

5. Measurement and Scaling Techniques

Bipin Kumar Dubey

Assistant Professor,
Dept. of Sociology, Oriental College,
Patna, Bihar.

Abstract:

Questionnaires are a crucial research tool. Before creating a survey or questionnaire, the researcher must determine how the data will be measured for collection. The area of measuring that deals with instrument construction is called scaling. When selecting a suitable scaling method for a questionnaire and which ones to apply during data analysis, several criteria need to be taken into account. In order to offer guidance when choosing a scaling method for a survey or questionnaire, this document summarizes the various kinds of scaling methods. We shall introduce the idea of measurement and scaling in this chapter. Along with discussing the fundamental scales of measurement, it will also categorize and explain scaling methods that are comparative and no comparative. In creating the ideal question, it will also go over how to select a suitable scaling strategy. Also, an in-depth examination of the ideas of validity and dependability will be included. This paper will talk about. Techniques for Measuring and Scaling.

Keywords:

Measurement, Scaling, Techniques, Research Tools, Questionnaire, Comparative, No comparative, Nominal Scale, Ordinal Scale, Interval Scale, Ratio Scale.

5.1 Introduction:

An essential component of research and a crucial facet of study design is the measurement of variables. Individuals and businesses have different characteristics from one another as well as from one organization to another.

Humans possess a variety of physical and qualitative traits, such as intelligence, integrity, attitude, creativity, etc., as well as certain quantitative traits like height and weight. Physical attributes that are easily measurable also apply to business organizations, such as personnel, sales, profit, etc. [1]

The process of measuring involves putting numbers or symbols on an object's attributes. We measure certain properties of the thing rather than the object itself. Thus, in research, just the subjects' views, attitudes, and other pertinent traits are measured rather than the subjects themselves or the consumers. The normal assignment of numbers is done for two reasons. Numbers allow for the statistical analysis of the resultant data in the first place, and they also make it easier to communicate measurement results. The attributes' numerical assignments must be isomorphic.

In research methodology, variables are categorized and/or quantified using measurement scales.

- Monitoring and documenting the data gathered for research purposes is the process of measurement.
- In accordance with specific guidelines, the observations may be recorded using numbers or other symbols that correspond to an object's properties.
- Standardized and uniform application of the rules governing number assignment is necessary. This cannot alter with time or with items.
- The process of assigning objects to numbers or meanings in accordance with a rule is known as scaling.

5.2 Levels of Measurement:

Nominal, ordinal, interval, and ratio are the four levels of measurement that are typically explained in marketing research texts; therefore, their discussion in this article will be brief. But it's a crucial subject since the kind of scale used for measurements has an immediate impact on the statistical methods that may be applied to the analysis in a proper manner.

Four distinct measurement scales exist. One of the four scales can be used to characterize the data. The following four scale types exist:

- a. Nominal Scale
- b. Ordinal Scale
- c. Interval Scale
- d. Ratio Scale

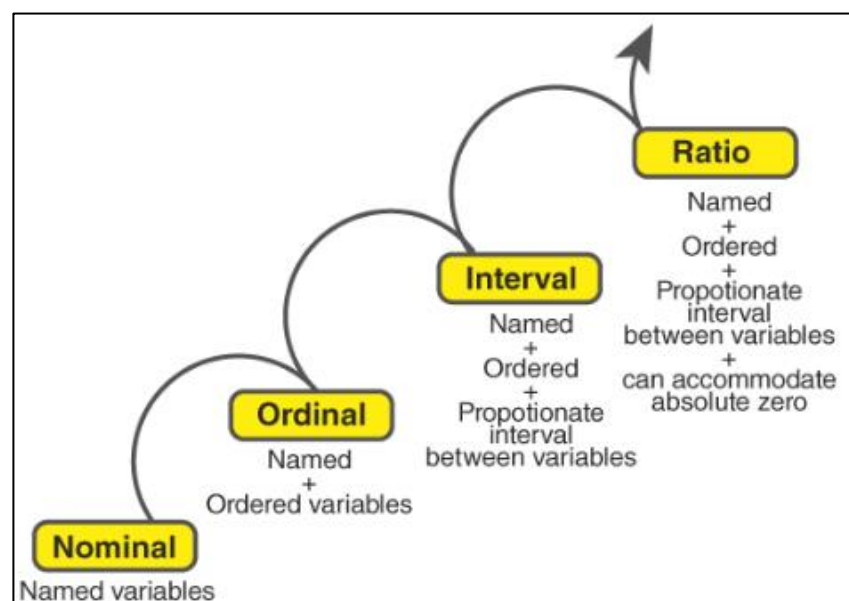


Figure 5.1: Levels of Measurements [2]

a. Nominal Scales:

This most basic scale of measurement divides people, businesses, goods, brands, and other entities into groups into which there is no inherent order. In fact, the term "categorical scale" is frequently applied to it. It does not position the object along a continuum; rather, it is a classification system.

It just entails counting how frequently cases are assigned to the different categories. If desired, nominal numbers can be provided to each category to indicate its label, as seen in the example below:

Table 5.1: An Example of a Nominal Scale:

Which of the following food items do you tend to buy at least once per month? (Please tick)		
Okra	Palm Oil	Milled Rice
Peppers	Prawns	Pasteurized milk

The numbers serve merely as labels and have no arithmetic features. Because this is just a collection of frequency counts, the mode is the only measure of average that can be applied. Data that has been gathered in the nominal form can be used for hypothesis testing. The Chi-square test seems to be the most likely. But it's important to remember that the Chi-square test measures the degree of association between two or more variables as well as whether they are related. It is unable to determine cause and effect; hence it cannot reveal anything about the nature of the link when one does exist.

b. Ordinal Scale:

The 2nd level of measurement, the ordinal scale, displays facts in an ordered and ranked manner without determining the degree of variance among them. Ordinal is a symbol for "order." Qualitative or categorical data are other names for ordinal data. It can be named, ranked, and categorized.

A. Characteristics of the Ordinal Scale:

- The ordinal scale shows the relative ranking of the variables
- It identifies and describes the magnitude of a variable
- Along with the information provided by the nominal scale, ordinal scales give the rankings of those variables
- The interval properties are not known
- The surveyors can quickly analyse the degree of agreement concerning the identified order of variables.

B. Example:

- ✚ Ranking of school students – 1st, 2nd, 3rd, etc.

- ✚ Ratings in restaurants
- ✚ Evaluating the frequency of occurrences
 - Very often
 - Often
 - Not often
 - Not at all
- ✚ Assessing the degree of agreement
 - Totally agree
 - Agree
 - Neutral
 - Disagree
 - Totally disagree

c. Interval Scale:

Since adjacent points on an interval scale have equal distances from one another, interval scales (definition of interval scale) go beyond the idea of ranking items in order. The Fahrenheit scale, for example, is an interval scale because there is no absolute zero point and every degree is equal.

This means that while we can multiply values and produce ratios (e.g., 100° is not twice as warm as 50°), we are unable to multiply or subtract degrees (e.g., 100° is 10° warmer than 90°). Behavioral scientists are willing to assume that most of their measures are interval scales because this allows the calculation of averages, mode, median, and mean, as well as range and standard deviation.

However, in an IQ scale, the intervals are not always equal (e.g., the difference between 105 and 110 is not really the same as between 80 and 85). This makes the underlying intent regarding the equal intervals crucial in determining whether a scale is considered interval or not.

Likert scales are typically regarded as interval scales, even though they are actually ordinal scales. Researchers can compute mean scores for this kind of agreement scale or attitudinal evaluation by treating it as an interval and then comparing them. Men's degree of agreement, for example, was 3.5 versus 4.1 for women, while first-time visitors' level of agreement was 3.3 versus 2.8.

d. Ratio Scale:

The ratio scale is one of the best methods of measurement available. Ratio scales are abstract number systems, much like interval scales. It has the extra benefit of starting from a set zero point and enabling measurement at appropriate intervals, order, classification, and distance. This is when the acquired ratio can be used to compare the two.

A health product manufacturer, for instance, conducted a poll to find out how common obesity is in a given area. The following survey instrument was made available:

Decide which category best describes your weight:

- Less than 40 kilograms
- 40-59 Kilograms
- 60-79 Kilograms
- 80-99 Kilograms
- 100-119 Kilograms
- 120 Kilograms and more

Table 5.2: The following table will better clarify the difference between all the four primary scaling techniques:

Particular	Nominal Scale	Ordinal Scale	Interval Scale	Ratio Scale
Characteristics	Description	Order	Distance	Description, Order, Distance & Origin
Sequential Arrangement	Not Applicable	Applicable	Applicable	Applicable
Fixed Zero Point	Not Applicable	Not Applicable	Not Applicable	Applicable
Multiplication & Division	Not Applicable	Not Applicable	Not Applicable	Applicable
Addition & Subtraction	Not Applicable	Not Applicable	Applicable	Applicable
Difference between Variables	Non-Measurable	Non-Measurable	Measurable	Measurable
Mean	Not Applicable	Not Applicable	Applicable	Applicable
Median	Not Applicable	Applicable	Applicable	Applicable
Mode	Applicable	Applicable	Applicable	Applicable

5.3 Types of Scaling Techniques:

Two groups of scaling strategies can be distinguished from each other in research: (a) comparative scales and (b) non-comparative scales. The respondent is asked to compare two objects in the comparative scaling exercise. One such question for the researcher to pose to the participants is whether they have a preference for detergent brands A or B.

Respondents just have to assess one item when using no comparative scaling, on the other hand. Their assessment is made without reference to the other item the researcher is looking at. Respondents that use a non-comparative scale choose any rating criteria that they believe is suitable. Continuous and itemized rating scales are examples of non-comparative approaches. These scaling methods are categorized in Figure 5.2. [3]

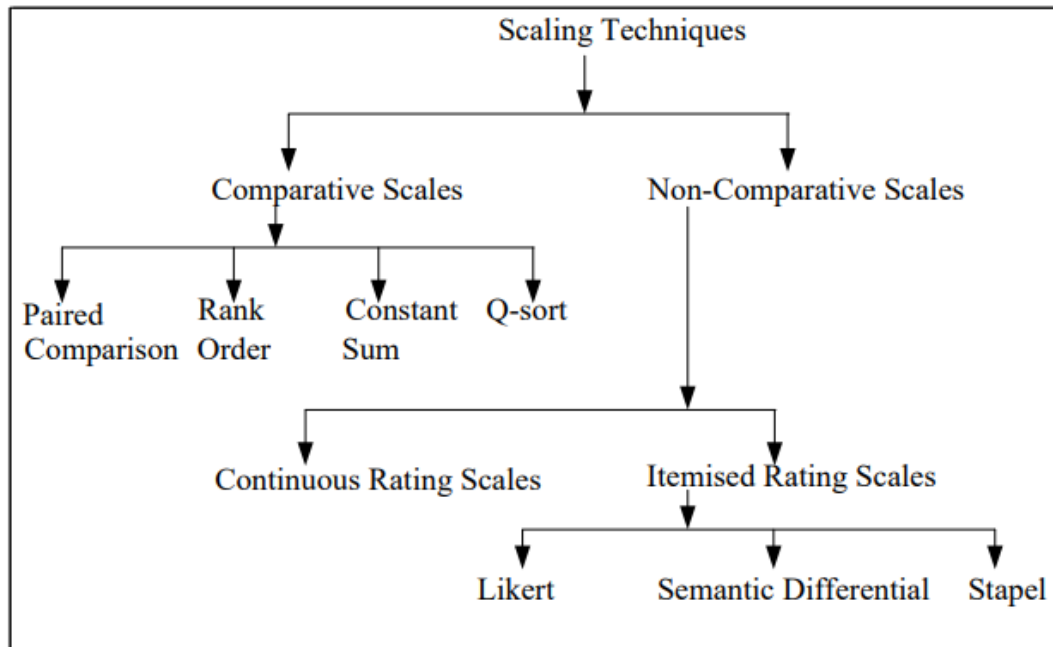


Figure 5.2: Scaling Techniques

5.3.1 Comparative Scales:

The four categories of scaling strategies that can be applied to the comparative scales are as follows: There are four different types of scales: (a) Q-sort, (b) Rank Order, (c) Constant Sum, and (d) Paired Comparison.

A. Paired Comparison Scale:

A respondent is asked to choose one thing (or rate between two objects at a time) based on a set of criteria in this comparative scaling technique, which presents two objects at once. The collected data are of an ordinal type. For instance, there are four varieties of cold beverages: Limca, Coke, Pepsi, and Sprite. Respondents may favor Coke over Sprite, Pepsi over Coke, etc. We can make the following six comparisons in total. [4]

- Coke–Pepsi
- Coke–Sprite
- Coke–Limca
- Pepsi–Sprite
- Pepsi–Limca
- Sprite–Limca

In general, with n brands we have $\frac{n(n-1)}{2}$ paired comparisons. The following is the data recording format using the paired comparisons.

Table 5.3:

Brand	Coke	Pepsi	Sprite	Limca
Coke	—	√		
Pepsi		—		
Sprite	√	√	—	
Limca	√	√	√	—
No. of times preferred	2	3	1	0

When a brand appears in a box with a \checkmark , it indicates that brand in that column was chosen over brand in that row. As you can hear in the recording above, there were two times where Coke was preferred over Sprite and Limca. Similar to how Pepsi was three times preferred above Coke, Sprite, and Limca in this instance. Therefore, adding up the \checkmark s in every column yields the total number of times a brand was chosen.

The paired comparison of data (assumed) for four different brands of cold drinks is provided in the following table. Table 5.4 [5]:

Brand	Coke	Pepsi	Sprite	Limca
Coke	—	0.90	0.64	0.14
Pepsi	0.10	—	0.32	0.02
Sprite	0.36	0.68	—	0.15
Limca	0.86	0.98	0.85	—

The percentage of respondents who favor "row" and "column" brands is indicated by the entries in the boxes. For instance, only 10% prefer Coke over Pepsi, and 90% prefer Pepsi over Coke, etc.

B. Rank Order:

Using this method, the respondent evaluates one item in comparison to others.

The respondent is asked to rank or arrange the various objects in the presence of a criterion.

Scaling of rank order is also ordinal.

With this strategy, only (n-1) scaling decisions need to be made.

As an illustration: The following soft drinks are asked to be rated by the respondent:

Drinks	Rank
Pepsi	2
Thumbs Up	1
Mountain dew	3
Mirinda	4

C. Constant Sum Scaling:

Respondents give each feature of a product a fixed number of points, say 100, to indicate how important they are.

The respondent gives an attribute zero points if it is unimportant.

An attribute is worth twice as many points if it is twice as significant as another attribute.

The total number of points is 100. Therefore, the scale's name. [6]

D. Q-Sort Scaling:

A method for selecting the best suitable objects from a vast number of provided variables is called Q-sort scaling. In order to create comparable heaps based on particular features, it places emphasis on rating the provided objects in descending order.

It is appropriate when there are between 60 and 90 objects, which is the most optimal range, but it is also appropriate when there are above 140 objects.

A marketing manager of a clothing manufacturing company, for instance, would rank the most effective marketing executives according to their track record, ability to generate sales income, growth, and dedication.

After 60 executives underwent Q-sort scaling, the marketing chief divided the executives into three piles according to their efficiency:

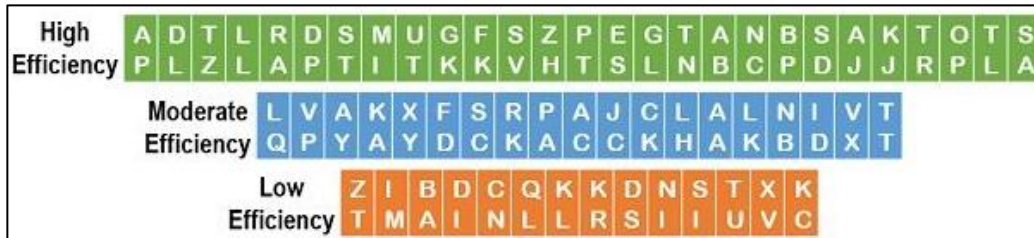


Figure 5.3: Q-sort scaling [7]

The employee names are indicated in the above graphic by their initials.

5.3.2 Non-Comparative Scales:

An individual product or object's performance is analyzed on various parameters using a non-comparative scale. Here are a few of its most prevalent varieties:

A. Continuous Rating Scales:

Respondents can freely arrange the object on the graphical rating scale by selecting its preferred location. In order to do this, a point that falls between two extreme criteria on a vertical or horizontal line is chosen and marked.

For instance, to determine the degree of client satisfaction with its new, cozy bedding, a mattress manufacturing company used a continuous rating system. The following possible interpretations of the response (expressed as variants here) are possible: [8]

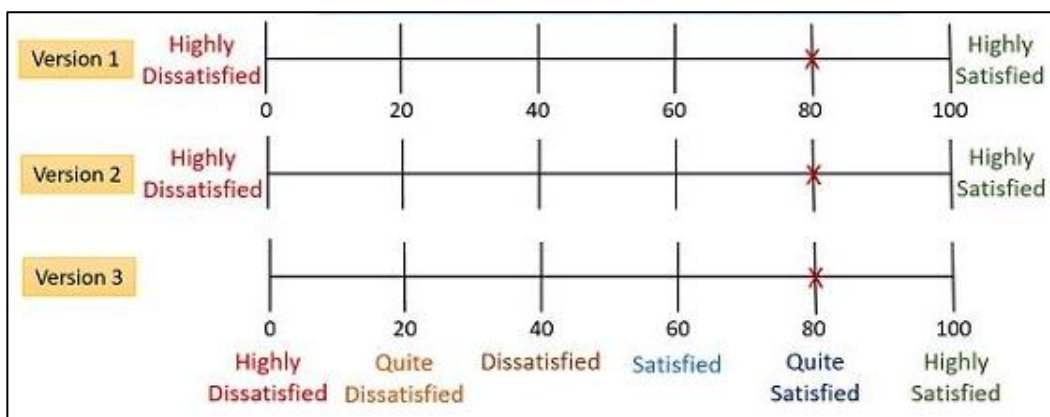


Figure 5.4: Continuous Rating Scales

An examination of a single product, comfortable bedding, is presented in a non-comparative manner in the diagram above. Consequently, it is evident that the clients are happy with the product and all of its characteristics.

B. Itemized Rating Scales:

Raters choose from a small number of categories that are arranged in some way. A typical number of categories is 2 to 11. A 3-point itemized scale is displayed on the right.

How eager are you to use an online service?

- Very interested
- Somewhat interested
- Not interested

a. Likert Scale:

One of the most popular scales in marketing research that measures agreement or disagreement is the Likert scale. Rensis Likert, the scale's creator, is honored by the scale's name. A sequence of statements regarding the stimulus objects are given to the respondent, who is then asked to indicate whether or not they agree with each statement. Five questions, ranging from "strongly disagree" to "strongly agree," make up a conventional Likert scale. In addition, researchers linked numbers to the Likert scale for statistical convenience. Figure 5.5 presents an illustration of a Likert scale. [9]

Q. Following are some statement relating too Newspaper X. Please indicate how strongly you agree or disagree with the statements using the scale provided by circling one of the numbers:

1 = Strongly disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5 = Strongly agree.

	Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
a. Newspaper X has high quality content	1	2	3	4	5
b. Newspaper X has the best writers	1	2	3	4	5
c. Newspaper X has a balance of local and national coverage	1	2	3	4	5
d. Newspaper X is my preferred newspaper	1	2	3	4	5

Figure 5.5: An Example of Likert Scale

Respondents find it simpler to respond to questions when the Likert scale is displayed. In addition, researchers employ a range of number systems other than 1 to 5, such as - 2 to + 2 or flipping the order of numbers from 5 to 1. The Likert scale analysis can be done item by item or based on the total score, which can be determined for each respondent by adding together all of the questions. The Likert scale is a useful tool for creating comparison systems. One way to compare the results would be to repeat the scale for Newspaper Y. Among the many benefits of the Likert scale are its ease of development and comprehension. Any type of survey administration can be used. Conversely, the Likert scale requires a lot of time to complete since participants must read each statement and give an answer that corresponds to it.

b. Stapel Scale:

The Stapel scale lacks a neutral point and has a single criterion in the midst of an even-numbered value range, from -5 to +5. Typically, the scale is displayed vertically. On the pre-established criterion, the responders are asked to select a specific number that best describes the stimuli object of concern. Figure 5.6 gives a thorough explanation of the Stapel scale. The graphic illustrates how the Stapel scale, which is represented by numbers, resembles the semantic differential scale in appearance. The Stapel scale data can be analyzed using the same methods as semantic differential scale data. The benefit of the Stapel scale is that, unlike the semantic differential scale, it does not require any sentences to attain bipolarity. The Stapel scale is the least used itemized evaluation system in the realm of marketing research. The primary reason for this is the belief that respondents won't be able to comprehend the scale and may provide a biased response. [10]

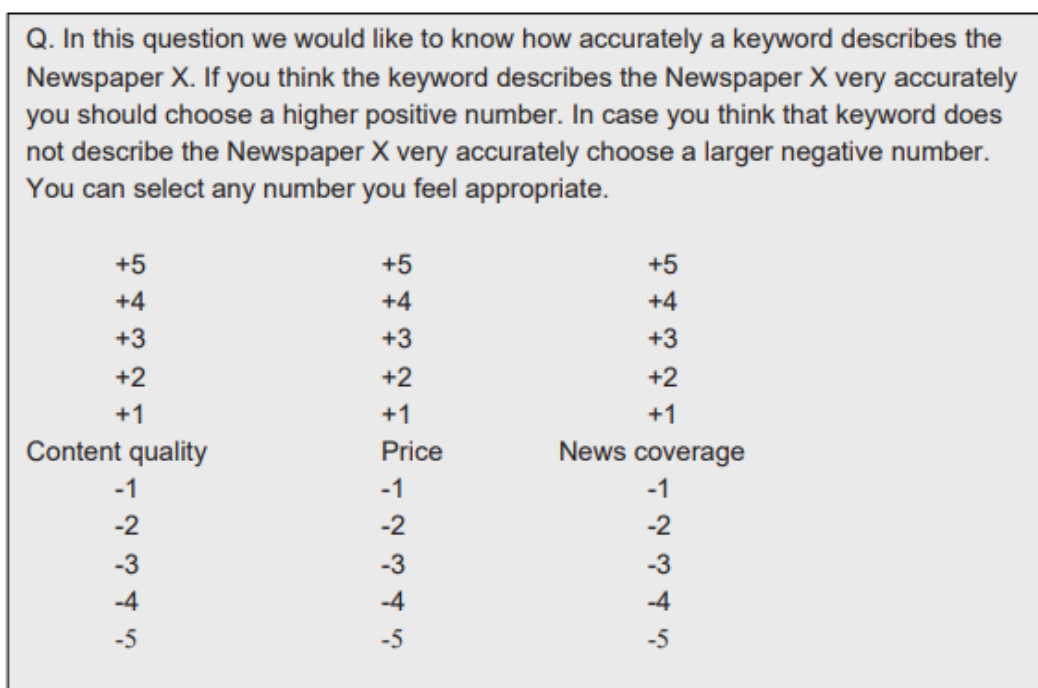


Figure 5.6: An example of Stapel scale

c. Semantic Differential Scale:

The respondent can mark on any of the seven points for each provided attribute of the object according to their personal preference on a bi-polar seven-point non-comparative rating scale. so illustrating the respondent's viewpoint or attitude regarding the item.

For instance, a well-known watch company used semantic differential scaling to ascertain how customers felt about their line of timepieces. This technique's visual depiction looks like this:

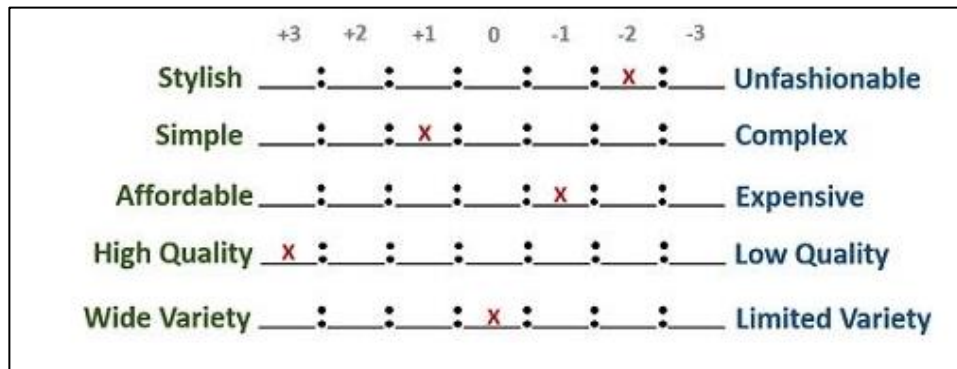


Figure 5.7: Semantic Differential Scale

The diagram above allows us to assess that although customers find the product to be of excellent quality, the brand has to put more effort into the way its timepieces are styled. [11]

5.4 Conclusion:

A clear and basic language is used to illustrate multiple types of measurement scales and scaling processes. With any luck, the scaling techniques module will enable you to properly structure your questionnaire in order to obtain the necessary primary data. To deal with data and carry out statistical analysis, one must have a basic understanding of scales of measures. Since the various measurement scales share some characteristics, it is crucial to appropriately assess the data to ascertain its measurement scale before selecting an analysis method. For the measurement of the same measuring scale, several scaling methods are possible. As a result, choosing a scaling technique specifically for research has no special significance.

5.5 References:

1. Daniles, E. and Lawford, J., 'The effect of order in the presentation of samples in paired comparison tests', *Journal of the Market Research Society* 16 (April 1974), 127–33.
2. Stapel, J., 'About 35 years of market research in the Netherlands', *Markonderzoek Kwartaalschrift* 2 (1969), 3–7
3. Spagna, G.J., 'Questionnaires: which approach do you use?', *Journal of Advertising Research* (February–March 1984), 67–70.

4. Bishop, G. F. (1987). Experiments with the middle response alternative in survey questions. *Public Opinion Quarterly*, 51, 220–232.
5. Golden, L., Brockett, P., Albaum, G., & Zatarain, J. (1992). The golden numerical comparative scale format for economical multi-object/multiattribute comparison questionnaires. *Journal of Official Statistics*, 8 (1), 77–86.
6. Haley, R., & Case, P. (1979, Fall). Testing thirteen attitude scales for agreement and brand determination. *Journal of Marketing*, 43, 20–32.
7. Hodge, D. R. & Gillespie, D. F. (2005). Phrase Completion Scales. In K. Kempf-Leonard (Editor). *Encyclopedia of Social Measurement*. (Vol. 3, pp. 53–62). San Diego: Academic Press.
8. Hodge, D. R. & Gillespie, D. F. (2003). Phrase Completions: An alternative to Likert scales. *Social Work Research*, 27(1), 45–55.
9. DeVellis, Robert F (2003), *Scale Development: Theory and Applications* (2nd ed.), London: SAGE Publications, ISBN 0-7619-2604-6, retrieved 11 August 2010
Paperback ISBN 0-7619-2605-4
10. Lodge, Milton (1981), *Magnitude Scaling: Quantitative Measurement of Opinions*, Beverly Hills & London: SAGE Publications, ISBN 0-8039-1747-3
11. McIver, John P. & Carmines, Edward G (1981), *Unidimensional Scaling*, Beverly Hills & London: SAGE Publications, ISBN 0-8039-1736-8, retrieved 11 August 2010.