives in a way that was meaningful to executives.

### Rising Customer Satisfaction

During the pretest of its 7-point “poor” to “excellent” scale, Company 2 discovered that it was rated a 6 or 7 by 96% of the respondents and that no one rated it below 4. While the company was delighted to know how satisfied its customers were, its researchers were dismayed because this left the firm almost no room to improve over time and provided poor discrimination and modeling ability.

Rather than increasing the number of points (this was a phone survey), the researchers decided to change the endpoint labels ranging from “poor” to “excellent” to a “good” to “excellent” range with a zero being used if respondents deemed the firm to be less than “good.” Exhibit 1 shows how the change in labels affected scores and provided much more discrimination.

### From Performance to Standards

Most research on scales assumes that the company is conducting its customer satisfaction survey to obtain strategic

information about what’s important to customers (potential or existing), how the company is performing, possibly how other firms (named or ideal) are performing, or what priorities for change will do the most to increase satisfaction.

However, an increasing number of companies are using surveys for tactical reasons, simply to determine whether they’re living up to customer-driven performance standards. Company 3, for example, conducted tactical satisfaction surveys to measure customer satisfaction with key events and determine whether its people had lived up to performance standards during those key events.

To illustrate, Company 3 learned from its strategic customer satisfaction research that its sales calls were vital to customer satisfaction. After conducting additional research on sales calls, the company set a variety of performance standards for its salespeople, including arriving on time for appointments, being familiar with the customer’s account history and records, and remembering to send all promised materials after the meeting.

After training its sales staff on the importance of these behaviors, Company 3 developed a sales call satisfaction tracking survey to measure overall satisfaction with the event and whether the salespeople had lived up to the new standards.

The results from this 10-point overall satisfaction survey became part of the incentive system for the sales department. The company also used the 2-point behavior results scale to suggest when remedial training was needed (e.g., “arrives on time for appointment,” 2-point “yes” or “no.”).

Company 3’s performance tracking scale is very similar to the one recommended by Wittink and Bayer. The only difference is that Company 3’s 2-point labels were “yes/no,” whereas Wittink and Bayer used “problem/no problem.” On a purely subjective basis, the “yes/no” scale appears to be more sensitive than the “problem/no problem” scale because customers may not necessarily call it a problem if a salesperson arrives a bit late for an appointment, but the company knows that this small annoyance (which would be picked up by the “yes/no” scale) actually decreases customer satisfaction.

### Literal-Minded Executives

Company 4 was just starting on its quality improvement/customer satisfaction journey. Its executives had never used survey research before and were very suspicious of scales that were not fully anchored. “How do I know what the customer is thinking if she gives us a 4?” asked one of the executives.

Although we knew that a 7-point scale with endpoints labeled “poor” and “excellent” would have strong statistical advantages, the company chose to use a 5-point scale (1 = poor, 2 = fair, 3 = good, 4 = very good, and 5 = excellent) because its executives could relate better to it. They were much more able to agree to act on scores that were only “poor,” “fair,” or “good” than on scores that were 1-4 out of 7 points. Company 4 has made considerable progress in the years since then and has decided to move to a 7-point scale for greater discrimination and modeling ability.

Exhibit 1 Effects of altering labels for Company 2

	Poor	1	2	3	4	5	6	7	Excellent
Old		1	2	3	4	5	6	7	
		-	-	-	1%	3%	41%	55%	
New	Good	1	2	3	4	5	6	7	Excellent
		2%	9%	10%	18%	19%	22%	20%	

### **The Bottom Line**

Grapentine makes an excellent point about the many difficulties associated with the expectations (“much better/much worse than expected”) and requirements (“exceeded/missed requirements”) scales. Chief among these for most companies I have worked with is the interpretation difficulty; executives cannot tell if changes in scores are due to changes in requirements/expectations or to changes in performance.

This makes it tricky for executives who champion change to convince top management that action must be taken and allows the “do little” executives to slow down the process with numerous debates. The SERVQUAL scales (which separate for expectations and performance evaluations) were developed by Valarie Zeithaml, A. Parasuraman, and Len Berry for just this reason.

Wittink and Bayer argue convincingly that using more points on the scale usually increases sensitivity, precision, and the ability to use statistical methods to infer importance. Unfortunately, Wittink and Bayer don’t tell us whether they used the same endpoint labels on their 5-point and 10-point scales or whether they worded their attributes identically, so

we have no way of knowing why they obtained different importance rankings for the independent variables using the different scales. Without this information, I remain uncertain about how superior the 10-point scale truly is for identifying the most important attributes.

Most people would probably agree that Wittink and Bayer’s 10-point plus 2-point scales make perfect sense for tracking firms’ ability to live up to performance standards and satisfy customers during key events (such as the sales call). But would they agree that the same 10- and 2-point scales are the best choice for a strategic customer satisfaction research project that aims to understand competitive strengths and weaknesses? Probably not.

The bottom line is that the choice of labels and number of points on the scale can and should be driven largely by company objectives and market situations. Given the almost infinite variety of objectives, no single scale is “perfect” in every situation. ●

---

Diane H. Schmalensee is president of Schmalensee Partners in Chestnut Hill, Mass.