
UNIT 5 MEASUREMENT AND SCALING TECHNIQUES

STRUCTURE

- 5.0 Objectives
- 5.1 Introduction
- 5.2 Measurement and Scaling
- 5.3 Issues in Attitude Measurement
- 5.4 Levels of Measurement Scales
- 5.5 Types of Scaling Techniques
 - 5.5.1 Comparative Scales
 - 5.5.2 Non-comparative Scales
- 5.6 Selection of an Appropriate Scaling Technique
- 5.7 Let Us Sum Up
- 5.8 Key Words
- 5.9 Answers to Self Assessment Exercises
- 5.10 Terminal Questions
- 5.11 Further Reading

5.0 OBJECTIVES

After studying this unit, you should be able to:

- 1 explain the concepts of measurement and scaling,
- 1 discuss four levels of measurement scales,
- 1 classify and discuss different scaling techniques, and
- 1 select an appropriate attitude measurement scale for your research problem.

5.1 INTRODUCTION

As we discussed earlier, the data consists of quantitative variables like price, income, sales etc., and qualitative variables like knowledge, performance, character etc. The qualitative information must be converted into numerical form for further analysis. This is possible through measurement and scaling techniques. A common feature of survey based research is to have respondent's feelings, attitudes, opinions, etc. in some measurable form. For example, a bank manager may be interested in knowing the opinion of the customers about the services provided by the bank. Similarly, a fast food company having a network in a city may be interested in assessing the quality and service provided by them. As a researcher you may be interested in knowing the attitude of the people towards the government announcement of a metro rail in Delhi. In this unit we will discuss the issues related to measurement, different levels of measurement scales, various types of scaling techniques and also selection of an appropriate scaling technique.

5.2 MEASUREMENT AND SCALING

Before we proceed further it will be worthwhile to understand the following two terms: (a) Measurement, and (b) Scaling.

- a) **Measurement:** Measurement is the process of observing and recording the observations that are collected as part of research. The recording of the

observations may be in terms of numbers or other symbols to characteristics of objects according to certain prescribed rules. The respondent's, characteristics are feelings, attitudes, opinions etc. For example, you may assign '1' for Male and '2' for Female respondents. In response to a question on whether he/she is using the ATM provided by a particular bank branch, the respondent may say 'yes' or 'no'. You may wish to assign the number '1' for the response yes and '2' for the response no. We assign numbers to these characteristics for two reasons. First, the numbers facilitate further statistical analysis of data obtained. Second, numbers facilitate the communication of measurement rules and results. The most important aspect of measurement is the specification of rules for assigning numbers to characteristics. The rules for assigning numbers should be standardised and applied uniformly. This must not change over time or objects.

- b) **Scaling:** Scaling is the assignment of objects to numbers or semantics according to a rule. In scaling, the objects are text statements, usually statements of attitude, opinion, or feeling. For example, consider a scale locating customers of a bank according to the characteristic "agreement to the satisfactory quality of service provided by the branch". Each customer interviewed may respond with a semantic like 'strongly agree', or 'somewhat agree', or 'somewhat disagree', or 'strongly disagree'. We may even assign each of the responses a number. For example, we may assign strongly agree as '1', agree as '2' disagree as '3', and strongly disagree as '4'. Therefore, each of the respondents may assign 1, 2, 3 or 4.

5.3 ISSUES IN ATTITUDE MEASUREMENT

When a researcher is interested in measuring the attitudes, feelings or opinions of respondents he/she should be clear about the following:

- a) What is to be measured?
- b) Who is to be measured?
- c) The choices available in data collection techniques

The first issue that the researcher must consider is 'what is to be measured'? The definition of the problem, based on our judgments or prior research indicates the concept to be investigated. For example, we may be interested in measuring the performance of a fast food company. We may require a precise definition of the concept on how it will be measured. Also, there may be more than one way that we can measure a particular concept. For example, in measuring the performance of a fast food company we may use a number of measures to indicate the performance of the company. We may use sales volume in terms of value of sales or number of customers or spread of network of the company as measures of performance. Further, the measurement of concepts requires assigning numbers to the attitudes, feelings or opinions. The key question here is that on what basis do we assign the numbers to the concept. For example, the task is to measure the agreement of customers of a fast food company on the opinion of whether the food served by the company is tasty, we create five categories: (1) strongly agree, (2) agree, (3) undecided, (4) disagree, (5) strongly disagree. Then we may measure the response of respondents. Suppose if a respondent states 'disagree' with the statement that 'the food is tasty', the measurement is 4.

The second important issue in measurement is that, who is to be measured? That means who are the people we are interested in. The characteristics of the people such as age, sex, education, income, location, profession, etc. may

have a bearing on the choice of measurement. The measurement procedure must be designed keeping in mind the characteristics of the respondents under consideration.

The third issue in measurement is the choice of the data collection techniques. In Unit 3, you have already learnt various methods of data collection. Normally, questionnaires are used for measuring attitudes, opinions or feelings.

5.4 LEVELS OF MEASUREMENT SCALES

The level of measurement refers to the relationship among the values that are assigned to the attributes, feelings or opinions for a variable. For example, the variable ‘whether the taste of fast food is good’ has a number of attributes, namely, very good, good, neither good nor bad, bad and very bad. For the purpose of analysing the results of this variable, we may assign the values 1, 2, 3, 4 and 5 to the five attributes respectively. The level of measurement describes the relationship among these five values. Here, we are simply using the numbers as shorter placeholders for the lengthier text terms. We don’t mean that higher values mean ‘more’ of something or lower values mean ‘less’ of something. We don’t assume that ‘good’ which has a value of 2 is twice of ‘very good’ which has a value of 1. We don’t even assume that ‘very good’ which is assigned the value ‘1’ has more preference than ‘good’ which is assigned the value ‘2’. We simply use the values as a shorter name for the attributes, opinions, or feelings. The assigned values of attributes allow the researcher more scope for further processing of data and statistical analysis.

Typically, there are four levels of measurement scales or methods of assigning numbers: (a) Nominal scale, (b) Ordinal scale, (c) Interval scale, and (d) Ratio scale.

- a) **Nominal Scale** is the crudest among all measurement scales but it is also the simplest scale. In this scale the different scores on a measurement simply indicate different categories. The nominal scale does not express any values or relationships between variables. For example, labelling men as ‘1’ and women as ‘2’ which is the most common way of labelling gender for data recording purpose does not mean women are ‘twice something or other’ than men. Nor it suggests that men are somehow ‘better’ than women. Another example of nominal scale is to classify the respondent’s income into three groups: the highest income as group 1. The middle income as group 2, and the low-income as group 3. The nominal scale is often referred to as a categorical scale. The assigned numbers have no arithmetic properties and act only as labels. The only statistical operation that can be performed on nominal scales is a frequency count. We cannot determine an average except mode.

In designing and developing a questionnaire, it is important that the response categories must include all possible responses. In order to have an exhaustive number of responses, you might have to include a category such as ‘others’, ‘uncertain’, ‘don’t know’, or ‘can’t remember’ so that the respondents will not distort their information by forcing their responses in one of the categories provided. Also, you should be careful and be sure that the categories provided are mutually exclusive so that they do not overlap or get duplicated in any way.

- b) **Ordinal Scale** involves the ranking of items along the continuum of the characteristic being scaled. In this scale, the items are classified according to

whether they have more or less of a characteristic. For example, you may wish to ask the TV viewers to rank the TV channels according to their preference and the responses may look like this as given below:

TV Channel	Viewers preferences
Doordarshan-1	1
Star plus	2
NDTV News	3
Aaaj Tak TV	4

The main characteristic of the ordinal scale is that the categories have a logical or ordered relationship. This type of scale permits the measurement of degrees of difference, (that is, 'more' or 'less') but not the specific amount of differences (that is, how much 'more' or 'less'). This scale is very common in marketing, satisfaction and attitudinal research.

Another example is that a fast food home delivery shop may wish to ask its customers:

<p>How would you rate the service of our staff?</p> <p>(1) <i>Excellent</i> • (2) <i>Very Good</i> • (3) <i>Good</i> • (4) <i>Poor</i> • (5) <i>Worst</i> •</p>

Suppose respondent X gave the response 'Excellent' and respondent Y gave the response 'Good', we may say that respondent X thought that the service provided better than respondent Y to be thought. But we don't know how much better and even we can't say that both respondents have the same understanding of what constitutes 'good service'.

In marketing research, ordinal scales are used to measure relative attitudes, opinions, and preferences. Here we rank the attitudes, opinions and preferences from best to worst or from worst to best. However, the amount of difference between the ranks cannot be found out. Using ordinal scale data, we can perform statistical analysis like Median and Mode, but not the Mean.

c) **Interval Scale** is a scale in which the numbers are used to rank attributes such that numerically equal distances on the scale represent equal distance in the characteristic being measured. An interval scale contains all the information of an ordinal scale, but it also one allows to compare the difference/distance between attributes. For example, the difference between '1' and '2' is equal to the difference between '3' and '4'. Further, the difference between '2' and '4' is twice the difference between '1' and '2'. However, in an interval scale, the zero point is arbitrary and is not true zero. This, of course, has implications for the type of data manipulation and analysis. We can carry out on data collected in this form. It is possible to add or subtract a constant to all of the scale values without affecting the form of the scale but one cannot multiply or divide the values. Measuring temperature is an example of interval scale. We cannot say 40°C is twice as hot as 20°C. The reason for this is that 0°C does not mean that there is no temperature, but a relative point on the Centigrade Scale. Due to lack of an absolute zero point, the interval scale does not allow the conclusion that 40°C is twice as hot as 20°C.

Interval scales may be either in numeric or semantic formats. The following are two more examples of interval scales one in numeric format and another in semantic format.

i) **Example of Interval Scale in Numeric Format**

<i>Food supplied is:</i>						<i>Indicate your score on the concerned blank and circle the appropriate number on each line.</i>
Fresh	1	2	3	4	5	
Tastes good	1	2	3	4	5	
Value for money	1	2	3	4	5	
Attractive packaging	1	2	3	4	5	
Prompt time delivery	1	2	3	4	5	

ii) **Example of Interval Scale in Semantic Format**

Please indicate your views on the food supplied by XXX Fast Food Shop by scoring them on a five points scale from 1 to 5 (that is, 1=Excellent, 2=Very Good, 3=Good, 4=Poor, 5=Worst). Indicate your views by ticking the appropriate responses below:

Food supplied is:	Excellent	Very Good	Good	Poor	Worst
Fresh					
Tastes good					
Value for money					
Attractive packaging					
Prompt time delivery					

The interval scales allow the calculation of averages like Mean, Median and Mode and dispersion like Range and Standard Deviation.

- d) **Ratio Scale** is the highest level of measurement scales. This has the properties of an interval scale together with a fixed (absolute) zero point. The absolute zero point allows us to construct a meaningful ratio. Examples of ratio scales include weights, lengths and times. In the marketing research, most counts are ratio scales. For example, the number of customers of a bank's ATM in the last three months is a ratio scale. This is because you can compare this with previous three months. Ratio scales permit the researcher to compare both differences in scores and relative magnitude of scores. For example, the difference between 10 and 15 minutes is the same as the difference between 25 and 30 minutes and 30 minutes is twice as long as 15 minutes. Most financial research that deals with rupee values utilizes ratio scales. However, for most behavioural research, interval scales are typically the highest form of measurement. Most statistical data analysis procedures do not distinguish between the interval and ratio properties of the measurement scales and it is sufficient to say that all the statistical operations that can be performed on interval scale can also be performed on ratio scales.

Now you must be wondering why you should know the level of measurement. Knowing the level of measurement helps you to decide on how to interpret the data. For example, when you know that a measure is nominal then you know that the numerical values are just short codes for longer textual names. Also, knowing the level of measurement helps you to decide what statistical analysis is appropriate on the values that were assigned. For example, if you know that a measure is nominal, then you would not need to find mean of the data values or perform a t-test on the data. (t-test will be discussed in Unit-16 in the course).

It is important to recognise that there is a hierarchy implied in the levels of measurement. At lower levels of measurement, assumptions tend to be less restrictive and data analyses tend to be less sensitive. At each level up the hierarchy, the current level includes all the qualities of the one below it and adds something new. In general, it is desirable to have a higher level of measurement (that is, interval or ratio) rather than a lower one (that is, nominal or ordinal).

Self Assessment Exercise-A

- 1) The main difference between interval scale and the ratio scale in terms of their properties is:

.....
.....
.....

- 2) Why should the researcher know the level of measurement?

.....
.....
.....

- 3) What are the main statistical limitations of nominal scale?

.....
.....
.....

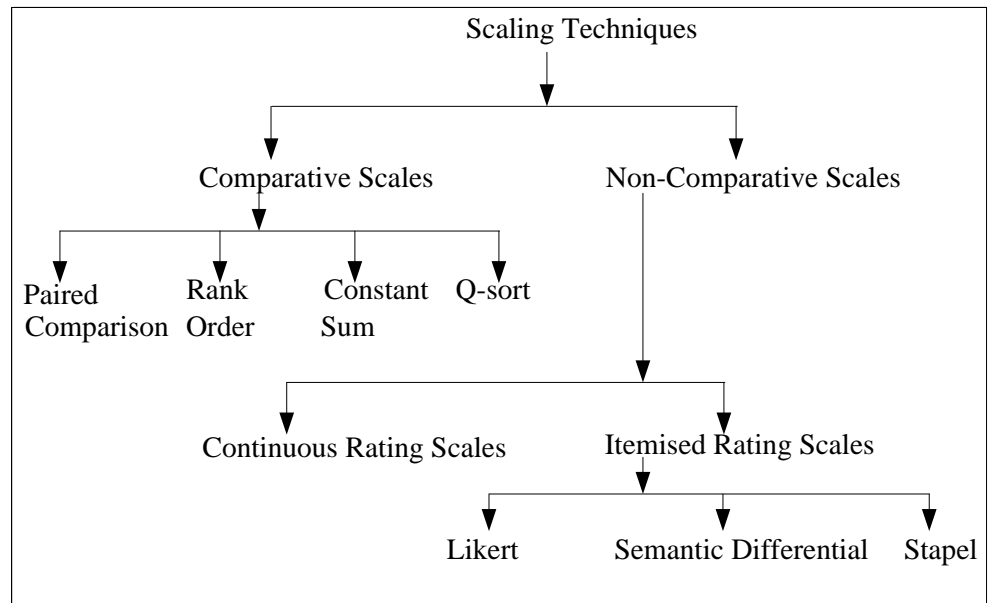
- 4) Indicate whether the following measures are nominal, ordinal, interval or ratio scales?

- a) Social status of a respondent:
b) Stock market prices:
c) The ranks obtained by students:
d) The Fahrenheit scale for measuring temperature:

5.5 TYPES OF SCALING TECHNIQUES

The various types of scaling techniques used in research can be classified into two categories: (a) comparative scales, and (b) Non-comparative scales. In **comparative scaling**, the respondent is asked to compare one object with another. For example, the researcher can ask the respondents whether they prefer brand A or brand B of a detergent. On the other hand, in non-comparative scaling respondents need only evaluate a single object. Their evaluation is independent of the other object which the researcher is studying. Respondents using a non-comparative scale employ whatever rating standard seems appropriate to them. Non-comparative techniques consist of continuous and itemized rating scales. Figure 5.1 shows the classification of these scaling techniques.

Figure 5.1: Scaling Techniques



5.5.1 Comparative Scales

The comparative scales can further be divided into the following four types of scaling techniques: (a) Paired Comparison Scale, (b) Rank Order Scale, (c) Constant Sum Scale, and (d) Q-sort Scale.

- a) **Paired Comparison Scale:** This is a comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object (rate between two objects at a time) according to some criterion. The data obtained are ordinal in nature. For example, there are four types of cold drinks - Coke, Pepsi, Sprite, and Limca. The respondents can prefer Pepsi to Coke or Coke to Sprite, etc. In all we can have the following six comparisons.

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 5.1

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

A \surd in a particular box means that the brand in that column was preferred over the brand in the corresponding row. In the above recording, Coke was preferred over Sprite, Coke over Limca, in this case the number of times coke preferred was 2 times. Similarly, Pepsi over Coke, Pepsi over Sprite, Pepsi over Limca, in this case Pepsi was 3 time preferred. Thus, the number of times a brand was preferred is obtained by summing the \surd s in each column.

The following table gives paired comparison of data (assumed) for four brands of cold drinks.

Table 5.2

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 entries in the boxes represent the proportion of respondents preferring ‘column brand’ and to ‘row’ brand. For example, 90% prefer Pepsi to Coke and only 10% prefer Coke to Pepsi, etc.

Paired comparison is useful when the number of brands are limited, since it requires direct comparison and overt choice. One of the disadvantages of paired comparison scale is violation of the assumption of transitivity may occur. For example, in our example (Table 5.1) the respondent preferred Coke 2 times, Pepsi 3 times, Sprite 1 time, and Limca 0 times. That means, preference-wise, Pepsi >Coke, Coke >Sprite, and Sprite >Limca. However, the number of times Sprite was preferred should not be that of Coke. In other words, if A>B and B >C then C >A should not be possible. Also, the order in which the objects are presented may bias the results. The number of items/brands for comparison should not be too many. As the number of items increases, the number of comparisons increases geometrically. If the number of comparisons is too large, the respondents may become fatigued and no longer be able to carefully discriminate among them. The other limitation of paired comparison is that this scale has little resemblance to the market situation, which involves selection from multiple alternatives. Also, respondents may prefer one item over certain others, but they may not like it in an absolute sense.

- b) **Rank Order Scale:** This is another type of comparative scaling technique in which respondents are presented with several items simultaneously and asked to rank them in the order of priority. This is an ordinal scale that describes the favoured and unfavoured objects, but does not reveal the distance between the objects. For example, if you are interested in ranking the preference of some selected brands of cold drinks, you may use the following format for recording the responses.

Table 5.3: Preference of cold drink brands using rank order scaling

Instructions: Rank the following brands of cold drinks in order of preference. Begin by picking out the one brand you like most and assign it a number 1. Then find the second most preferred brand and assign it a number 2. Continue this procedure until you have ranked all the brands of cold drinks in order of preference. The least preferred brand should be assigned a rank of 4. Also remember no two brands receive the same rank order.

Format:

Brand	Rank
(a) Coke	3
(b) Pepsi	1
(c) Limca	2
(d) Sprite	4

Like paired comparison, the rank order scale, is also comparative in nature. The resultant data in rank order is ordinal data. This method is more realistic in obtaining the responses and it yields better results when direct comparison are required between the given objects. The major disadvantage of this technique is that only ordinal data can be generated.

- c) **Constant Sum Scale:** In this scale, the respondents are asked to allocate a constant sum of units such as points, rupees, or chips among a set of stimulus objects with respect to some criterion. For example, you may wish to determine how important the attributes of price, fragrance, packaging, cleaning power, and lather of a detergent are to consumers. Respondents might be asked to divide a constant sum to indicate the relative importance of the attributes using the following format.

Table 5.4: Importance of detergent attributes using a constant sum scale

Instructions: Between attributes of detergent please allocate 100 points among the attributes so that your allocation reflects the relative importance you attach to each attribute. The more points an attribute receives, the more important the attribute is. If an attribute is not at all important, assign it zero points. If an attribute is twice as important as some other attribute, it should receive twice as many points.

Format:

Attribute	Number of Points
(a) Price	50
(b) Fragrance	05
(c) Packaging	10
(d) Cleaning Power	30
(e) Lather	05
Total Points	100

“If an attribute is assigned a higher number of points, it would indicate that the attribute is more important.” From the above Table, the price of the detergent is

the most important attribute for the consumers followed by cleaning power, packaging. Fragrance and lather are the two attributes that the consumers cared about the least but preferred equally.” The advantage of this technique is saving time. However, there are two main disadvantages. The respondents may allocate more or fewer points than those specified. The second problem is rounding off error if too few attributes are used and the use of a large number of attributes may be too taxing on the respondent and cause confusion and fatigue.

- d) **Q-Sort Scale:** This is a comparative scale that uses a rank order procedure to sort objects based on similarity with respect to some criterion. The important characteristic of this methodology is that it is more important to make comparisons among different responses of a respondent than the responses between different respondents. Therefore, it is a comparative method of scaling rather than an absolute rating scale. In this method the respondent is given statements in a large number for describing the characteristics of a product or a large number of brands of a product. For example, you may wish to determine the preference from among a large number of magazines. The following format shown in Table 5.5 may be given to a respondent to obtain the preferences.

Table 5.5: Preference of Magazines Using Q-Sort Scale Procedure

<p>Instructions: The bag given to you contain pictures of 90 magazines. Please choose 10 magazines you ‘prefer most’, 20 magazines you ‘like’, 30 magazines to which you are ‘neutral (neither like nor dislike)’, 20 magazines you ‘dislike’, and 10 magazines you ‘prefer least’. Please list the sorted magazine names in the respective columns of the form provided to you.</p> <p>Format:</p>				
Prefer Most	Like	Neutral	Dislike	Prefer Least
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
(10)	_____	_____	_____	(10)
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	(20)	_____	(20)	

		(30)		

Note that the number of responses to be sorted should not be less than 60 or not more than 140. A reasonable range is 60 to 90 responses that result in a normal or quasi-normal distribution. This method is faster and less tedious than paired comparison measures. It also forces the subject to conform to quotas at each point of scale so as to yield a quasi-normal distribution. The utility of Q-sort in marketing research is to derive clusters of individuals who display similar preferences, thus representing unique market segments.

5.5.2 Non-Comparative Scales

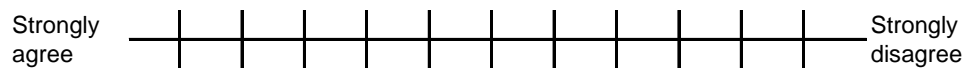
The non-comparative scaling techniques can be further divided into: (a) Continuous Rating Scale, and (b) Itemised Rating Scale.

a) Continuous Rating Scales

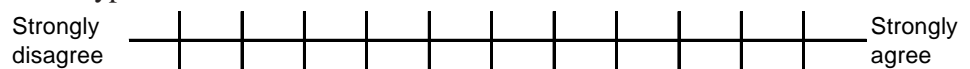
It is very simple and highly useful. In continuous rating scale, the respondent's rate the objects by placing a mark at the appropriate position on a continuous line that runs from one extreme of the criterion variable to the other. Examples of continuous rating scale are given below:

Question: How would you rate the TV advertisement as a guide for buying?

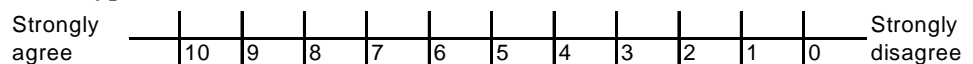
Scale Type A



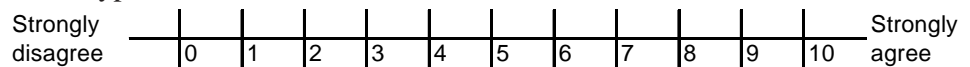
Scale Type B



Scale Type C



Scale Type D


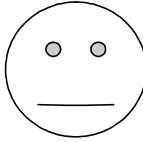



When scale type A and B are used, the respondents score is determined either by dividing the line into as many categories as desired and assigning the respondent a score based on the category into which his/her mark falls, or by measuring distance, in millimeters, centimeters, or inches from either end of the scale. Which ever of the above continuous scale is used, the results are normally analysed as interval scaled.

b) Itemised Rating Scales

Itemised rating scale is a scale having numbers or brief descriptions associated with each category. The categories are ordered in terms of scale position and the respondents are required to select one of the limited number of categories that best describes the product, brand, company, or product attribute being rated. Itemised rating scales are widely used in marketing research.

The itemised rating scales can be in the form of : (a) graphic, (b) verbal, or (c) numeric as shown below:

Itemised Graphic Scale	Itemised Verbal Scale	Itemised Numeric Scale
 Favourable	Completely satisfied	-5
	Somewhat satisfied	-4
		-3
		-2
 Indifferent	Neither satisfied nor dissatisfied	-1
		0
	Somewhat dissatisfied	+1
		+2
 Unfavourable	Completely dissatisfied	+3
		+4
		+5

Some rating scales may have only two response categories such as : agree and disagree. Inclusion of more response categories provides the respondent more flexibility in the rating task. Consider the following questions:

- How often do you visit the supermarket located in your area of residence?
 - Never, • Rarely, • Sometimes, • Often, • Very often
- In your case how important is the price of brand X shoes when you buy them?
 - Very important, • Fairly important, • Neutral, • Not so important

Each of the above category scales is a more sensitive measure than a scale with only two responses since they provide more information.

Wording is an extremely important factor in the usefulness of itemised scales. Table 5.6 shows some common wordings for categories used in itemised scales.

Table 5.6: Some common words for categories used in Itemised Rating scales

<i>Quality:</i>				
Excellent	Good	Not decided	Poor	Worst
Very Good	Good	Neither good nor bad	Fair	Poor
<i>Importance:</i>				
Very Important	Fairly important	Neutral	Not so important	Not at all important
<i>Interest:</i>				
Very interested	Somewhat interested	Neither interested nor disinterested	Somewhat uninterested	Not very interested
<i>Satisfaction:</i>				
Completely satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Completely dissatisfied
<i>Frequency:</i>				
All of the time	Very often	Often	Sometimes	Hardly ever
Very often	Often	Sometimes	Rarely	Never
<i>Truth:</i>				
Very true	Somewhat true	Not very true	Not at all true	
<i>Purchase Interest:</i>				
Definitely will buy	Probably will buy	Probably will not buy	Definitely will not buy	
<i>Level of Agreement:</i>				
Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
<i>Dependability:</i>				
Completely dependable	Somewhat dependable	Not very dependable	Not at all dependable	
<i>Style:</i>				
Very stylish	Somewhat stylish	Not very stylish	Completely unstylish	
<i>Cost:</i>				
Extremely expensive	Expensive	Neither expensive nor inexpensive	Slightly inexpensive	Very inexpensive
<i>Ease of use:</i>				
Very ease to use	Somewhat easy to use	Not very easy to use	Difficult to use	
<i>Modernity:</i>				
Very modern	Somewhat modern	Neither modern nor old-fashioned	Somewhat old fashioned	Very old fashioned
<i>Alert:</i>				
Very alert	Alert	Not alert	Not at all alert	

In this section we will discuss three itemised rating scales, namely (a) Likert scale, (b) Semantic Differential Scale, and (c) Stapel Scale.

- a) **Likert Scale:** In business research, the Likert scale, developed by Rensis Likert, is extremely popular for measuring attitudes, because, the method is simple to administer. With the Likert scale, the respondents indicate their own attitudes by checking how strongly they agree or disagree with carefully worded statements that range from very positive to very negative towards the attitudinal object. Respondents generally choose from five alternatives (say strongly agree, agree, neither agree nor disagree, disagree, strongly disagree).

Consider the following example of a study or measuring attitudes towards cricket.

	Strongly agree	Agree	Not sure	Disagree	Strongly disagree
It is more fun to play a tough, competitive cricket match than to play an easy one.	5	4	3	2	1

To measure the attitude, the researchers assign weights or scores to the alternative responses. In the above example the scores 5 to 1 are assigned to the responses. Strong agreement of the respondent indicates the most favourable attitudes on the statement, and the score 5 is assigned to it. On the other hand, strong disagreement of the respondent indicates the most unfavourable attitude on the statement, and the score 1 is assigned to it. If a negative statement towards the object is given, the corresponding scores would be reversed. In this case, the response 'strongly agree' will get a score of 1 and the response 'strongly disagree' will get a score of 5.

A Likert scale may include a number of items or statements. Each statement is assumed to represent an aspect of an attitudinal domain. For example, Table 5.7 shows the items in a Likert Scale to measure opinions on food products.

Table 5.7: A Likert Scale for studying opinions on food products.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
If the price of raw materials fall, firms too should reduce the price of the food products.	1	2	3	4	5
There should be uniform price through out the country for food products	1	2	3	4	5
The food companies should concentrate more on keeping hygiene while manufacturing food products.	1	2	3	4	5
The expiry dates should be printed on the food products before they are delivered to consumers in the market.	1	2	3	4	5
There should be government regulations on the firms in keeping acceptable quality and on the prices	1	2	3	4	5
Now-a-days most food companies are concerned only with profit making rather than taking care of quality.	1	2	3	4	5

Each respondent is asked to circle his opinion on a score against each statement. The final score for the respondent on the scale is the sum of their ratings for all the items. The very purpose of Likert's Scale is to ensure the final items evoke a wide response and discriminate among those with positive and negative attitudes. Items that are poor (because they lack clarity or elicit mixed response patterns) are detected from the final statement list. This will ensure us to discriminate between high positive scores and high negative scores. However, many business researchers do not follow this procedure and you may not be in a position to distinguish between high positive scores and high negative scores because all scores look alike. Hence a disadvantage of the Likert Scale is that it is difficult to know what a single summated score means. Many patterns of response to the various statements can produce the same total score. The other disadvantage of Likert Scale is that it takes longer time to complete than other itemised rating scales because respondents have to read each statement. Despite the above disadvantages, this scale has several advantages. It is easy to construct, administer and use.

- b) **Semantic Differential Scale:** This is a seven point rating scale with end points associated with bipolar labels (such as good and bad, complex and simple) that have semantic meaning. The Semantic Differential scale is used for a variety of purposes. It can be used to find whether a respondent has a positive or negative attitude towards an object. It has been widely used in comparing brands, products and company images. It has also been used to develop advertising and promotion strategies and in a new product development study. Look at the following Table, for examples of Semantic Differential Scale.

Table 5.8: Examples of Semantic Differential Scale

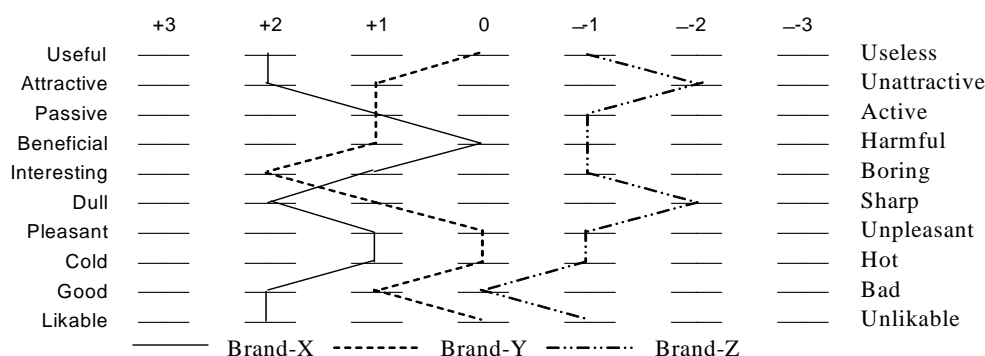
Modern	—	—	—	—	—	—	—	Old-fashioned
Good	—	—	—	—	—	—	—	Bad
Clean	—	—	—	—	—	—	—	Dirty
Important	—	—	—	—	—	—	—	Unimportant
Expensive	—	—	—	—	—	—	—	Inexpensive
Useful	—	—	—	—	—	—	—	Useless
Strong	—	—	—	—	—	—	—	Weak
Quick	—	—	—	—	—	—	—	Slow

In the Semantic Differential scale only extremes have names. The extreme points represent the bipolar adjectives with the central category representing the neutral position. The in between categories have blank spaces. A weight is assigned to each position on the scale. The weights can be such as +3, +2, +1, 0, -1, -2, -3 or 7,6,5,4,3,2,1. The following is an example of Semantic Differential Scale to study the experience of using a particular brand of body lotion.

	In my experience, the use of body lotion of Brand-X was:							
	+3	+2	+1	0	-1	-2	-3	
Useful	_____	_____	_____	_____	_____	_____	_____	Useless
Attractive	_____	_____	_____	_____	_____	_____	_____	Unattractive
Passive	_____	_____	_____	_____	_____	_____	_____	Active
Beneficial	_____	_____	_____	_____	_____	_____	_____	Harmful
Interesting	_____	_____	_____	_____	_____	_____	_____	Boring
Dull	_____	_____	_____	_____	_____	_____	_____	Sharp
Pleasant	_____	_____	_____	_____	_____	_____	_____	Unpleasant
Cold	_____	_____	_____	_____	_____	_____	_____	Hot
Good	_____	_____	_____	_____	_____	_____	_____	Bad
Likable	_____	_____	_____	_____	_____	_____	_____	Unlikable

In the semantic Differential scale, the phrases used to describe the object form a basis for attitude formation in the form of positive and negative phrases. The negative phrase is sometimes put on the left side of the scale and sometimes on the right side. This is done to prevent a respondent with a positive attitude from simply checking the left side and a respondent with a negative attitude checking on the right side without reading the description of the words.

The respondents are asked to check the individual cells depending on the attitude. Then one could arrive at the average scores for comparisons of different objects. The following Figure shows the experiences of 100 consumers on 3 brands of body lotion.



In the above example, first the individual respondent scores for each dimension are obtained and then the average scores of all 100 respondents, for each dimension and for each brand were plotted graphically. The maximum score possible for each brand is +30 and the minimum score possible for each brand is -30. Brand-X has score +14. Brand-Y has score +7, and Brand-Z has score -11. From the scale we can identify which phrase needs improvement for each Brand. For example, Brand-X needs to be improved upon benefits and Brand-Y on pleasantness, coldness and likeability. Brand Z needs to be improved on all the attributes.

- c) **Staple Scale:** The Stapel scale was originally developed to measure the direction and intensity of an attitude simultaneously. Modern versions of the Stapel scale place a single adjective as a substitute for the Semantic differential when it is difficult to create pairs of bipolar adjectives. The modified Stapel scale places a single adjective in the centre of an even number of numerical values (say, +3, +2, +1, 0, -1, -2, -3). This scale measures how close to or how distant from the adjective a given stimulus is perceived to be. The following is an example of a Stapel scale.

Instructions: Select a plus number for words that you think describe personnel banking of a bank accurately. The more accurately you think the word describes the bank, the larger the plus number you should choose. Select a minus number for words you think do not describe the bank accurately. The less accurately you think the word describes the bank, the larger the minus number you should choose.

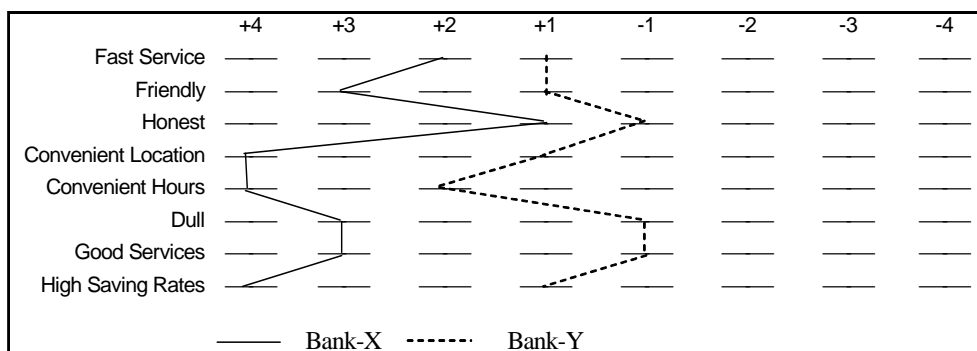
Format:

+5	+5
+4	+4
+3	+3
+2	+2
+1	+1
Friendly Personnel	Competitive Loan Rates
-1	-1
-2	-2
-3	-3
-4	-4
-5	-5

The following format shows an example of Stapel scale that illustrates respondents description on personnel banking of a bank.

	+4	+3	+2	+1	-1	-2	-3	-4
Fast Services	_____	_____	_____	_____	_____	_____	_____	_____
Friendly	_____	_____	_____	_____	_____	_____	_____	_____
Honest	_____	_____	_____	_____	_____	_____	_____	_____
Convenient Location	_____	_____	_____	_____	_____	_____	_____	_____
Convenient Hours	_____	_____	_____	_____	_____	_____	_____	_____
Dull	_____	_____	_____	_____	_____	_____	_____	_____
Good Services	_____	_____	_____	_____	_____	_____	_____	_____
High Saving Rates	_____	_____	_____	_____	_____	_____	_____	_____

Each respondent is asked to circle his opinion on a score against each phrase that describes the object. The final score of the respondent on a scale is the sum of their ratings for all the items. Also, the average score for each phrase is obtained by totaling the final score of all the respondents for that phrase divided by the number of respondents of the phrase. The following Figure shows the opinions of 100 respondents on two banks.



In the above example first the individual respondent's scores for each phrase that describes the selected bank are obtained and then the average scores of all

100 respondents for each phrase are plotted graphically. The maximum score possible for each bank is +32 and the minimum possible score for each brand is -32. In the example, Bank-X has score +24, and Bank-Y has score +3. From the scale we can identify which phrase needs improvement for each Bank.

The advantages and disadvantages of the Stapel scale are very similar to those for the Semantic differential scale. However, the Stapel scale tends to be easier to construct and administer, especially over telephone, since the Stapel scale does not call for the bipolar adjectives as does the Semantic differential scale. However, research on comparing the Stapel scale with Semantic differential scale suggests that the results of both the scales are largely the same.

5.6 SELECTION OF AN APPROPRIATE SCALING TECHNIQUE

In this unit, so far, you have learnt some of the important scaling techniques that are frequently used in attitudinal research for the measurement of attitudes. Each of these techniques has some advantages and disadvantages. Now you may ask which technique is more appropriate to use to measure attitudes. Virtually any technique can be used to measure the attitudes. But at the same time all techniques are not suitable for all purposes. As a general rule, you should use a scaling technique that will yield the highest level of information feasible in a given situation. Also, if possible the technique should permit you the use of a variety of statistical analysis. A number of issues decide the choice of scaling technique. Some significant issues are:

- 1) **Problem Definition and Statistical Analysis:** The Choice between ranking, sorting, or rating techniques is determined by the problem definition and the type of statistical analysis likely to be performed. For example, ranking provides only ordinal data that limits the use of statistical techniques.
- 2) **The Choice between Comparative and Non-comparative Scales:** Some times it is better to use a comparative scale rather than a non-comparative scale. Consider the following example:

How satisfied you are with the brand- X detergent that you are presently using?

<i>Completely satisfied</i>	<i>Somewhat satisfied</i>	<i>Neither satisfied nor dissatisfied</i>	<i>Somewhat dissatisfied</i>	<i>Completely dissatisfied</i>
---------------------------------	-------------------------------	---	----------------------------------	------------------------------------

This is a non-comparative scale since it deals with a single concept (the brand of a detergent). On the other hand, a comparative scale asks a respondent to rate a concept. For example, you may ask:

Which one of the following brands of detergent you prefer?

<i>Brand-X</i>	<i>Brand-Y</i>
----------------	----------------

In this example you are comparing one brand of detergent with another brand. Therefore, in many situations, comparative scaling presents 'the ideal situation' as a reference for comparison with actual situation.

- 3) **Type of Category Labels:** We have discussed different types of category labels used in constructing measurement scales such as verbal categories and numeric categories. Many researchers use verbal categories since they believe that these categories are understood well by the respondents. The maturity and the education level of the respondents influences this decision.

- 4) **Number of Categories:** While there is no single, optimal number of categories, traditional guidelines suggest that there should be between five and nine categories. Also, if a neutral or indifferent scale response is possible for at least some of the respondents, an odd number of categories should be used. However, the researcher must determine the number of meaningful positions that are best suited for a specific problem.
- 5) **Balanced versus Unbalanced Scale:** In general, the scale should be balanced to obtain objective data.
- 6) **Forced versus Nonforced Categories:** In situations where the respondents are expected to have no opinion, the accuracy of data may be improved by a non forced scale that provides a 'no opinion' category.

Self Assessment Exercises B

- 1) In paired comparison, the order in which the objects are presented may _____ results.
- 2) A researcher wants to measure consumer preference between 7 brands of bath soap and has decided to use the Paired comparisons scaling technique. How many pairs of brands will the researcher present the respondents?:

- 3) In a semantic differential scale there are 20 scale items. Should all the positive adjectives be on the left side and all the negative adjectives be on the right side. Can you explain?
.....
.....
.....
- 4) Indicate the type of scale used in the following examples.
 - a) Do you favour or oppose the return of the ruling party in the next elections?
(i) Favour (ii) Neutral (iii) Oppose
.....
 - b) Which one of the following pairs of companies would you like to invest your money?
 - i) MTNL - Reliance
 - ii) MTNL - BPL
 - iii) Reliance - BPL
 - c) Suppose Rs. 1,000 is given to you. How do you spend it?

Items	Amount (Rs.)
(a) Books	
(b) Clothes	
(c) Fast Food	
Total	1000

5.7 LET US SUM UP

There are four levels of measurements: nominal, ordinal, interval, and ratio. These constitute a hierarchy where the lower scale of measurement, nominal, has far fewer statistical applications than those further up this hierarchy of scales. Nominal scales yield data on categories; ordinal scales give sequences; interval scales begin to reveal the magnitude between points on the scale and ratio scales explain both order and the absolute distance between any two points on the scale.

The measurement scales, commonly used in marketing research, can be divided into two types; comparative and non-comparative scales. Comparative scales involve the respondent in signaling where there is a difference between two or more firms, brands, services, or other stimuli. The scales under this type are: (a) Paired Comparison, (b) Rank Order, (c) Constant Sum, and (d) Q-sort. Further, The non-comparative scales can be classified into: (a) Continuous Rating Scales and (b) Itemised Rating Scales. The Itemised Rating scales can further be classified into: (a) Likert Scale, (b) Semantic Differential Scale, and (c) Stapel Scale.

A number of scaling techniques are available for measurement of attitudes. There is no unique way that you can use to select a particular scaling technique for your research study. A number of issues, such as problem definition and statistical analysis, choice between comparative and non-comparative scales, type of category labels, number of categories etc., discussed in this unit should be considered before you arrive at a particular scaling technique.

5.8 KEY WORDS

Comparative Scales : In comparative scaling, the respondent is asked to compare one object with another.

Constant Sum Scale : In this scale, the respondents are asked to allocate a constant sum of units such as points, rupees, or chips among a set of stimulus objects with respect to some criterion.

Continuous Rating Scales : Here the respondents rate the objects by placing a mark at the appropriate position on a continuous line that runs from one extreme of the criterion variable to the other.

Itemised Rating Scales : Itemised rating scale is a scale having numbers or brief descriptions associated with each category.

Interval Scale : In this scale, the numbers are used to rank attributes such that numerically equal distances on the scale represent equal distances in the characteristic being measured.

Likert Scale : With the Likert scale, the respondents indicate their own attitudes by checking how strongly they agree or disagree with carefully worded statements that range from very positive to very negative towards the attitudinal object.

Measurement : Measurement is the process of observing and recording the observations that are collected as part of research.

Non-comparative Scales : In non-comparative scaling, respondents need only evaluate a single object.

Nominal Scale : In this scale, the different scores on a measurement simply indicate different categories.

Ordinal Scale : In this scale, the items are ranked according to whether they have more or less of a characteristic.

Paired Comparison Scale : This is a comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object according to some criterion.

Q-Sort Scale : This is a comparative scale that uses a rank order procedure to sort objects based on similarity with respect to some criterion.

Rank Order Scale : In this scale, the respondents are presented with several items simultaneously and asked to order or rank them according to some criterion.

Ratio Scale : Ratio scales permit the researcher to compare both differences in scores and relative magnitude of scores.

Scaling : Scaling is the assignment of objects to numbers or semantics according to a rule.

Semantic Differential Scale : This is a seven point rating scale with end points associated with bipolar labels (such as good and bad, complex and simple) that have semantic meaning.

Staple Scale : The Staple scale places a single adjective as a substitute for the Semantic differential when it is difficult to create pairs of bipolar adjectives.

5.9 ANSWERS TO SELF ASSESSMENT EXERCISES

- A.** 1) Interval scale does not have a fixed (absolute) zero point whereas ratio scale has a fixed zero point that allows us to construct a meaningful ratio.
- 2) Knowing the level of measurement helps in interpreting the data and performing statistical analysis of the data.
- 3) In nominal scale the assigned numbers have no arithmetic properties and act only as labels. The only statistical operation that can be performed on nominal scales is frequency count.
- 4) a) Nominal Scale, b) Ratio Scale, c) Ordinal Scale, d) Interval Scale.
- B.** 1) Bias
- 2) 21
- 3) No. Some of the positive adjectives may be placed on the left side and some on the right side. This prevents the respondent with positive (negative) attitude from simply checking the left (right) side without reading the description of the words.
- 4) a) Itemised rating scale, b) Paired comparison scale, c) Constant sum scale.

5.10 TERMINAL QUESTIONS

- 1) Discuss briefly different issues you consider for selecting an appropriate scaling technique for measuring attitudes.
- 2) What are the different levels of measurement? Explain any two of them.
- 3) How do you select an appropriate scaling technique for a research study? Explain the issues involved in it.

- 4) Discuss briefly the issues involved in attitude measurement.
- 5) Differentiate between ranking scales and rating scales. Which one of these scales is better for measuring attitudes?
- 6) In what type of situation is the Q-sort technique more appropriate?
- 7) Name any four situations in commerce where you can use the Likert scale.
- 8) Construct a Rank Order Scale to measure toothpaste preferences. Discuss its advantages and disadvantages.
- 9) Construct a Semantic differential scale to measure the experiences of respondents in using Brand-X of shaving cream (assume that all the respondents use that brand).

Note: These questions/exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the university for assessment. These are for your practice only.

5.11 FURTHER READING

The following text books may be used for more indepth study on the topics dealt with in this unit.

Aaker, David A. and George S. Day. (1983) *Marketing Research*, John Wiley, New York.

Bailey, Kenneth D. (1978) *Methods of Social Research*, The Free Press, New York.

Coombs, C.H.(1953) “Theory and Methods of Social Measurement”, in *Research Methods in the Behavioral Sciences*, eds. Feslinger, L. and Ratz, D., Holt, Rinehart and Winston.

Donald S. Tull and Gerald S. Albaum. (1973) *Survey Research: A Decisional Approach*, Index Educational Publishers, New York.

Meister, David. (1985) *Behavioural Analysis and Measurement Methods*, John Wiley, New York.

Rodger, Lesile W. (1984) *Statistics for Marketing*, McGraw-Hill (UK), London.