Discussion Paper on Scales for Measuring Customer Satisfaction

Prepared by
Market Directions'

Market Directions
200 Mt. Vernon St.,   Boston
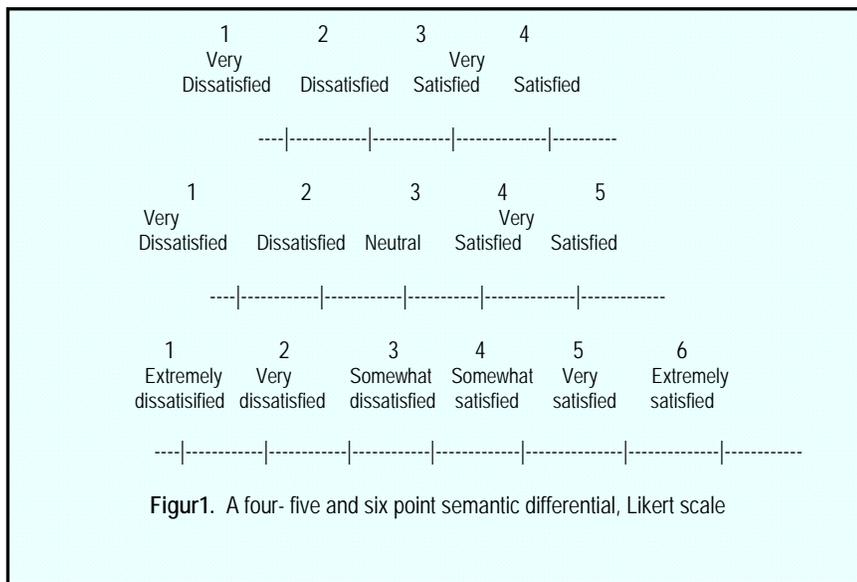**800-475-9808**
info@makretdirectionsmr.com

Contents

Scales for Measuring Satisfaction

Survey design research reveals inconclusive results on which scales to use in measuring customer satisfaction. However, there is reliable research that indicates that your results can be impacted by the scales that you choose for your survey. The most common scales used to measure customer satisfaction are classical "Likert[1]" scales – named after its creator. These scales have numeric values that describe the range of possible attitudes about something --- from "very dissatisfied" to "very satisfied."   When constructing simple rating scales, a researcher must make a number of major decisions including (1) the number of scale categories to be used, (2) the verbal stimuli (or numbers) that should be used to label scale categories and (3) the physical scale form.

This paper briefly describes and discusses different measurement scales and some strengths and weaknesses of each. The ideas are presented to encourage thought and debate as to how you may wish to design your next survey rating scales.
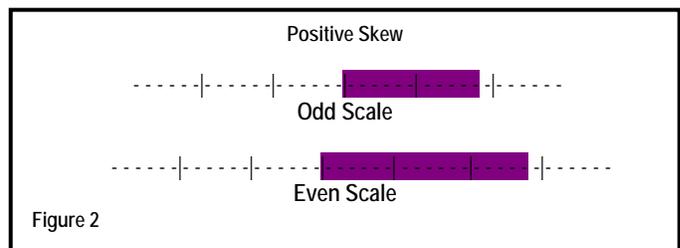
The format of a typical Likert looks like this:

1	2	3	4
Very Dissatisfied	Dissatisfied	Very Satisfied	Satisfied

----|------------|------------|-------------|----------

1	2	3	4	5
Very Dissatisfied	Dissatisfied	Neutral	Very Satisfied	Satisfied

----|------------|------------|----------|-------------|------------

1	2	3	4	5	6
Extremely dissatisfied	Very dissatisfied	Somewhat dissatisfied	Somewhat satisfied	Very satisfied	Extremely satisfied

----|------------|------------|------------|-------------|-------------|-------------|----------

Figur1. A four- five and six point semantic differential, Likert scale

Likert allows you to manage qualitative data by ascribing quantitative value to attitudes which makes the data useful for statistical analysis.   Likert scales usually have five potential choices (strongly agree, agree, neutral, disagree, strongly disagree) but sometimes go up to ten or more.   The numerical value assigned to each potential choice (the anchor) enables you to calculate a mean figure for all the responses. This final average score represents overall level of accomplishment, attitude or satisfaction toward the subject matter. The wording of the questionnaire, scale choice, and the anchor choices will make a difference in satisfaction measurements.

The objective of a customer satisfaction survey is for a positive trend over time--- it should enable you to discover problems, take action, measure, and look for improvement in results—implement results and resurvey. A common problem when using Likert scales can be that a majority of answers are at the upper extreme. Some studies indicate that most customer satisfaction surveys score in the eighty to ninety percent range. This commonly leads executives to believe that '80% of customers rate their satisfaction at a four or five; without realizing that it simply means that 80% are only at least **somewhat satisfied**. In statistical language this is referred to positive skew. When distributions have a positive skew, **the mean is higher than the median**.
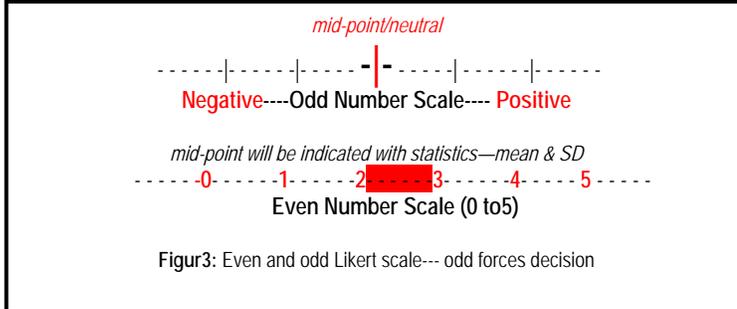
With these high ratings managers have little incentive to improve or to repeat the survey when changes to processes or procedures are made. A scale with good dispersion properties allows for measuring improvement trends and some research believe that scales with fewer points seem more susceptible to score inflation.   A well designed survey should have a dispersion of responses

Positive Skew

- - - - -|- - - - -|- - - - -|- - - - -|- - - - -|- - - - -
Odd Scale

- - - - -|- - - - -|- - - - -|- - - - -|- - - - -|- - - - -
Even Scale

Figure 2

---

[1] Rensis Likert (American social scientist))
sought to find effective and systematic means of studying human attitudes and the factors that influence them. His research led him to develop a scale for attitude measurement. Now known as the Likert Scale, it offers a means of determining attitudes along a continuum of choices, such as "strongly agree," "agree," and "strongly disagree."

to separate out the "critical few" on which to focus attention.

The use of a middle or neutral

mid-point/neutral

Negative----Odd Number Scale---- Positive

mid-point will be indicated with statistics—mean & SD

-0- - - -1- - - -2- - - -3- - - -4- - - 5

Even Number Scale (0 to5)

Figur3: Even and odd Likert scale--- odd forces decision

points (e.g. even point vs. an odd point scale). Some scholars advocate an odd point scale where respondents can have a "neutral" middle point referred to as bipolar measurement either positive or negative response to a statement. Sometimes a four-point scale is used whereas others prefer to "force" people to select a negative or positive position with an even point scale. Studies are inconclusive if a "don't know" option increases accuracy of responses.

Central tendency is somewhat problematic and some research indicates surveys are subject to a central tendency form-related error. This statistical incident occurs with many evaluation, training and satisfaction scores, especially employee evaluations.

Scale design impacts how respondents will answer a question. And when managers' performance is tied to satisfaction measurements, the incentive to design survey questions and scales without bias is threatened. The goals

Central Tendency Ranges

Odd Scale

Even Scale

Figur4: Even and odd Likert scale--- odd forces decision

of a performance measurement system using surveys -- where higher scores are wanted -- can compromise the role of that same survey program in identifying areas for quality improvement.
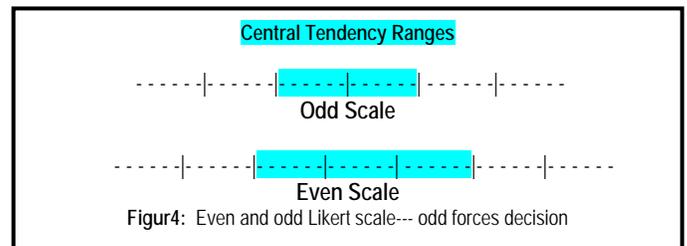
Overall when Likert scales (from 5 to 7 points) are used for satisfaction measurements, research of results indicates that there is a predisposition to be positively skewed or respondents choose "middle" of the road scores.

It All in the Results and Analysis

While there is no perfect way to design a survey without some bias towards some groups--- being aware of the bias is a first good step in over coming it. This is where a good analyst is an invaluable asset in the survey process. When results are tallied graphed and finalized, the final report should not merely present distributions without reminders as to what bias may be influencing the skew of the results. The analysis and presentation of the results is where these issues should be addressed and discussed.

There are several theories on how to acquire satisfaction measurements through surveys to reflect broad dispersion properties. These include using:

> **1.** larger number scales such as,
> > a. a ten point scale or
> > b. eleven point scale (0-10);
> **2.** verbal descriptions instead of numbers;
> **3.** scoring where 'zero' is used as a midpoint and
> **4.** a two-stage format of simple rating scales

These survey design elements are discussed below.

1. Larger Numbered Scales

Ten- and eleven- point scales: Some researchers believe a five-point scale is totally inappropriate for customer satisfaction studies because of the statistical errors that are discussed above and because scales with fewer points can be more susceptible to inflated results. They choose ten point scales because it is easier for customers to think in terms of percentages.

Researchers claim that a ten point scale is intuitive to respondents because of the familiarity they have with this rating system and is a standard that is being adopted by many of the world's leading companies. In addition, ten point + scales are suitability for correlation and regression, which provides standardization for benchmark data.

On a scale of 1 to 10 how satisfied are you? Where 10 is extremely satisfied and 1 is extremely dis-satisfied

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
❑	❑	❑	❑	❑	❑	❑	❑	❑	❑

Ten point scales can look a bit overwhelming for customer satisfaction survey participants, resulting in a limited range of responses. Because of this, they can be misleading. For example, 7.5 out of 10 might sound like a good result, but some research indicates people will only score between 6 and 9 unless they feel very strongly about the issue. Taking this into account, a rating of 6 actually indicates that a respondent is somewhat dissatisfied, 9 is satisfied, and the mid point of 7.5 is only neutral.

2. Verbal Descriptive Scales

Some studies indicate that presenting a scale as a series of verbal descriptions -- for example, ***"Are you extremely satisfied, very satisfied, somewhat satisfied, somewhat dissatisfied, very dissatisfied, or extremely dissatisfied?***" -- lead to more dispersion and less clustering of responses.

Overall are you?
Extremely Satisfied………………………………………..
Very Satisfied……………………………………………….
Somewhat Satisfied………………………………………..
Somewhat Dis-satisfied…………………………………….. .
Very Dis-satisfied…………………………………………..
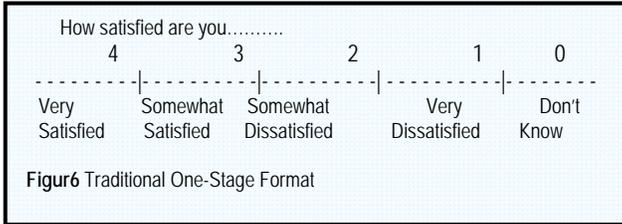*Extremely Dis-satisfied…………………………………….*

3. Midpoint=0, whereas 0 = "met expectations"

Some organizations' goal in customer satisfaction isn't just to satisfy customers but to ***delight*** customers by providing service that goes beyond expectations. This ambition to "delight" customers began aggressively in the "quality " campaigns of the eighties when the U.S. was scurrying to compete with the reputations of Japanese

Has our services……….
-2	-1	0	+1	+2

Fell Short of Your Expectations   Met Your Expectations   Greatly Exceeded your Expectations

Figur5 Expectations
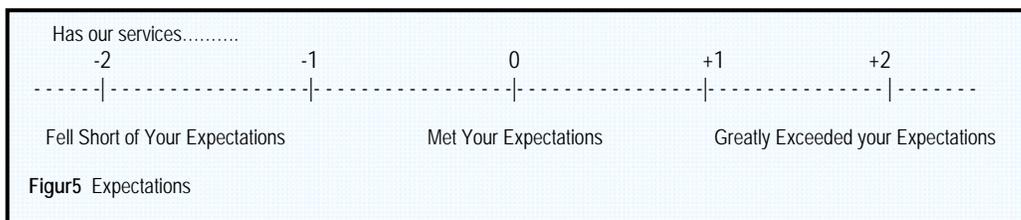
manufacturers. The scale in Figure 5 above does just that. The level of service **delivered just as promised** puts you in the center of the scale, and the numbering scheme, while still a 5-point scale, reinforces the midpoint as an average position. However, the zero can be a problem for some respondents. So using this scale with only the verbal descriptors may minimize that problem. This scale also offers is a good compliment to using just the "Ultimate Question[i]" in survey satisfaction design. (This Ultimate question asks….. *Would you recommend this service to others?* Whereas, Market Directions asks….. ***On a scale of 1 to ten how likely are you to recommend this service to others?***)
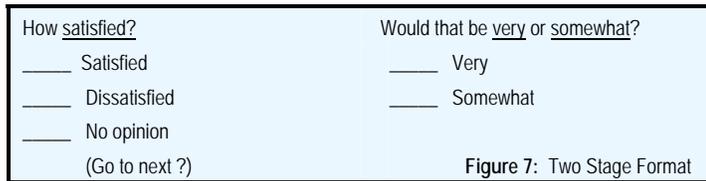
4. Two Stage Format

How satisfied are you..........

4	3	2	1	0
Very Satisfied	Somewhat Satisfied	Somewhat Dissatisfied	Very Dissatisfied	Don't Know

Figur6 Traditional One-Stage Format

A study in 2007[ii] experimented with a two stage format in receiving feelings on customer satisfaction--- they tested they traditional "one stage format" v. a two stage format as depicted in Figure 6.

The experiment concluded that the greatest percentage of extreme position responses across scales was generated

by the two-stage format of simple rating scales, and suggests that one stage format is possibly subject to a central tendency form-related error.

How satisfied?	Would that be very or somewhat?
_____ Satisfied	_____ Very
_____ Dissatisfied	_____ Somewhat
_____ No opinion	
(Go to next ?)	**Figure 7:** Two Stage Format

Conclusion

Designing a survey is easy--- designing a great survey is time consuming and requires expertise, knowledge and experience. With the dawn of, dare I say it *"Survey Monkey"* everyone believes they can design and implement a survey. After all they have taken plenty of them so they know what a survey should look like. And that's what drives companies to take on the task of building their own customer satisfaction surveys. Most people like to believe that they are fair and honest (and most are)--- but when your job, your reputation, and your salary depend upon customer satisfaction--- it is almost humanly impossible not to incorporate bias into your own survey.

There is a strong argument for using wording that leads to a lower percentage of top box scores, this is what leads you to improvement. However, to receive an initial low score can be damaging to both morale and ego. But it is the long run picture that management should strive for--- improvement over time means that you have to begin at the bottom point to rise to the top.

If I had my wish--- (and I rarely do) I would have my clients use the "met expectations" scale for customer satisfaction surveys, my second choice is the two stage format. When people have a dialogue with one another they don't say: *"we just tried a new restaurant and I was somewhat satisfied."* People tend to speak in black and white with a neutral middle--- so when companies speak to their customers with a survey it should be poised as a dialogue-- so I also like scales with mid-points, where the mid-point by management should be considered-- *so so* rating.

I don't like ten point scales unless the survey is administered over the phone the respondent on the other end can quickly comprehend the dialogue ***"on a scale of 1 to ten please rate……"*** and quickly adjust. However, on an online survey or mail survey the 1 to ten scale is cumbersome in design and overwhelming to the participant.

Alternative scales are a challenge to present to management. The management of most companies don't want to deviate from past surveys or from what others are doing. But we keep trying.

Resources

"The Optimal Length of Rating Scales to Maximize Reliability and Validity"

Brad Bortner, principal analyst with Forrester Research, "Best Practices: Why Customer Satisfaction Studies Fail"

Fred Reichheld, Fellow with Bain & Company, *The Ultimate Question*, p. 98-99

Fred Van Bennekom, Dr.B.A., Principal Great Brook Consulting

Dawes, John (2008), "Do Data Characteristics Change According to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales," International Journal of Market Research,

"Likert Scale." Encyclopædia Britannica. 2009. Encyclopædia Britannica Online. 29 Apr. 2009

---

[i] A net promoter score (NPS) is obtained by asking this single question and responses are categorized into three groups: Promoters, Passives and Detractors. By asking customers whether they would put their own credibility on the line by recommending the CLIENT to a friend, net promoter scores, can be indicators of loyalty and future behavior and, therefore, sales growth. Market Directions does not use the NPS score exclusively because of the limitations that a yes or no answer provides. NPS do not provide a holistic view of customers' perceptions or experiences, it does not tell you why customers are satisfied or dissatisfied or where the problems exist.
Fred Reichheld, Fellow with Bain & Company is the designer of this score.

[ii] *ARE THERE CENTRAL TENDENCY ERRORS IN SIMPLE RATING SCALES?:*
IMPLICATIONS FOR CROSS-NATIONAL/CULTURAL RESEARCH
Gerald Albaum, University of New Mexico
Robert D. Rogers, Market Metrics Southwest
Catherine Roster, University of New Mexico
Julie H. Yu, The Chinese University of Hong Kong